Structure and Expression of the *Lactococcus lactis* Gene for Phospho- β -galactosidase (*lacG*) in *Escherichia coli* and *L. lactis*

By WILLEM M. DE VOS¹* AND MIKE J. GASSON²

¹Molecular Genetics Group, Department of Biophysical Chemistry, Netherlands Institute for Dairy Research (NIZO), PO Box 20, 6710 BA Ede, The Netherlands

²Genetics and Microbiology Department, AFRC Institute of Food Research, Norwich Laboratory, Colney Lane, Norwich NR4 7UA, UK

(Received 5 January 1989; revised 9 March 1989; accepted 30 March 1989)

The *Lactococcus lactis* subsp. *lactis* 712 *lacG* gene encoding phospho- β -galactosidase was isolated from the lactose mini-plasmid pMG820 and cloned and expressed in *Escherichia coli* and *L. lactis*. The low phospho- β -galactosidase activity in *L. lactis* transformed with high-copy-number plasmids containing the *lacG* gene contrasted with the high activity found in *L. lactis* containing the original, low-copy-number lactose plasmid pMG820, and indicated that the original lactose promoter was absent from the cloned DNA. In *E. coli* the phospho- β -galactosidase could be overproduced using the strong inducible λ P_L promoter, which allowed a rapid purification of the active enzyme. The complete nucleotide sequence of the *L. lactis lacG* gene and its surrounding regions was determined. The deduced amino acid sequence was confirmed by comparison with the amino acid composition of the purified phospho- β -galactosidase and its amino-terminal sequence. This also allowed the exact positioning of the *lacG* gene and identification of its characteristic Gram-positive translation initiation signals. The homologous expression data and the sequence organization of the *L. lactis lacG* gene indicate that the gene is organized into a large lactose operon which contains an intergenic promoter located in an inverted repeat immediately preceding the *lacG* gene. The organization and sequence of the *L. lactis lacG* gene were compared with those of the highly homologous *lacG* gene from *Staphylococcus aureus*. A remarkable bias for leucine codons was observed in the *lacG* genes of these two species. Heterogramic homology was observed between the deduced amino acid sequence of the *L. lactis* phospho- β -galactosidase, that of the functionally analogous *E. coli* phospho- β -glucosidase, and that of an *Agrobacterium* β -glucosidase (cellobiase).

INTRODUCTION

Multicomponent phosphoenolpyruvate (PEP)-dependent sugar : phosphotransferase systems (PTS) are widespread in facultatively and obligately anaerobic Gram-negative and Gram-positive bacteria (Postma & Lengeler, 1985). The lactose-PTS, however, is almost exclusively found in Gram-positive bacteria, and the enzymology of the reactions involved has been studied extensively in *Staphylococcus aureus* (Hengstenberg, 1977; Hengstenberg & Deutscher, 1987). Lactose is transported by the membrane-bound Enzyme II-lactose and concomitantly phosphorylated. The resultant lactose 6-phosphate is subsequently cleaved into glucose and galactose 6-phosphate by a phospho- β -galactosidase (EC 3.2.1.85, Hengstenberg *et al.*, 1970). A similar mechanism of lactose transport and hydrolysis operates in the lactic streptococci *Lactococcus lactis* subsp. *lactis* and *L. lactis* subsp. *cremoris*, which are used in industrial dairy

Abbreviations: ONPG, *o*-nitrophenyl β -D-galactopyranoside; ONPG-6P, *o*-nitrophenyl β -D-galactopyranoside 6-phosphate; PTS, phosphotransferase system.

fermentations. It has been suggested that the presence of lactose-PTS is a prerequisite for the rapid homolactic fermentation of these starter bacteria (Lawrence *et al.*, 1976). Early studies of lactose metabolism implied that the lactose-specific PTS components and phospho- β -galactosidase were plasmid encoded in various lactococcal strains (McKay *et al.*, 1970; McKay, 1983). Genetic mapping experiments showed that these lactose genes were contained within an approximately 7 kb region on the *L. lactis* subsp. *lactis* 712 lactose mini-plasmid pMG820 (Gasson, 1983). These observations facilitated the cloning and expression in *Escherichia coli* of the pMG820-encoded phospho- β -galactosidase gene (Maeda & Gasson, 1986), previously designated *pbg* (de Vos, 1987) but in analogy with the terminology proposed for *S. aureus* (Breidt *et al.*, 1987) now described as *lacG*. Similar approaches have been used to clone the plasmid-encoded *lacG* genes from *Lactobacillus casei* (Lee *et al.*, 1982) and *L. lactis* subsp. *cremoris* (Inamine *et al.*, 1986) in *E. coli*. However, since cloning systems for these lactic acid bacteria have only recently been developed, studies on the expression of the *lacG* genes have been confined to heterologous hosts (LeBlanc, 1987; Chassy, 1987; de Vos, 1987). Here we describe the cloning of the pMG820 *lacG* gene in its homologous host, the study of its expression in *L. lactis* and *E. coli*, and the analysis of its gene product purified from an overproducing *E. coli*. The organization of the *L. lactis lacG* gene is further analysed by the determination of the nucleotide sequence of the gene and its surrounding regions, and its nucleotide and deduced amino acid sequences are compared with those of *lacG* genes from other Gram-positive bacteria. In addition, heterogramic homology is demonstrated between the Gram-positive phospho- β -galactosidases and Gram-negative counterparts, not only for the *E. coli* phospho- β -glucosidase (Schnetz *et al.*, 1987) as noted previously (Porter & Chassy, 1988), but also for an *Agrobacterium* β -glucosidase which has a high affinity for the non-phosphorylated substrate cellobiose (Wakarchuk *et al.*, 1988).

METHODS

Bacterial strains and plasmids. These are listed in Table 1.

Media, reagents and enzymes. Media based on M17 broth (Difco) containing 0.5% glucose or lactose (Terzaghi & Sandine, 1975) and L-broth (1% tryptone, 0.5% yeast extract, 0.5% NaCl) were used for the growth of *L. lactis* and *E. coli*, respectively. Ampicillin (Ap), chloramphenicol (Cm), *o*-nitrophenyl β -D-galactopyranoside (ONPG) and *o*-nitrophenyl β -D-galactopyranoside 6-phosphate (ONPG-6P) were purchased from Sigma. Unless stated otherwise, the antibiotics were used at final concentrations of 50 μ g ml⁻¹ and 10 μ g ml⁻¹, respectively. Restriction enzymes, T4 DNA polymerase, T4 DNA ligase, calf intestinal phosphatase and the Klenow fragment of *E. coli* DNA polymerase were obtained from Bethesda Research Laboratories and Boehringer Mannheim Biochemicals, and used under conditions as suggested by the suppliers.

DNA manipulations and transfer. Plasmid DNA was isolated essentially according to the alkaline lysis method of Birnboim & Doly (1979), using partially protoplasted cells which were obtained by incubating cells in 50 mM-Tris/HCl (pH 7.4) containing 25% (w/v) sucrose, 0.5 mM-EDTA and 10 mg lysozyme ml⁻¹ for 10 min at 4 °C (*E. coli*) or 37 °C (*L. lactis*). Further purification of plasmid DNA by CsCl/ethidium bromide density-gradient centrifugation, and all subsequent manipulations of DNA *in vitro* and in *E. coli*, were performed as described by Maniatis *et al.* (1982). *L. lactis* protoplasts were prepared from exponentially growing cultures as described previously (Gasson, 1980) and were transformed according to the method of Kondo & McKay (1984). Transformants were isolated on M17 agar plates containing 0.5 M-sucrose and 5 μ g chloramphenicol ml⁻¹ using a top agar overlay. In some cases use was made of protoplasts kept at -80 °C. Conjugation of pMG820 to *L. lactis* MG1363 containing pNZ12 derivatives was performed using the filter mating technique as described previously (Gasson & Davies, 1980) with counterselection for chloramphenicol resistance.

Determination of nucleotide sequences. DNA fragments were subcloned into single-stranded DNA vectors, either M13Mp8 or M13Mp9, using *E. coli* JM103 as host (Messing, 1983). Nucleotide sequences of both strands were determined by the dideoxy chain termination method (Sanger *et al.*, 1977).

Assembly and analysis of sequence data. Sequence data were assembled and analysed using the PC/GENE program version 5.01 (Genofit, Geneva). The facilities of the Netherlands CAOS/CAMM Center (University of Nijmegen) were used to screen the protein databases NBRF/PIR release 18.0, NBRF/NEW release 36.0 and SWISS-PROT release 9.0.

Phospho- β -galactosidase and β -galactosidase assays. Cell-free extracts of exponentially growing cultures were prepared by ultrasonication (*E. coli*) or obtained from osmotically shocked protoplasts (*L. lactis*). Phospho- β -galactosidase and β -galactosidase activities were determined at 37 °C using the chromogenic substrates ONPG-6P and ONPG, respectively (Maeda & Gasson, 1986; Miller, 1972). For rapid qualitative assays of phospho- β -

Table 1. *Bacterial strains and plasmids*

Strain or plasmid	Genotype or phenotype	Source or reference
<i>L. lactis</i>		
SH4109	NCDO 712 cured of ϕ T712 prophage, Lac ⁺	Gasson (1983)
MG1363	Plasmid-free derivative of SH4109, Lac ⁻	Gasson (1983)
MG1820	MG1363 containing pMG820, Lac ⁺	Maeda & Gasson (1986)
<i>E. coli</i> K12		
MC1061	<i>hsdR rpsL araD139</i> Δ (<i>ara-leu</i>)7697 Δ <i>lac74 galK</i>	Casadaban <i>et al.</i> (1980)
Δ H1 Δ <i>trp</i>	<i>lacZam</i> Δ (<i>bio-uvrB</i>) <i>trpEA2</i> (λ <i>Nam7 Nam53 cI857</i> Δ H1)	Remaut <i>et al.</i> (1981)
JM83	<i>ara</i> Δ (<i>lac-proAB</i>) <i>rpsL</i> ϕ 80 <i>dlacZ</i> Δ M15	Vieira & Messing (1982)
JM103	Δ (<i>lac-pro</i>) <i>supE thi strA endA sbcB15 hsdR4</i> (<i>F'</i> <i>traD36 proAB lacP</i> Δ M15)	Messing (1983)
<i>Plasmids</i>		
pMG820	Lactose mini-plasmid of <i>L. lactis</i>	Maeda & Gasson (1986)
pNZ12	Cm ^R Km ^R wide-host-range vector	de Vos (1987)
pNZ30	Cm ^R Km ^R pNZ12 derivative with <i>XhoI</i> fragment of pMG820 containing the <i>lacG</i> gene	This work
pNZ31	as pNZ30, opposite orientation	This work
pSM76	Ap ^R pAT153 with <i>ClaI</i> - <i>XhoI</i> fragment of pMG820 containing the <i>lacG</i> gene	Maeda & Gasson (1986)
pNZ32	Cm ^R Km ^R pNZ12 derivative with <i>ClaI</i> - <i>NaeI</i> fragment of pSM76 containing the <i>lacG</i> gene	This work
pNZ36	Cm ^R Km ^R pNZ12 derivative with <i>SstI</i> - <i>NaeI</i> fragment of pSM76 containing the <i>lacG</i> gene	This work
pPL603	Km ^R pUB110 derivative with promoterless <i>Bacillus pumilis</i> CAT-86 gene	Williams <i>et al.</i> (1981)
pNZ220	Km ^R pNZ12 derivative containing the promoterless CAT-86 gene	This work
pPLc28	Ap ^R pBR322 derivative with P _L promoter	Remaut <i>et al.</i> (1981)
pUC7	Ap ^R cloning vector	Vieira & Messing (1982)
pNZ316	Ap ^R pPLc28 derivative with the <i>lacG</i> gene cloned under control of the P _L promoter	This work
pMC1403	Ap ^R pBR322 derivative with the <i>E. coli lacZ</i> gene	Casadaban <i>et al.</i> (1980)
pNZ231	Ap ^R pMC1403 derivative with <i>SstI</i> - <i>ApaI</i> fragment of pNZ32, <i>lacZ</i> ⁺	This work

galactosidase activity, toluene-treated whole cells were used. Protein was measured according to Bradford (1976) with bovine serum albumin as standard.

Overproduction, purification and analysis of the lacG gene product. A 1.5 kb *HinfI*-*AhaIII* fragment (see Figs 1 and 2 for location) containing the structural *lacG* gene from which the *HinfI* site had been made blunt-ended by treatment with Klenow polymerase was cloned into the *HindIII* site of pUC7. The *lacG* gene was subsequently isolated as a *Bam*HI fragment and cloned under control of the λ P_L promoter in the *Bam*HI site of the expression vector pPLc28. Cells of *E. coli* Δ H1 Δ *trp* carrying the resulting plasmid pNZ316 were grown to an OD₆₀₀ of 0.5 at 28 °C, at which time the culture was divided into two equal parts. One part was further incubated at 28 °C and the other part at 42 °C to inactivate the thermosensitive λ repressor. Samples (1 ml) were taken after 1, 2 and 3 h and the cells were collected by centrifugation (12000 g, 2 min). The cell pellets were each then resuspended in 100 μ l of the lysis buffer [2% (w/v) sodium dodecyl sulphate (SDS), 5% (v/v) β -mercaptoethanol, 10% (v/v) glycerol, 50 mM-Tris/HCl, pH 6.8, and 0.001% bromophenol blue] and boiled for 3 min; 30 μ l of each sample was then electrophoresed on a 10% SDS-polyacrylamide gel, and the protein bands were visualized by staining with Coomassie blue (Laemmli, 1970). The *lacG* gene product was isolated from an overproducing *E. coli* containing pNZ316 as described by de Vos & Simons (1988), yielding approximately 15 mg purified phospho- β -galactosidase from one litre of induced culture. The purified protein was subjected to acid hydrolysis (6 M-HCl, 110 °C for 24 and 96 h, *in vacuo*) and the total amino acid composition was determined using an LKB type 4151 amino acid analyser (courtesy of Dr S. Visser and Mr P. van Rooijen). In addition, the N-terminal amino acid sequence of the purified phospho- β -galactosidase was determined with the aid of a gas-phase sequenator (Applied Biosystems) after transfer of the protein to glass sintered filter paper (Vandekerhoven *et al.*, 1985).

RESULTS

Cloning and expression of the lacG gene in E. coli and L. lactis

A 4.4 kb *XhoI* fragment of the *L. lactis* lactose mini-plasmid pMG820 containing the *lacG* gene (Maeda & Gasson, 1986; Fig. 1) was inserted into the unique *SalI* site of the high-copy-

Fig. 1. Physical maps of pMG820 and derivative plasmids containing the *L. lactis* phospho- β -galactosidase (P- β -GAL) gene. The cleavage sites of the restriction enzymes *Aha*III, *Cla*I, *Eco*RI, *Hind*III, *Sst*I and *Xho*I are shown. The hatched areas indicate regions of DNA deleted from plasmids pNZ32 and pNZ36. The open areas indicate the *Sal*I-*Nae*I fragment of pAT153 which was introduced during the manipulation. The vector (pNZ12)-located genes encoding chloramphenicol acetyltransferase (CAT), kanamycin nucleotidyltransferase (KNT) and control of plasmid replication (REPA) are indicated.

number wide-host-range lactic streptococcal vector pNZ12 and cloned into *E. coli* MC1061. Two plasmids differing with respect to the orientation of the inserted DNA (pNZ30 and pNZ31) were selected and transformed into the plasmid-free, lactose-deficient *L. lactis* MG1363. The resulting strains were not able to ferment lactose, indicating the absence of the other lactose-specific PTS enzymes.

The level of phospho- β -galactosidase activity in both hosts (Table 2) was affected by the orientation of the inserted fragment, indicating the involvement of a vector-located sequence in the expression of the *lacG* gene. A properly orientated consensus promoter sequence is known to be present in the sequence upstream of the *repA* gene in the replication origin of the vector pNZ12 (de Vos, 1987). The CAT-194 gene of pNZ12 was replaced by the promoterless *B. pumilis* CAT-86 gene derived from pPL603 (Williams *et al.*, 1981). When correctly orientated as in plasmid pNZ220, resistance to 25 μ g chloramphenicol ml⁻¹ was conferred on *E. coli*, demonstrating that transcriptional activity can proceed into the DNA cloned in pNZ12. In addition, by reducing the distance between the presumed location (Maeda & Gasson, 1986) of the *lacG* gene and the vector-located promoter, resulting in pNZ32 and pNZ36, an increase in *lacG* gene expression could be observed in both *E. coli* and *L. lactis* (Table 2). The smallest plasmid, pNZ36, gave the highest specific phospho- β -galactosidase activity. In spite of the approximately 10-fold higher copy number of pNZ36 as compared to that of pMG820

Table 2. Phospho- β -galactosidase activity of plasmid-containing *E. coli* and *L. lactis* strains

Plasmids were introduced into *E. coli* MC1061 or *L. lactis* MG1363. Cell-free extracts were prepared and specific activities determined as described in Methods.

Plasmid	Phospho- β -galactosidase activity [nmol min ⁻¹ (mg protein) ⁻¹]	
	<i>E. coli</i>	<i>L. lactis</i>
pMG820	–	130
pNZ30	76	12
pNZ31	168	46
pNZ32	186	50
pNZ36	375	75
pNZ12	<1	<1

–, Does not replicate in *E. coli*.

(unpublished observations), it appeared that *L. lactis* containing the original lactose plasmid pMG820 showed a higher phospho- β -galactosidase activity than *L. lactis* containing pNZ36 (Table 1). These results indicate either that the original *lacG* gene promoter is not present on the cloned DNA fragments or that it is present but requires an activating component not specified by the cloned DNA. To discriminate between these possibilities pMG820 was conjugated into *L. lactis* MG1363 containing pNZ32. Since no effect of gene dosage was observed in the resulting *L. lactis* strain containing both compatible plasmids, it was concluded that the 4.4 kb *Xho*I fragment contains the structural *lacG* gene but lacks the promoter which is responsible for the high expression of the *lacG* gene when located on pMG820.

Nucleotide sequence of the *L. lactis lacG* gene

The organization of the pMG820-derived *lacG* gene was further investigated by determining the nucleotide sequence of the insert present in smallest phospho- β -galactosidase-expressing plasmid, pNZ36. The nucleotide sequence contained one open reading frame with the size and orientation expected from the genetic analysis of the lactose region of pMG820 and from estimates of the molecular mass of purified phospho- β -galactosidase enzyme (Maeda & Gasson, 1986; de Vos & Simons, 1988; Fig. 2). The exact start of the *lacG* gene was concluded from the analysis of its expression product (see below), which allowed the identification of its ribosome-binding site and initiation codon (Fig. 2). The *lacG* gene extended from the ATG initiation codon (position 121–123) to the ochre termination codon TAA (position 1526–1528). The 1407 bp *lacG* gene encodes a 468-residue polypeptide with a molecular mass of 54078 Da.

The *L. lactis lacG* gene is preceded by a 84 bp sequence which contains the major part of a consensus Gram-positive promoter sequence (Graves & Rabinowitz, 1986; de Vos, 1987), with the canonical sequences TTGATT and TATTTG, which are spaced by 17 bp. The –10 region of this putative intergenic promoter is located in a DNA sequence which is capable of forming a stable stem-loop structure as is shown in Fig. 3. Support for the functionality of this promoter has been obtained from preliminary transcription analysis in *E. coli* (unpublished results). In addition, gene fusion studies have been performed to show the functionality of the postulated signals involved in the transcription and translation initiation of the *lacG* gene. In those experiments use was made of the unique *Apa*I site located at a suitable position in the *lacG* coding region (position 207 in Fig. 2). A 261 bp *Sst*I–*Apa*I fragment was isolated from pNZ32 and made blunt-ended by T4 DNA polymerase. Based on the sequence information of the *lacG* gene a correct in-frame *lacZ* gene fusion was expected when this blunt-ended DNA fragment was inserted in the expression probe vector pMC1403 linearized and made blunt-ended by digestion with *Eco*RI and *Bam*HI followed by a filling-in reaction using the Klenow fragment of *E. coli* DNA polymerase. Blue colonies were obtained on L-plates containing X-Gal and ampicillin when the resulting ligation mixture was transformed into *E. coli* MC1061, whereas

Fig. 2. Nucleotide sequence of the *L. lactis lacG* gene and its deduced amino acid sequence compared to that of *S. aureus* (only differences are indicated). Inverted repeats are indicated by arrows, and the ribosome-binding site by asterisks. The underlined amino acid sequence has been determined from the purified *L. lactis* phospho- β -galactosidase. Hyphens indicate nucleotides or amino acids not present in the corresponding sequence. The *HinfI* site unique to this sequence is located at position 81.

only white colonies were observed when the vector alone was ligated and transformed. The blue colonies appeared to contain slightly enlarged pMC1403 derivatives. A representative plasmid, designated pNZ231, was isolated from a blue colony and showed the expected 0.26 kb insert (the *EcoRI* and *BamHI* sites become reconstituted by the manipulations) which hybridized with

Fig. 3. Residues 5 to 133 of the *L. lactis lacG* gene sequence capable of forming a stem-loop structure. The -10 and -35 sequences of a putative promoter are indicated. The stop codons of the preceding open reading frame (*lacE*; de Vos *et al.*, unpublished results) are overlined, as are the out-of-phase stop codons near the *lacG* initiation codon, which is underlined. The ribosome-binding site of the *lacG* gene is indicated by asterisks.

pNZ32 DNA (not shown). These results confirm the sequence results and show that the *L. lactis lacG* transcription and translation initiation signals are functional in *E. coli*. Their combined efficiency, however, is not high since the specific activity of the β -galactosidase specified by pNZ31 [0.063 nmol min⁻¹ (mg protein)⁻¹] is relatively low, although significantly higher than that specified by pMC1403 [less than 0.008 nmol min⁻¹ (mg protein)⁻¹].

Overproduction and analysis of the phospho- β -galactosidase

To confirm the identity of the putative *lacG* gene product, its presence in an overproducing *E. coli* was demonstrated. To this purpose the *lacG* gene was excised as a *Hinf*I-*Aha*III fragment (for location of sites see Figs 1 and 2) and inserted into the expression vector pPLc28 so as to place the *lacG* gene under control of the λ P_L promoter. This results in plasmid pNZ316 (see Methods). Since *Hinf*I cuts in the inverted repeat preceding the *lacG* gene (position 81; see Figs 2 and 3), any possible effect of this secondary structure on the phospho- β -galactosidase expression is prevented. Fig. 4 compares the intracellular proteins of an *E. coli* host containing the thermosensitive λ repressor and carrying pNZ316 when grown at inducing and non-inducing temperatures. A protein with an apparent molecular mass of 58 kDa is overproduced under inducing conditions. Concomitantly with this induction, the phospho- β -galactosidase specific activity of the *E. coli* cells increases from approximately 3 nmol min⁻¹ (mg protein)⁻¹ (after 3 h at 28 °C) to 8900 nmol min⁻¹ (mg protein)⁻¹ (after 3 h at 42 °C).

The phospho- β -galactosidase was purified from an overproducing *E. coli* to apparent homogeneity using a one-step purification method (de Vos & Simons, 1988). The amino acid composition of this purified protein (Fig. 5) was in good agreement with that deduced from the

Fig. 4. Overproduction of the *L. lactis* phospho- β -galactosidase in *E. coli*. Cultures of *E. coli* Δ H1 Δ trp containing pNZ316 were either non-induced or induced, and grown for various times as described in Methods. Total protein was analysed by SDS-polyacrylamide gel electrophoresis. Lanes 1, 3 and 5, non-induced, grown for 1, 2 and 3 h, respectively; lanes 2, 4 and 6, induced, grown for 1, 2 and 3 h, respectively; lane P, approximately 2 μ g of purified (de Vos & Simons, 1988) phospho- β -galactosidase; lane M, molecular mass standards of 200, 116, 92.5, 66.2 and 45.0 kDa.

Table 3. Amino acid composition (mol mol^{-1}) of the purified phospho- β -galactosidase compared to that deduced from the sequence of the *lacG* gene

L. lactis phospho- β -galactosidase was purified from an overproducing *E. coli* strain containing pNZ316 (de Vos & Simons, 1988), hydrolysed and used to determine the amino acid composition as described in Methods. Asx, Asp and Asn; Glx, Glu and Gln; ND, not determined.

Amino acid	Protein	DNA	Amino acid	Protein	DNA
Ala	32.9	33	Lys	35.3	34
Arg	13.2	14	Met	6.8	8
Asx	61.5	62	Phe	25.1	25
Cys	1.5	2	Pro	20.5	20
Glx	50.1	47	Ser	15.4	15
Gly	38.1	38	Thr	19.7	20
His	17.8	18	Trp	ND	10
Ile	27.6	28	Tyr	36.4	37
Leu	29.4	28	Val	27.5	28

nucleotide sequence (Table 3). In addition, the sequence of the first ten amino-terminal amino acids of the purified protein was determined. The sequence Met-Thr-Lys-Thr-Leu-Pro-Lys-Asp-Phe-Ile was found, which completely matched that predicted from the nucleotide sequence (see underlined sequence in Fig. 2).

DISCUSSION

The cloning and expression of the *L. lactis* 712 *lacG* gene in both *E. coli* and *L. lactis* is described. Several lines of evidence indicate that the cloned fragments, which cover more than

2 kb of DNA upstream of the *lacG* gene, do not contain the original lactose promoter, and suggest that the *lacG* gene is part of an operon. The main evidence is the low specific phospho- β -galactosidase activity found in *L. lactis* carrying the *lacG* gene cloned on a high-copy-number vector, which contrasts with the high activity found in *L. lactis* containing the original, low-copy-number lactose mini-plasmid pMG820 (Table 2).

Sequence analysis of approximately 2.5 kb upstream of the *lacG* gene shows the presence of several open reading frames, including those encoding the lactose-specific PTS enzymes Factor III (*lacF*) and Enzyme II (*lacE*), which, like those in *S. aureus* (Breidt & Stewart, 1987), are organized in an operon structure with small intercistronic regions (W. M. de Vos and others, unpublished). Further support for an operon structure of the lactose genes comes from RNA studies which show the presence of a long (approximately 6 kb) mRNA specific for the *lacG* gene (W. M. de Vos and others, unpublished observations) and from the isolation of a strong functional promoter located approximately 5 kb upstream of the phospho- β -galactosidase gene (P. H. Anderson, M. J. Gasson & W. M. de Vos, unpublished data).

An operon structure of the lactose genes in *L. lactis* implies that the promoter located immediately upstream of the *lacG* gene is an intergenic promoter. Evidence for its functionality has been derived from *lacZ* fusion studies in *E. coli* and from the observation that low, but significant, expression of the *lacG* gene is observed in *L. lactis* containing pNZ30. In this plasmid the *lacG* gene is most likely not under control of a vector-located promoter since the chloramphenicol resistance (CAT-194) gene (Fig. 1) is followed by a functional rho-independent terminator (Horinuchi & Weisblum, 1982). In the complete lactose region, it is therefore possible that expression of the phospho- β -galactosidase gene is under dual control of a strong, presumably inducible *lac* promoter, located far upstream of the *lacG* gene, and a weak, intergenic promoter adjacent to the *lacG* gene. If so, this intergenic promoter may ensure a low and possibly constitutive level of phospho- β -galactosidase in the cell. An advantage of this situation could be the degradation of traces of lactose 6-phosphate, which may be toxic to the cells. In addition, the degradation of lactose 6-phosphate results in the formation of galactose 6-phosphate which is assumed to be the inducer of the lactose genes in *L. lactis* (McKay, 1983).

The -10 region of the intergenic promoter is located within a region of dyad symmetry (Fig. 3). The function of this inverted repeat, which is followed by a stretch of T residues, is not clear. By analogy with the functions of dyad symmetries found in other PTS operons (Schnetz *et al.*, 1987; Mahadevan & Wright, 1987; Schnetz & Rak, 1988), it may be involved in the regulation of *lacG* gene expression and may act (i) as a regulation site for the intergenic promoter, (ii) as an intergenic terminator, or (iii) as a sequence involved in stabilization of the *lac* operon mRNA. Alternatively, folding of the *lac* mRNA in such a stable loop [ΔG value of $-18.2 \text{ kcal mol}^{-1}$ ($-76.1 \text{ kJ mol}^{-1}$), as calculated following the rules of Tinoco *et al.* (1973)] could reduce the distance between the end of the *lacE* gene and the start of the *lacG* gene.

The *L. lactis lacG* sequence was compared with that of the *S. aureus lacG* sequence (Breidt & Stewart, 1987). A remarkable degree of homology, both between the nucleotide sequences and between the deduced amino acid sequences (80% and 82% identity, respectively), was found (Fig. 2). The *S. aureus lacG* gene is 6 bp longer than the *L. lactis* counterpart due to an insertion at position 1273. In both sequences a consensus Gram-positive ribosome-binding site (McLaughlin *et al.*, 1981) is located immediately upstream of the initiation codon. In *L. lactis* this is the sequence GAAAGGA, which is highly complementary to the 5' end of the *L. lactis* 16S rRNA [Ludwig *et al.*, 1985; $\Delta G -14 \text{ kcal mol}^{-1}$ ($-58.6 \text{ kJ mol}^{-1}$), as calculated following the rules of Tinoco *et al.* (1973)] and is 9 bp away from the ATG initiation codon.

The sequences flanking the *L. lactis lacG* gene do not share homology with sequences located in the immediate vicinity of the *S. aureus lacG* gene (Breidt & Stewart, 1987). Whereas the *S. aureus lacG* coding sequence is followed by a rho-independent terminator, it appears that no such structure is found downstream of the *L. lactis lacG* gene, which, in contrast, is followed by an intercistronic region without significant secondary structure and another open reading frame (de Vos & Simons, 1988). Both *lacG* genes are preceded by another gene *lacE* (encoding Enzyme II-lactose; Breidt & Stewart, 1987; de Vos & Simons, 1988), the termination codons of which are shown in Fig. 2. However, in the *S. aureus* lactose region, the intergenic promoter and the

P- β -GAL	MTKT	LPKDFIFGGA	TAAVQAEGAT	HTDGKGPVAV	DKYLE-----
β -GLU	MTDPNTLAAR	FPGDFLPGVA	TASFQIEGST	KADGRKPSIW	DAFCN-----
P- β -GLU	MKA	FPETFLWGGA	TAANQVEGAW	QEDGKGISTS	DLQPHGVMMGK
P- β -GAL	-----DNYWY	TAEPASDFYH	KYPVDLELAE	EYGVNGIRIS	IAWSRIFFTG
β -GLU	MPGHVFGGRHN	-GDIACDHYN	RWEEDLDLIK	EMGVEAYRFS	LAWPRIIPDG
P- β -GLU	MEPRILGKEN	IKDVAIDFYH	RYPEDIALFA	EMGFTCLRIS	IAWARIFPQG
P- β -GAL	-YGEVNEKGV	EFYHKLFAEC	HKRHVEFPVT	LHHFDTPEAL	HSN-GDFLNR
β -GLU	-FGPINEKGL	DFYDRLVDGC	KARGIKTYAT	LYHWDLPLTL	MGD-GGWASR
P- β -GLU	DEVEPNNEAGL	AFYDRLFDEM	AQAGIKPLVT	LSHYEMPYGL	VKNYGGWANR
P- β -GAL	ENIEHFIDYA	AFCFEFPE-	VNYWTFNEI	GPIGDGQYLV	GKFFPGIKYD
β -GLU	STAHAFQRYA	KTVMARLGDR	LDAVATFNPE	WCAVWLSHLY	GVHAPGERN-
P- β -GLU	AVIGHFEHYA	RTVFTRYQHK	VALWLTFNEI	NMSL-HAPFT	GVGLAEESG-
P- β -GAL	LAKVFQSHEN	MMVSHARAVK	LYKDKGYKGE	IGVVHALPTK	YPYPDENPAD
β -GLU	MEAALAMMH	INLAHGFGVE	ASRHVAPKVP	VGLVLNAHSA	IP-ASDGEAD
P- β -GLU	EAEVYQAIHH	QLVASARAVK	ACHSLLEPAK	IGNMLLGGLV	YP-LTCQPQD
P- β -GAL	-VRAAELEDI	IHNKFILDAT	YLGHYSDKTM	EGVNHILAEN	GGELDLRDED
β -GLU	-LKAAERAFAQ	FHNGAFFDPV	FKGEYPAEMM	E-----	ALGDRMPVVE
P- β -GLU	MLQAMEENRR	W--MFFGDVQ	ARGQYPGYMQ	R----FFRDH	NITIEMTESD
P- β -GAL	FQALDAAKDL	NDFLGINYMM	SDWMQAFDGE	TEIIHNGKGE	KGSSKYQIKG
β -GLU	AEDLGIISQK	LDWWGLNYTT	PMRVADDATP	GVEFFA----	-----
P- β -GLU	AEDL---KHT	VDFISFSYMM	TGCVSHDESI	NKNAQG----	-----
P- β -GAL	VGRRVAPDYV	PRTDWDWIY	PEGLYDQIMR	VKNDYPNYKK	IYITENGLGY
β -GLU	TMPAPAVSDV	K-TDIGWEVY	APALHTLVET	LYERY-DLPE	CYITENGACY
P- β -GLU	ILNMIPNPHL	KSSEWGWOID	PVGLRVLLNT	LWDRY--QKP	LFIVENGLGA
P- β -GAL	KDEFV-DNTV	YDDGRIDYVK	QHLEVLSDAI	ADGANVKGYF	IWSLMDVFSW
β -GLU	NMGVE-NGEV	NDQPRLDYYA	EHLGIVADLI	RDGYPMRGYF	AWSLMDNFEW
P- β -GLU	KDSVEADGSI	QDDYRIAYLN	DHLVQVNEAI	ADGVDIMGYT	SWGPIDLVSA
P- β -GAL	-SNGYEKRYG	LFYVDFDTQE	RYPKSAHWY	KKLAETQVIE	
β -GLU	-AEGYRMRFG	LVHVDYETQV	RTVKNSGKWY	SALA	
P- β -GLU	SHSQMSKRYG	FIYVDRDDNG	EGSLTRTRKK	SFRMVCAEVI	KTRGLSLKKI TIKAP

Fig. 5. Homologies in the deduced amino acid sequences of the *L. lactis* phospho- β -galactosidase (P- β -GAL; 468 amino acids, this study), an *Agrobacterium* β -glucosidase (β -GLU; 458 amino acids, Wakarchuck *et al.*, 1988) and the *E. coli* phospho- β -glucosidase (P- β -GLU; 471 amino acids, Schnetz *et al.*, 1987). The amino acid sequences are given in the one-letter code and have been aligned by introducing gaps (hyphens) to maximize identity. Identical amino acids are shown in bold face. Percentage identity for pairwise comparisons were 29.1% for *L. lactis* P- β -GAL and *Agrobacterium* β -GLU, 28.3% for *E. coli* P- β -GLU and *Agrobacterium* β -GLU, and 28.2% for *L. lactis* P- β -GAL and *E. coli* P- β -GLU. Dots above the P- β -GAL sequence refer to amino acids which are identical in the three phospho- β -galactosidase sequences from *L. lactis*, *S. aureus* and *Lactobacillus casei* (Porter & Chassy, 1988).

inverted repeat of the *L. lactis* operon are absent. This may reflect the differences in the regulation of *lacG* gene expression, which is tight and under repressor control in *S. aureus* (Oskouian & Stewart, 1987) but relatively relaxed in *L. lactis* (Maeda & Gasson, 1986).

The presence of the NH₂-terminal methionine was demonstrated in the *L. lactis* phospho- β -galactosidase protein. This contrasts with the phospho- β -galactosidase purified from *S. aureus*,

Table 4. Leucine codon usage in *L. lactis* and *S. aureus*

Codon	<i>L. lactis</i>	<i>S. aureus</i>
UUA	5	21
UUG	8	3
CUU	12	1
CUC	1	0
CUA	2	1
CUG	0	0
	<hr/> 26	<hr/> 26

which starts with a threonine residue (Hengstenberg *et al.*, 1970). The *L. lactis* lacG product expressed in *E. coli* is clearly not processed at its amino-terminus. Whether this is a result of the overproduction in *E. coli* (Sharma, 1986) or an inherent property of the *L. lactis* phospho- β -galactosidase remains to be determined. In any case, these results on the NH₂-terminal sequence provide the first confirmation of the functionality of a presumed lactic streptococcal translation initiation sequence (de Vos, 1987).

The lacG gene investigated in this study has been cloned, via the lactose mini-plasmid pMG820, from the large, 56.5 kb plasmid pLP712 present in the multiplasmid host *L. lactis* 712 (Gasson, 1983). The nucleotide sequence is completely identical to the sequence of the lacG gene of the *L. lactis* strain Z268, which was independently determined (A. Mercenier, personal communication; Boizet *et al.*, 1988). In the latter strain the lacG gene is located on a large 54 kb plasmid, pUCL13, which greatly differs from pLP712 in its restriction map (Boizet *et al.*, 1988; Gasson, 1983). These results suggest that plasmid evolution in the mesophilic lactic streptococci is of recent date.

The *L. lactis* lacG coding region is highly homologous to that of *S. aureus* both at the protein level and at the nucleotide level. Recently, the *Lactobacillus casei* lacG nucleotide sequence has also been determined and included in a comparison of the deduced phospho- β -galactosidase sequences of *L. lactis*, *S. aureus* and *Lb. casei* (Porter & Chassy, 1988). Considering the location of the amino-terminal sequence of the *L. lactis* lacG gene product reported here, all three phospho- β -galactosidases have a similar length and show a degree of homology exceeding 50%, with the *L. lactis*-*S. aureus* pair being the most homologous. Also the codon usage for the *L. lactis* and *S. aureus* lacG genes was found to be very similar, with a predominance of adenine or uracil in the third position, reflecting the relatively low GC content of those genes and species (Porter & Chassy, 1988). However, a careful comparison of the codon usage showed a significantly different codon usage for the Leu codon (see Table 4). The preference for a CUU codon in *L. lactis* is not shared by *S. aureus*, where a strong preference for the UUA codon is found. In addition, the optimal *E. coli* Leu codon CUG (Aota *et al.*, 1988) is not used at all in either phospho- β -galactosidase gene, and this apparently does not affect the overproduction in *E. coli* of the *L. lactis* phospho- β -galactosidase (Fig. 4). Inspection of the Leu codons in the phospho- β -galactosidase sequences (Fig. 2) shows nine instances (amino acid positions 5, 38, 123, 237, 252, 274, 284, 294, 423) where a *S. aureus* UUA Leu codon has been changed into a CUU Leu codon in *L. lactis* as a result of two separate mutations. This is a relatively unlikely event, not observed for other codons, and suggests that there has been a strong selective pressure both at the amino acid level (Leu amino acid) and the nucleotide level (Leu codon usage) in these Gram-positive bacteria.

The high overall homology between the phospho- β -galactosidases of the lactic acid bacteria and *S. aureus* is not expected from the phylogenetic position of these species within the *Clostridium* tree of the Gram-positive bacteria (Stackebrandt & Teuber, 1988) and has been used to suggest a recent and rapid horizontal evolution of these lactose genes (Porter & Chassy, 1988). The deduced amino acid sequence of the lacG gene shows no homology with the *E. coli* β -galactosidase but, in contrast, a remarkably high homology with that of the *E. coli* phospho- β -glucosidase (*bglB*) gene (Schnetz *et al.*, 1987, Porter & Chassy, 1988; Fig. 5). This is the second case of a PTS-related enzyme showing heterogramic homology, the other example being the

Bacillus subtilis Enzyme II-sucrose, which is homologous to Enzymes II of various enteric species, including that of the *E. coli* *bgl* operon (Fouet *et al.*, 1987). A further discussion on the implication of the observed homologies for the evolution of these functionally homologous phospho- β -glycosidases has been presented elsewhere (Porter & Chassy, 1988; de Vos & Simons, 1988). Unexpectedly, a similar high degree of homology is found between the deduced amino acid sequence of the two phospho- β -glycosidases and that of a β -glucosidase from *Agrobacterium* sp. ATCC 21400 (Wakarchuk *et al.*, 1988), which is specific for cellobiose (Fig. 5). Regions of identity and high homology are present along the whole polypeptide sequence and include amino acids which are identical in the three Gram-positive phospho- β -galactosidases (Fig. 5). These homologous sequences may constitute domains which are involved in the common function of these enzymes, which all hydrolyse β -glycosidic linkages. In addition, the few regions of homology shared only by the two phospho- β -glycosidases may be specific for the binding and/or cleavage of the phosphorylated substrates. The high heterogramic homology between the *L. lactis* and *E. coli* phospho- β -glycosidases and the *Agrobacterium* β -glucosidase suggests a common ancestor for those PTS- and non-PTS-related enzymes. Partial homology between PTS- and non-PTS-related enzymes has been shown before for the *B. subtilis* sucrose-6-phosphate hydrolase. This PTS-related enzyme is partially homologous to yeast invertase (Fouet *et al.*, 1986), and contains, as does the homologous sucrose-6-phosphate hydrolase from *Streptococcus mutans*, a sucrose box which is conserved in the sequence of fructosyltransferase and levansucrase (Sato & Kuramitsu, 1988).

The authors wish to thank Herma Buys and Rutger van Rooijen for expert technical assistance. The help of Peter van Rooijen and Ser Visser in the amino acid analysis is gratefully acknowledged. We are grateful to Annick Mercenier for communicating and discussing results prior to publication. This work was performed under contracts BAP-0011-NL and BAP-0009-UK of the Biotechnology Action Programme of the Commission of the European Communities.

REFERENCES

- AOTA, S., GOJOBORI, T., ISIBASHI, F., MARUYAMA, T. & IKEMURA, T. (1988). Codon usage tabulated from the GenBank Genetic Sequence Data. *Nucleic Acids Research* **16**, r315-r403.
- BIRNBOIM, H. C. & DOLY, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* **7**, 1513-1519.
- BOIZET, B., VILLEVAL, D., SLOS, P., NOVEL, M., NOVEL, G. & MERCENIER, A. (1988). Isolation and structural analysis of the phospho- β -galactosidase gene from *Streptococcus lactis* Z268. *Gene* **62**, 249-261.
- BRADFORD, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein, utilizing the principle of protein-dye binding. *Analytical Biochemistry* **12**, 248-254.
- BREIDT, F., JR & STEWART, G. C. (1987). Nucleotide sequence and deduced amino acid sequence of the *Staphylococcus aureus* phospho- β -galactosidase gene. *Applied and Environmental Microbiology* **53**, 969-973.
- BREIDT, F., JR, HENGSTENBERG, W., FINKELDEI, U. & STEWART, G. C. (1987). Identification of the genes for the lactose-specific components of the phosphotransferase system in the *lac* operon of *Staphylococcus aureus*. *Journal of Biological Chemistry* **262**, 16444-16449.
- CASADABAN, M. J., CHOU, J. & COHEN, S. N. (1980). *In vitro* gene fusions that join an enzymatically active β -galactosidase segment to amino-terminal fragments of exogenous proteins: *E. coli* plasmid vectors for the detection and cloning of translational signals. *Journal of Bacteriology* **143**, 971-980.
- CHASSY, B. M. (1987). Prospects for genetic manipulation in lactobacilli. *FEMS Microbiology Letters* **46**, 297-312.
- FOUET, A., KLIER, A. & RAPOPORT, G. (1986). Nucleotide sequence of the sucrase gene of *Bacillus subtilis*. *Gene* **45**, 221-225.
- FOUET, A., ARNAUD, M., KLIER, A. & RAPOPORT, G. (1987). *Bacillus subtilis* sucrose-specific enzyme II of the phosphotransferase system: expression in *Escherichia coli* and homology to enzymes II from enteric bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 8773-8777.
- GASSON, M. J. (1980). Production, regeneration and fusion of protoplasts in lactic streptococci. *FEMS Microbiology Letters* **9**, 99-102.
- GASSON, M. J. (1983). Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic streptococci after protoplast-induced curing. *Journal of Bacteriology* **154**, 1-9.
- GASSON, M. J. & DAVIES, F. L. (1980). High-frequency conjugation associated with *Streptococcus lactis* donor cell aggregation. *Journal of Bacteriology* **143**, 1260-1264.
- GRAVES, M. C. & RABINOWITZ, J. C. (1986). *In vivo* and *in vitro* transcription of the *Clostridium pasteurianum* ferredoxin gene. Evidence for extended promoter elements in Gram-positive organisms. *Journal of Biological Chemistry* **261**, 11409-11415.
- HENGSTENBERG, W. (1977). Enzymology of carbohydrate transport in bacteria. *Current Topics in Microbiology and Immunology* **77**, 97-126.

- HENGSTENBERG, W. & DEUTSCHER, J. (1987). Chemical characterization of proteins of the phosphoenolpyruvate-dependent phosphotransferase system. In *Sugar Transport and Metabolism in Gram-positive Bacteria*, pp. 215–234. Edited by J. Reizer & A. Peterkofsky. Chichester: Ellis Horwood.
- HENGSTENBERG, W., PENBERTHY, W. K. & MORSE, M. L. (1970). Purification of the staphylococcal 6-phospho- β -galactosidase. *European Journal of Biochemistry* **14**, 27–32.
- HORINOCHI, S. & WEISBLUM, B. (1982). Nucleotide sequence of and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. *Journal of Bacteriology* **150**, 815–825.
- INAMINE, J. M., LEE, L. N. & LEBLANC, D. J. (1986). Molecular and genetic characterization of lactose metabolic genes of *Streptococcus cremoris*. *Journal of Bacteriology* **167**, 855–862.
- KONDO, J. K. & MCKAY, L. L. (1984). Plasmid transformation of *Streptococcus lactis* protoplasts: optimization and use in molecular cloning. *Applied and Environmental Microbiology* **48**, 252–259.
- LAEMMLI, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature, London* **227**, 680–685.
- LAWRENCE, C. R., THOMAS, T. D. & TERZAGHI, B. E. (1976). Reviews of the progress of dairy science: cheese starters. *Journal of Dairy Research* **43**, 141–193.
- LEBLANC, D. J. (1987). Genetic and molecular aspects of carbohydrate metabolism in streptococci. In *Sugar Transport and Metabolism in Gram-positive Bacteria*, pp. 386–418. Edited by J. Reizer & A. Peterkofsky. Chichester: Ellis Horwood.
- LEE, L.-J., HANSEN, J. B., JAGUSZTYN-KRYNICKA, E. K. & CHASSY, B. M. (1982). Cloning and expression of the β -D-phosphogalactoside galactohydrolase gene of *Lactobacillus casei* in *Escherichia coli* K-12. *Journal of Bacteriology* **152**, 1138–1146.
- LUDWIG, W., SEEWALDT, E., KILPPER-BALZ, R., SCHLEIFER, K. H., MAGRUM, L., WOESE, C. R., FOX, G. E. & STACKEBRANDT, E. (1985). The phylogenetic position of *Streptococcus* and *Enterococcus*. *Journal of General Microbiology* **131**, 543–551.
- MAEDA, S. & GASSON, M. J. (1986). Cloning, expression and location of the *Streptococcus lactis* gene for phospho- β -D-galactosidase. *Journal of General Microbiology* **132**, 331–340.
- MAHADEVAN, S. & WRIGHT, A. (1987). A bacterial gene involved in transcription antitermination: regulation at a rho-independent terminator in the *bgl* operon of *E. coli*. *Cell* **50**, 485–494.
- MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982). *Molecular Cloning, a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- MCKAY, L. L. (1983). Functional properties of plasmids in lactic streptococci. *Antonie van Leeuwenhoek* **49**, 259–274.
- MCKAY, L. L., MILLER, A., III, SANDINE, W. E. & ELLIKER, P. E. (1970). Mechanism of lactose utilization by lactic acid streptococci: enzymatic and genetic analysis. *Journal of Bacteriology* **102**, 804–809.
- MCLAUGHLIN, J. R., MURRAY, C. L. & RABINOWITZ, J. C. (1981). Unique features in ribosome binding site sequences of the Gram-positive *Staphylococcus aureus* β -lactamase gene. *Journal of Biological Chemistry* **255**, 8819–8830.
- MESSING, J. (1983). New M13 vectors for cloning. *Methods in Enzymology* **101**, 20–78.
- MILLER, J. H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- OSKOUIAN, B. & STEWART, G. C. (1987). Cloning and characterization of the repressor gene of the *Staphylococcus aureus* lactose operon. *Journal of Bacteriology* **169**, 5459–5465.
- PORTER, E. V. & CHASSY, B. M. (1988). Nucleotide sequence of the β -D-phosphogalactoside galactohydrolase gene from *Lactobacillus casei*: comparison to β -D-phosphogalactoside galactohydrolase genes from other Gram-positive organisms. *Gene* **62**, 263–276.
- POSTMA, P. W. & LENGELER, J. W. (1985). Phosphoenolpyruvate:carbohydrate phosphotransferase system of bacteria. *Microbiological Reviews* **49**, 232–269.
- REMAUT, E., STANSENS, P. & FIERIS, W. (1981). Plasmid vectors for high-efficiency expression controlled by the P_L promoter of coliphage lambda. *Gene* **15**, 81–93.
- SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463–5467.
- SATO, Y. & KURAMITSU, H. K. (1988). Sequence analysis of the *Streptococcus mutans* *scrB* gene. *Infection and Immunity* **56**, 1956–1960.
- SCHNETZ, K. & RAK, B. (1988). Regulation of the *bgl* operon of *Escherichia coli* by transcriptional anti-termination. *EMBO Journal* **7**, 3271–3277.
- SCHNETZ, K., TOLOCZYKI, C. & RAK, B. (1987). β -Glucoside (*bgl*) operon of *Escherichia coli* K12: nucleotide sequence, genetic organization, and possible evolutionary relationship to regulatory components of two *Bacillus subtilis* genes. *Journal of Bacteriology* **169**, 2579–2590.
- SHARMA, S. K. (1986). On the recovery of genetically engineered proteins from *Escherichia coli*. *Separation Science & Technology* **21**, 701–726.
- STACKEBRANDT, E. & TEUBER, M. (1988). Molecular taxonomy and phylogenetic position of lactic acid bacteria. *Biochimie* **70**, 317–324.
- TERZAGHI, B. K. & SANDINE, N. R. (1975). Improved medium for lactic streptococci, and their bacteriophages. *Applied and Environmental Microbiology* **29**, 807–813.
- TINOCO, I., BORER, P. N., DENGLE, B., LEVINE, M. D., UHLENBECK, O. C., CROTHERS, D. M. & GRALLA, J. (1973). Improved estimation of secondary structure in ribonucleic acids. *Nature New Biology* **246**, 40–41.
- VANDEKERKHOVEN, J., BAUW, G., PUYE, M., VAN DAMME, J. & VAN MONTAGU, M. (1985). Protein blotting on polybrene-coated glass fiber sheets. *European Journal of Biochemistry* **152**, 9–19.
- VIEIRA, J. & MESSING, J. (1982). The pUC plasmids, an M13mp7 derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**, 259–268.
- DE VOS, W. M. (1987). Gene cloning and expression in lactic streptococci. *FEMS Microbiology Reviews* **46**, 281–295.
- DE VOS, W. M. & SIMONS, G. (1988). Molecular cloning of lactose genes in dairy lactic streptococci: the

- phospho- β -galactosidase and β -galactosidase genes and their expression products. *Biochimie* **70**, 461–473.
- WAKARCHUK, W. W., GREENBERG, N. M., KILBURN, D. G., MILLER, R. C., JR & WARREN, R. A. J. (1988). Structure and transcription analysis of the gene encoding a cellobiase from *Agrobacterium* sp. strain ATCC 21400. *Journal of Bacteriology* **170**, 301–307.
- WILLIAMS, D. M., DUVAL, E. J. & LOVETT, P. S. (1981). Cloning restriction fragments that promote expression of a gene in *Bacillus subtilis*. *Journal of Bacteriology* **148**, 1162–1165.