Arch Microbiol (2000) 173 : 383–389
Digital Object Identifier (DOI) 10.1007/s002030000159

ORIGINAL PAPER

Andrea McCracken · Mark S. Turner · Phil Giffard
Louise M. Hafner · Peter Timms

Analysis of promoter sequences from *Lactobacillus* and *Lactococcus* and their activity in several *Lactobacillus* species

Received: 27 September 1999 / Revised: 8 March 2000 / Accepted: 8 March 2000 / Published online: 6 April 2000

Abstract Promoter-active fragments were isolated from the genome of the probiotic organism *Lactobacillus rhamnosus* strain GG using the promoter-probe vector pNZ272. These promoter elements, together with a promoter fragment isolated from the vaginal strain *Lactobacillus fermentum* BR11 and two previously defined promoters (*Lactococcus lactis lacA* and *Lactobacillus acidophilus* ATCC 4356 *slpA*), were introduced into three strains of *Lactobacillus*. Primer-extension analysis was used to map the transcriptional start site for each promoter. All promoter fragments tested were functional in each of the three lactobacilli and a purine residue was used to initiate transcription in most cases. The promoter elements encompassed a 52- to 1140-fold range in promoter activity depending on the host strain. *Lactobacillus* promoters were further examined by surveying previously mapped sequences for conserved base positions. The *Lactobacillus* hexamer regions (–35: **T**Tgaca and –10: **TA**tAA**T**) closely resembled those of *Escherichia coli* and *Bacillus subtilis*, with the highest degree of agreement at the –10 hexamer. The TG dinucleotide upstream of the –10 hexamer was conserved in 26% of *Lactobacillus* promoters studied, but conservation rates differed between species. The region upstream of the –35 hexamer of *Lactobacillus* promoters showed conservation with the bacterial UP element.

Key words *Lactobacillus* · Promoter · Consensus sequence · Gene expression ·

Introduction

Transcription and its regulation have been studied extensively in gram-negative bacteria, particularly *Escherichia coli*. However, with the exception of *Bacillus subtilis*, the study of transcription has been largely neglected in gram-positive organisms. The primary σ-factors of *E. coli* and *B. subtilis* make site-specific contacts with key bases harboured within two conserved hexamers located at approximately –35 (TTGACA) and –10 (TATAAT) with respect to the transcriptional start site (Helmann 1995; Lisser and Margalit 1993). The length of the spacer region connecting the two hexamers is also conserved (17±1 bp). Several other elements within close proximity of these hexamers have also been shown to influence transcription in these organisms. These include the UP element, an AT-rich sequence upstream of the –35 hexamer which is contacted by the C-terminal domain of the RNA polymerase α-subunit (Ross et al. 1993) and a TG motif located upstream of the –10 hexamer (Harley and Reynolds 1987).

Transcription has received little attention in the important gram-positive organism *Lactobacillus*. No genes encoding σ-factors have been identified and it is not yet clear what constitutes an optimal promoter region in this genus. A survey of 11 *Lactobacillus* promoters with mapped transcriptional start sites identified typical –35 and –10 sequences which, when aligned, resembled those of both *E. coli* and *B. subtilis* promoters (Mercenier et al. 1994; Pouwels and Leer 1993). Interestingly, strong consensus promoters of foreign genes are not always transcribed in lactobacilli (Natori et al. 1988). It is likely, therefore, that additional sequences outside the traditional –35 and –10 boxes are important in defining a *Lactobacillus* promoter region.

The development of *Lactobacillus* strains expressing heterologous proteins at sufficient levels for application in industry and public health has been largely hindered by the lack of detailed knowledge of gene expression control in these organisms. The current study describes an analysis of promoters in *Lactobacillus*. We isolated seven promoter-active fragments from the genome of the important probiotic strain *Lactobacillus rhamnosus* GG, as well as a single promoter fragment from the vaginal strain *Lactobacillus fermentum* BR11, and evaluated the ability of each fragment to drive transcription of the *gusA* reporter

A. McCracken · M.S. Turner · P. Giffard · L.M. Hafner
P. Timms (✉)
Centre for Molecular Biotechnology, School of Life Sciences,
Queensland University of Technology,
2 George Street, Brisbane, QLD, Australia
e-mail: p.timms@qut.edu.au,
Tel.: +61-7-38642120, Fax: +61-7-38641534

gene in three strains of *Lactobacillus*. Previously defined promoters of *Lactobacillus* and lactococcal origin were used for comparison. We then compiled a database of mapped *Lactobacillus* promoters and analysed them for conserved motifs that may represent determinants of promoter activity in this genus.

Materials and methods

Bacterial strains and growth conditions

Lactobacillus acidophilus ATCC 4356, *L. rhamnosus* GG (ATCC 53103) and the guinea pig vaginal isolates *L. fermentum* BR11 (Rush et al. 1994) and *Lactobacillus plantarum* BR3 (this laboratory) were grown in MRS (de Mann, Rogosa and Sharpe) medium (Oxoid). Luria-Bertani (LB) medium (Sambrook et al. 1989) was used for the growth and propagation of *E. coli* KW1. Chloramphenicol was used at a final concentration of 10 μg ml^{-1} or 25 μg ml^{-1} for the culture of recombinant lactobacilli and *E. coli*, respectively. Histochemical screening of *gusA*-positive clones used 5-bromo-4-chloro-3-indoyl-β-D-glucuronide (X-gluc; AGP Technologies) in MRS and LB agars at a final concentration of 0.5 mM.

DNA isolation and manipulation

Plasmid DNA was isolated from *E. coli* KW1 using the alkaline lysis method (Sambrook et al. 1989). Plasmid DNA was isolated from *Lactobacillus* according to the method of O'Sullivan and Klaenhammer (1993). Total DNA was isolated from lactobacilli essentially as described by Scheirlinck et al. (1989). Electrocompetent *E. coli* KW1 cells were prepared and transformed by electroporation as described by Sambrook et al. (1989). Plasmid DNA was introduced into *Lactobacillus* strains using the methods of Rush et al. (1994) and Wei et al. (1995). Penicillin concentrations of 1, 5 and 10 μg ml^{-1} were used for *L. fermentum* BR11, *L. plantarum* BR3 and *L. rhamnosus* GG, respectively. Buffers containing $MgCl_2$ were used for *L. fermentum* BR11 and *L. plantarum* BR3 only. Restriction and modification enzymes were obtained from Boehringer Mannheim and used in accordance with manufacturer's instructions. DNA was sequenced using an Applied Biosystems 373A automated DNA sequencer.

RNA isolation, primer-extension mapping and quantitative Northern blot analysis

Total RNA was isolated from exponentially (OD_{600nm} 0.6–0.8) growing lactobacilli using the method of Platteeuw et al. (1994), except that Macaloid clay suspension was not used. RNA was treated with 10 U of RNase-free DNase (Boehringer Mannheim) prior to final ethanol precipitation. Primer-extension mapping of total RNA was carried out with end-labeled [γ^{33}P]-ATP oligonucleotide GUS-AS (5′-GGGTTGGGGTTTCTACAGGACGTA-3′) as described by McCracken and Timms (1999). The quantitative Northern blot procedure and preparation of digoxigenin-labeled *gusA* and 16S rRNA probes have been described previously (McCracken and Timms 1999). Quantification of signals involved standardisation for linear response range and RNA/signal correlations using serial RNA dilutions in preliminary experiments. To standardise the Northern blot data, the *gusA* signal was divided by the 16S rRNA signal of the same RNA sample. Triplicate RNA samples, prepared independently, were analysed separately for each clone. Student's *t*-test was used for comparison of means for determining statistical significance.

Relative copy number determination

The relative copy number of plasmids within *Lactobacillus* species was determined essentially as described by Platteeuw et al. (1994). Random-prime ^{32}P-labeled *gusA* and 16S rRNA PCR fragments (McCracken and Timms 1999) were used as probes. The signals were quantified with a Fujifilm BAS1500 phosphoimager (Fuji) and AIS software (Imaging Research). The relative copy number within a single recombinant *Lactobacillus* strain was calculated as the ratio of the average *gusA* signal to the average 16S rRNA signal.

Isolation of *L. rhamnosus* GG genomic promoter fragments

Lactobacillus rhamnosus GG genomic DNA was digested to completion with *Sau*3AI and ligated to *Bam*HI-digested, alkaline phosphatase-treated pNZ272, a *Lactobacillus–E. coli* shuttle vector harbouring the promoterless *E. coli gusA* reporter gene downstream of a multiple cloning site (Platteeuw et al. 1994). *L. rhamnosus* GG was transformed directly with the ligation mixture and plated onto selective agar for the isolation of recombinant lactobacilli. Recombinant clones were subsequently transferred to MRS agar containing 10 μg chloramphenicol ml^{-1} and 0.5 mM X-gluc for histochemical screening. Constructs originating from recombinant *L. rhamnosus* GG that displayed a blue phenotype were extracted and transferred into *E. coli* KW1 for sequencing and purification. Plasmids were subsequently introduced into *L. fermentum* BR11 and *L. plantarum* BR3, where possible.

Isolation and characterisation of the putative *L. fermentum* BR11 *clpC* promoter

A genomic library of *L fermentum* BR11 was immunoscreened with an antiserum raised against *L. fermentum* BR11 whole cells. Construction of the library, preparation of the antiserum, and immunoscreening have been described previously (Turner et al. 1997). Sequence analysis of an immunoreactive clone indicated that it potentially encoded a polypeptide similar to previously characterised ClpC ATPase chaperones (Lazazzera and Grossman 1997), with a second open reading frame (*orf1*) immediately upstream of the *clpC* coding region. A 453-bp sequence encompassing the intergenic region upstream from the beginning of the putative *orf1* gene was amplified by PCR using *Pst*I site-tagged primers, cleaved with *Pst*I and ligated to *Pst*I-cleaved pNZ272. The resulting plasmid, pNZclp, was used to transform *L. rhamnosus* GG, *L. fermentum* BR11 and *L. plantarum* BR3.

Control promoters *lacA* and *slpA*

Plasmid pNZ276, harbouring the *Lactococcus lactis lacA* promoter downstream of the divergently transcribed *lacR* gene (Platteeuw et al. 1994), was obtained (W. de Vos, NIZO) and introduced into each *Lactobacillus* strain. The *L. acidophilus* ATCC 4356 *slpA* promoter region, from position –286 to –20 (relative to the translational start codon) was amplified by PCR using the primer pair 4356–5P (5′-AAAAGGATCCTGCTTGTGGGGTAAGCGG-3′) and 4356–3P (5′-TTTTCTGCAGATATAAAAAAATGTAATAGGCC-3′), digested with *Bam*HI/*Pst*I and ligated to *Bam*HI/*Pst*I-digested pNZ272. The resulting plasmid, pNZslp, was introduced into *L. rhamnosus* GG, *L. fermentum* BR11 and *L. plantarum* BR3 by electroporation.

Computer analysis

Lactobacillus promoter sequences were obtained from GenBank or EMBL via The Australian National Genome Information Service. Sequences (available upon request) were aligned either by the transcriptional start site (89 promoters, extending from position –80 to +26 where available) or by the proposed –35 and –10 regions (96 promoters, extending from 45 bp upstream of the –35 region to 13 bp downstream of the –10 hexamer, where available). Aligned sequences were analysed using the Consensus (Genetics Computer Group) alignment program. The frequency of each base

Table 1 Nucleotide sequences and transcriptional start sites of promoters in *Lactobacillus* sp. (*bold typeface* indicates identity with consensus; / indicates long –10-transcriptional start site spacer region; *dots* are used for alignment of hexamer sequences and start site; consensus element sequences: UP element Estrem et al. (1998); hexamers, Pouwels and Leer (1993)

Promoter	UP element AA A	–35	–10	Transcriptional start site
Consensus	nnAAA T TTTTnnAAAAnnn	.TTGACA.....	.....TATAAT.....	.A/G.....
	TT T			
P-GL1	CACGG**TTTT**AAAATG**A**GCGTTGT**TTG**T**CA**GCAAGCATTT	...TATCCCG**TATA**G**T**TAT	...TAAC.**A**.TAATAAGTT	
P-GL2	CGTC**AATTA**C**TT**ATATT**AA**AAGC**T**AACA**A**AATCAGCGGCT	..TTTAAGA**TAC**A**AT**CGT	...TGTG.**A**.ATCAGAAAG	
P-GL3-GG	AAGG**AT**GA**A**GAAGGCACTGGCGG**T**AAGA**A**AAAGCTGTTT	...AGTCGTGA**AG**CG**T**CGT	....TTT.**A**.AGAACCCGG	
P-GL3–11	GC**AA**CGC**T**GCACAGTGCT**A**AGCCAA**G**GATGAAGAAGGCA	...CTGGCGG**TA**AG**A**AAAA	....GCT.**G**.TTTAGTCGT	
P-GL4–1	GA**AA**G**AA**CC**T**G**T**GATTGTCAAGA**TT**A**A**GGCTAAGGAACTGA	.CAGCAGA**TA**CCG**T**TGA	.....GC.**A**.ATATTATGT	
P-GL4–2	ACTG**A**C**A**GCAGATACCGTTGAGCAAT**A**TTATGTTCGGGC	...CAAGGAC**TATG**A**A**AAG	...TTCG.**A**.TGTCATGAC	
P-GL9c	TTTTT**A**GACCAAAAAA**A**CTTCGCC**TTGA**TGCCAGTTCGAG	...TTGGCGC**TATTTT**AAA	....CAT.**G**.TGCCTCTCT	
P-GL9–3	GCC**A**G**TT**CGAG**TT**GGCGCTATTT**T**AA**ACA**TGTGCCTCTCT	..TGACACT**TATTA**GTCG	.ACCACT.**A**.CATACCTAT	
P-GL10c	TGCC**AA**CA**T**GA**T**CTAGCCGCTGG**TT**CCCG	GAAGGTGGCC	...GTCTCCG**T**TAT**AT**GTT	./.TGCC.**A**.TTTTACCAG
P-GL10-GG	AGCT**AAA**A**A**AAGAACG**AT**A**T**ATCT**TTG**C**CA**ACATGATCTAGC	.CCGCTGG**T**TCCCG	GAA.....	G.**G**.TGGCCGTCT
P-GL21c	TGGCC**A**AGCCCGGGCC**A**TCACGC**T**CA**AG**GACACTTATGC	...TCCGGTT**TCTAA**CGG	...GCTG.**G**.CTCACGCTC	
P-GL21-GG	CCGGGC**T**GGC**T**CACGCTCTTTGA**TTGAC**TTGCGTTGTC	....AAATCTT**TATAC**TGAA	....ACA.**A**.GTAATGATC	
P-*clpC*	CTT**AAAATAT**AG**T**CAT**AG**A**A**TTA**GG**G**C**GTCGGTTTTAAA	...AGTGTGG**TAAAAT**AAT	.....G.**G**T.CAAGATTA	
P-*lacA-1*	AA**AC**A**AA**A**A**TT**A**AC**AAAA**ATAG**TTG**CGTTTTGTTTGAA	...TGTTTGA**TTTCAT**ATA	.....AA.C.AAAGAAATG	
P-*slpA*	TAGT**ATTTTT**CGGTC**A**TTTTAAC**TTG**C**TA**TTTCTTGAAG	...AGGTTAG**TAC**A**AT**ATG	...AATC.**G**.TGGTAAGTA	

at each position of the sequence was determined and the highest value at each position was used to generate histograms. Promoters of *Lactobacillus delbrueckii* subsp. *lactis* (13), *L. plantarum* (9), *Lactobacillus helveticus* (12) and *L. rhamnosus* (12) were aligned for analysis at the species level.

Results

Isolation and characterisation of promoter-active fragments in *Lactobacillus* sp.

Seven recombinant *L. rhamnosus* GG clones harbouring *L. rhamnosus* GG genomic fragments produced detectable primer-extension products. All *L. rhamnosus* GG genomic fragments were also transcriptionally active in the heterologous hosts *L. fermentum* BR11 and *L. plantarum* BR3 via primer-extension analysis. Identical transcriptional start sites (primarily purine residues) were used in the three organisms for most of the genomic promoter fragments examined (Table 1). The only exception was fragment GL3, where the functional start sites used in *L. rhamnosus* GG and *L. fermentum* BR11 were different (designated P-GL3-GG and P-GL3–11, respectively). Several promoter fragments directed transcription of the *gusA* gene from multiple start sites and thus from multiple promoters. This was particularly evident in *L. rhamnosus* GG where multiple promoters were utilised from three of the seven genomic promoter-active fragments (GL4, GL10 and GL21, in addition to the *lacA* promoter, see below).

Potential –35 and –10 hexamers were identified by analogy with the previously defined *Lactobacillus* consensus sequences (Pouwels and Leer 1993). Preference was given to motifs that matched the *Lactobacillus* consensus sequence at the most conserved positions of the hexamers and gave rise to a –35/–10 spacer of 17±1 bp. The most consistent hexamer sequences were within promoters P-GL1 and P-GL21-GG, both of which displayed five out of six matches to the consensus with each hexamer (Table 1). It was difficult to identify hexamer motifs in sev-

Table 2. Relative transcriptional activities of promoters in lactobacilli (% relative to P-GL1 for individual strains, mean±SD (%) for three independent determinations). *ND* Not determined; *NT* no transcript detected

Promoter	Relative *gusA* transcript level (%)		
	Lactobacillus rhamnosus GG	*Lactobacillus fermentum* BR11	*Lactobacillus plantarum* BR3
P-GL1 (100%)	100 ±10	100 ± 14.8	100 ± 30.5
P-GL2	29.5± 8.5	12.8± 2.6	21.3± 4.4
P-GL3-GG	31.4± 2.1	NT	ND
P-GL3–11	NT	20.2± 1.6	ND
P-GL4–1	40.1± 8.3	20.2± 1.7	ND
P-GL4–2	26.4± 5.5	23.1± 1.9	ND
P-GL9c	5.6± 2.7	29.3± 1.9	40 ± 2.1
P-GL9–3	NT	NT	67.5± 3.6
P-GL10c	2.7± 0.6	15.7± 4.1	42 ± 2.0
P-GL10-GG	11.9± 2.5	NT	NT
P-GL21c	80.5± 6.0	32.4± 5.9	ND
P-GL21-GG	56.6± 4.2	NT	ND
P-*clp*C	714 ±58.5	540 ±410	1551 ±440
P-*lacA*-1	529 ±47.6	14.6± 2.4	NT
P-*lacA*-2	546 ±49.1	NT	NT
P-*slpA*	2069 ±641	668 ± 90.5	10889 ±786

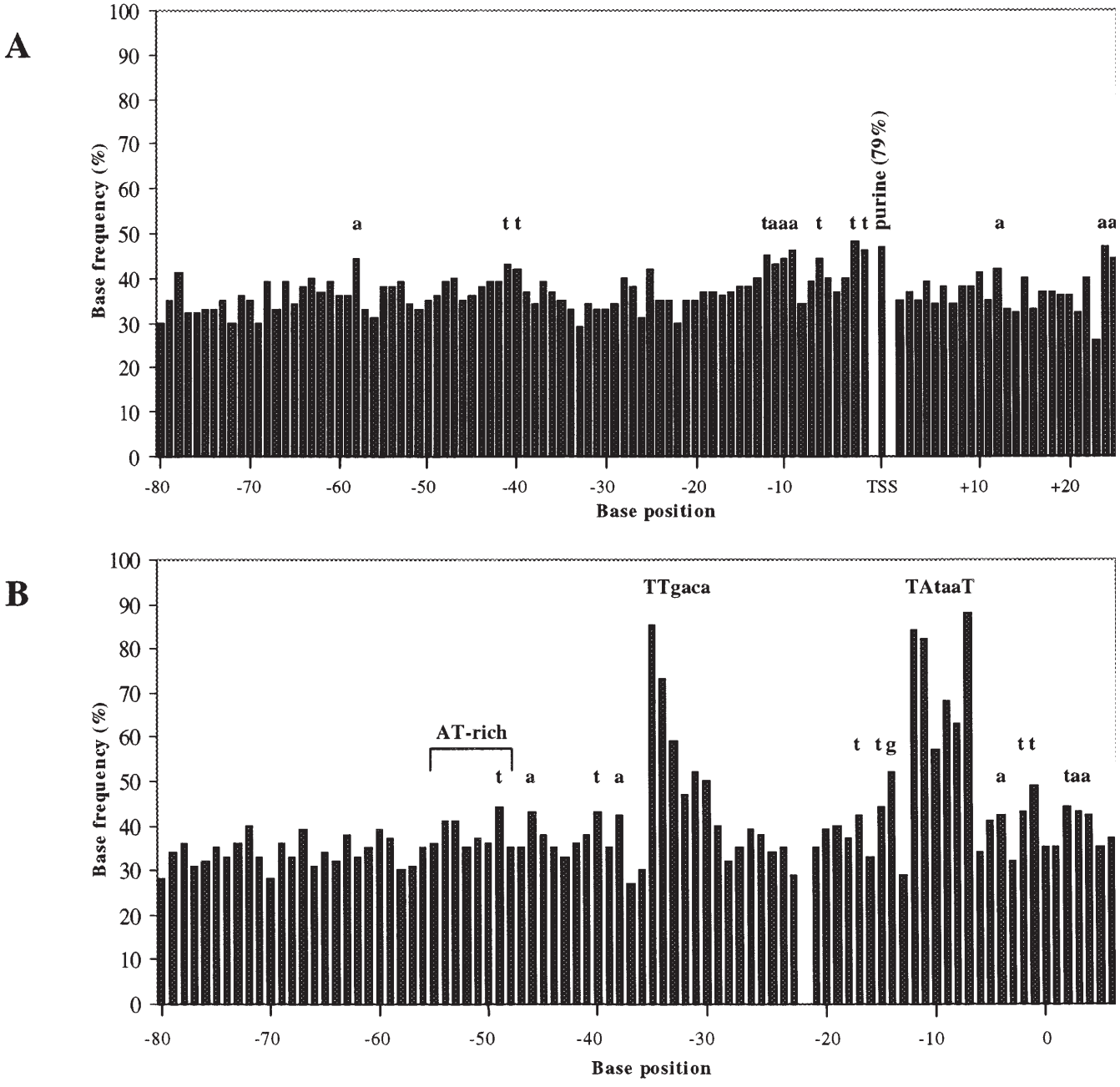

Fig. 1 Base conservation within transcriptionally mapped *Lactobacillus* promoter sequences aligned to the transcriptional start site (**A**), or the –35 and –10 hexamers (**B**). The abundance of the most frequent base is plotted as a function of sequence position

eral of the promoters, such as P-GL3-GG and P-GL3–11, as they were not highly homologous to the consensus sequences. Weakly homologous structures were postulated in such cases (Table 1). Dual adjacent residues located within the *clpC* promoter region were utilised for transcription initiation in each of the three lactobacilli. The transcriptional start sites were preceded by a high-identity –10 hexamer and a low-identity –35 hexamer. A TG motif was also observed upstream of the P-*clpC* –10 hexamer.

Transcription was initiated at the published P-*lacA* lactococcal start site (Van Rooijen et al. 1992) in *L. rhamnosus* GG (designated P-*lacA*-1). Transcription was also initiated from a second start site located approximately 100–120 bp upstream of the published site (designated P-*lacA*-2). Due to heterogeneous primer-extension products, however, neither the P-*lacA*-2 start site nor upstream sequences could be precisely mapped. Quantitative analysis indicated that P-*lacA*-1 and P-*lacA*-2 were utilised by *L. rhamnosus* GG with equal efficiency. P-*lacA*-driven *gusA* mRNA was produced at low levels in *L. fermentum* BR11 but was not detected in *L. plantarum* BR3. Primer-extension data indicated that the *slpA* promoter sequence was functional in each of the three lactobacilli, initiating transcription from the published *L. acidophilus* ATCC 4356

Table 3. *Lactobacillus* promoter consensus sequences (*f* a/t, *h* a/g, ***T*** ≥75%, *T* 60–74%, *t* 40–59%)

Species (no. promoters)	Sequence position						Reference
	–50	–35	–24	–15	–10	–1	
Lactobacillus (96)	aa...t..a.....a.t..	**T**Tgaca	t.......	a.t.tg.	**TA**t**AAT**	.aa.tt	This study
Lactobacillus delbrueckii subsp. *lactis* (13)	....aT.....tt.tTt.a	**T**Tg.ca	t.ttT.a.	..t.**T**Gc	**T**A.a**AT**	t...**T**T	This study
Lactobacillus helveticus (12)	...f.T..attta..t...	**TTG**acf	tttctct.	..ttTgg	**TA**t**AAT**	aTaaat	This study
L.plantarum (9)	......**T**.taaA.a.aTat	TT.aca	atTt.TaA	.AtA...	**TATA**c**T**	A**A**aatt	This study
Lactobacillus rhamnosus (12)	aa..........t...G.	**T**Thaca	..t..t..	t....gg	**TA**taa**T**	.a...t	This study
Bacillus subtilis	a.....t...aaaa...t.	**TTG**Aca	....	..t.tg.	**TA**t**AAT**	aa.a..	Helmann (1995)
Escherichia coli	.........a......t..	**TT**Gaca	....	..t.tg.	**TA**taa**T**	......	Chassy and Murphy (1993)

start site (Boot et al. 1995) in all organisms examined. Quantitative analysis revealed that P-*slpA* was used with a 4.5-fold greater efficiency in *L. rhamnosus* GG than in *L. fermentum* BR11 and *L. plantarum* BR3.

The promoter elements encompassed a 1140-fold, 52-fold and 512-fold range of activity in *L. rhamnosus* GG, *L. fermentum* BR11 and *L. plantarum* BR3, respectively (Table 2). In all organisms P-*slpA* directed the highest levels of *gusA* mRNA production. This was particularly pronounced in *L. plantarum* BR3, where *gusA*-specific mRNA levels produced from the *slpA* promoter were 109- to 512-fold higher than from any of the randomly isolated, genomic promoter-active fragments examined ($p<0.01$). The P-*lacA*-1 and P-*lacA*-2 promoters were also used at significantly higher levels (5.3- to 206-fold) in *L. rhamnosus* GG than were the promoters of random genomic origin ($p<0.01$). The *clpC* promoter directed transcription of *gusA* mRNA to higher levels than any of the randomly isolated promoter fragments of *L. rhamnosus* GG origin. Of the randomly isolated *L. rhamnosus* GG genomic promoters, P-GL1 produced the most *gusA*-specific mRNA in all three lactobacilli, 1.8- to 17.9-fold higher than any other genomic fragment (Table 2). The other promoter fragments displayed varying activities depending upon the strain in which they were harboured (Table 2).

Defining the *Lactobacillus* sp. promoter consensus sequence

Promoters (of both *Lactobacillus* gene and bacteriophage origin) with experimentally defined transcriptional start sites, including the promoters isolated in the current study, were aligned. When aligned to the transcriptional start site, no single base was more than 50% conserved at any one position (Fig. 1A). Despite this lack of overall base conservation some conserved features of *Lactobacillus* promoters were observed. The start site was a purine residue in 79% of all *Lactobacillus* promoters. Other residues with a moderate degree of conservation included a T at –1, –2, –40 and –41 and an A residue at positions –58, +25 and +26. The region from –12 to –7 was rich in A and T residues (average 73% A+T).

When *Lactobacillus* promoters were aligned to the –35 and –10 hexamers (Fig. 1B) the pattern of nucleotide conservation resembled that observed for *E. coli* and *B. subtilis*. The consensus sequences for promoters of *Lactobacillus* at the genus and species levels are represented in Table 3. Highest conservation was observed in the –10 element where the last T residue was the most conserved nucleotide position. Conservation within the –35 hexamer was largely limited to the initial T residues of the hexamer. Many other positions within the promoter region exhibited a degree of sequence conservation in *Lactobacillus* promoters when aligned to the hexamer sequences, albeit to a lesser degree (Fig. 1B). Several of these base positions were also conserved at the species level (Table 3). The entire region upstream of the –35 hexamer (-80 to –35) was enriched for short A and T tracts, with an AT-rich region (average 72% A+T) centred around position –53. A and T clusters, centred between position –60 and –40, were also observed upstream of the –35 hexamer of *L. delbrueckii* subsp. *lactis*, *L. helveticus*, *L. plantarum* and *L. rhamnosus* promoters (data not shown). The average distance between the –35 and –10 hexamers of *Lactobacillus* promoters was 17.4±1.6 bp, whilst the average distance between the –10 hexamer and the transcriptional start site was 6.9±2.7 bp.

The T residue at –15 and the G residue at –14 were both moderately well-conserved in the *Lactobacillus* promoters analysed in this study (Fig.1B). The bases occurred as a dinucleotide in 26% of all *Lactobacillus* promoters. Importantly, conservation of the TG motif appeared to be species-dependent. The motif was moderately conserved in *L. delbrueckii* subsp. *lactis* (54%) and *L. helveticus* (33%) promoters, but was not a conserved feature of the *L. plantarum* or *L. rhamnosus* promoters analysed in this study (Table 3).

Discussion

With the exception of the *Lactococcus lactis lacA* promoter, which is regulated by the divergently transcribed *lacR* repressor gene in the natural lactococcal host and in *E. coli* (Van Rooijen et al. 1992), all promoter fragments tested directed transcription in all three lactobacilli investigated. These results indicate that promoters recognised by RNA polymerase of *L. rhamnosus* GG, *L. fermentum* BR11 and *L. plantarum* BR3 share significant similarities. Differences between promoter strengths in each strain, however, indicate that discrete but significant differences defining promoters exist between strains of lactobacilli. These differences were reflected in the consensus sequence variations between *Lactobacillus* species defined in the current study.

Conserved elements with consensus sequences closely resembling the –35 and –10 hexamer motifs of vegetative promoters from *E. coli* and *B. subtilis* were observed upstream of *Lactobacillus* transcriptional start sites in the current work, including those promoters isolated from the genome of *L. rhamnosus* GG. Agreement with the *E. coli* and *B. subtilis* consensus sequences was higher at the –10 hexamer of *Lactobacillus* promoters than the –35 hexamer, both at the genus and species levels. Promoters functioning with highest efficiency in *Lactobacillus* in the current study showed consistently good identity with the consensus hexamers, particularly the –10 hexamer, substantiating the importance of these motifs in promoter function and strength. In contrast to observations in other organisms, however, good identity at the –35 and –10 hexamer positions alone was not a prerequisite for promoter function in *Lactobacillus* (e.g. P-GL3-GG and P-GL21c). In addition, the presence of high-identity –35 and –10 hexamers did not always correlate with efficient promoter activity in the organism (e.g. P-GL1 and P-GL9c). These results clearly suggest that the –35 and –10 hexamers are important, but not ultimate factors determining promoter strength in lactobacilli.

A number of additional base positions outside of the traditional hexamers displayed varying degrees of conservation in *Lactobacillus* promoters. The functional significance of these bases remains unclear; however, they may play important roles in defining a promoter sequence in *Lactobacillus*. Conserved bases corresponding to the consensus AT-rich UP element (Estrem et al. 1998) were identified in *Lactobacillus* promoters, both at the genus and species levels. Most notable were the T at position –50 and the A at position –41 which were conserved in 44% and 43% of all *Lactobacillus* promoters, respectively. Interestingly, the promoters displaying greatest strength in the current study, P-*slpA* and P-*lacA*-1 (in *L. rhamnosus* GG), identified very well with the UP element consensus (Table 1). The weaker promoters, even those which conformed well to the –35 and –10 hexamer consensus sequences, identified to the UP element consensus to a reduced extent. This observation suggests that the UP element, which has been shown to contribute significantly to promoter strength in other organisms (Estrem et al. 1998; Ross et al. 1998), may play a role in transcription in lactobacilli. The TG motif was conserved in 26% of *Lactobacillus* promoters. Conservation of the motif varied considerably depending on the species (up to 54%), perhaps indicating that functional significance of the TG motif is determined at the species level. The TG motif has been associated with promoter strength in several gram-positive organisms (Van der Vossen et al. 1987; Voskuil et al. 1995), including lactobacilli (McCracken and Timms 1999).

The *L. acidophilus* ATCC 4356 *slpA* promoter directed the highest levels of transcription in all *Lactobacillus* strains examined in the current study. Combined with data from other researchers indicating that the *slpA* promoter is highly efficient in the natural host and in *L. casei* ATCC 393 (Boot et al. 1996), this suggests that the *slpA* promoter sequence harbours motifs that interact effectively with RNA polymerase of several lactobacilli. This finding has striking implications for the development of heterologous gene-expression cassettes in lactobacilli by offering a powerful tool for achieving high levels of foreign protein production in these organisms. The *clpC* promoter also displayed high activity in all three *Lactobacillus* strains tested when present as a transcriptional fusion with the *gusA* gene, although significantly less active than P-*slpA*. Recent evidence suggests that P-*clpC* is likely to be regulated by the product of the *orf1* gene (Derré et al. 1999). The strong basal activity of P-*clpC* in the absence of *orf1*, however, indicates that it is potentially useful for driving heterologous gene expression in lactobacilli.

The use of multiple transcriptional start sites was a common finding in the current study, particularly in *L. rhamnosus* GG. This organism utilised promoter sequences with considerable promiscuity and readily initiated transcription from two promoters when presented in a synthetic tandem arrangement (data not shown). Tandem promoters have been identified upstream of several highly expressed gram-positive genes including the *Bacillus brevis* cell wall protein genes (Yamagata et al. 1987) and the *Lactobacillus brevis* S-layer gene (Vidgrén et al. 1992). Multiple promoters would be expected to lead to higher levels of transcript than would a single promoter. This property of *L. rhamnosus* GG may prove useful in the design of expression cassettes for the organism.

It is possible that the new promoters isolated in the current study are recognised by an alternative RNA polymerase σ-factor. No *Lactobacillus* alternative σ-factors have been identified to date, nor have alternative σ-factor-dependent promoter sequences been mapped. None of the promoters analysed in the current work showed identity to promoter sequences determined for alternative σ-factors of *E. coli* or *B. subtilis.* Assuming that alternative σ-factors exist in lactobacilli, it is likely that much of their activities may be growth-phase dependent or environmentally regulated (Lonetto et al 1992). In the current work, all analyses were done with cells taken from the exponential phase of growth; hence it is less likely that differences in promoter activity could be explained solely by alterna-

tive σ-factor binding. It is possible that promoters analysed in the current study were subject to regulation at the transcriptional level. Such promoter sequences may be expected to show less structured identity with consensus hexamers, particularly at the –35 region.

The current study has identified attributes that are important in the control of transcription in lactobacilli. This information will no doubt benefit the design of heterologous gene expression cassettes in this genus. Detailed mutagenesis of promoter regions displaying conservation would be necessary to gain further information on the role of such elements in defining promoter function and strength in lactobacilli.

Acknowledgements We are grateful to W. de Vos (NIZO, The Netherlands) for kindly providing plasmids pNZ272 and pNZ276. We thank S. Mathews for technical advice. This work was funded by the National Health and Medical Research Council of Australia, grant number 941114.

References

Boot HJ, Kolen CPAM, Pouwels PH (1995) Identification, cloning and nucleotide sequence of a silent S-layer protein gene of *Lactobacillus acidophilus* ATCC 4356 which has extensive similarity with the S-layer protein gene of this species. J Bacteriol 177:7222–7230

Boot HJ, Kolen CPAM, Andreadaki FJ, Leer RJ, Pouwels PH (1996) The *Lactobacillus acidophilus* S-layer protein gene expression site comprises two consensus promoter sequences, one of which directs transcription of stable mRNA. J Bacteriol 178:5388–5394

Chassy BM, Murphy CM (1993) *Lactococcus* and *Lactobacillu*s. In*:* Sonenshein AL, Hoch JA, Losick R (eds) *Bacillus subtilis* and other gram-positive bacteria. Biochemistry, physiology and molecular genetics. American Society for Microbiology, Washington DC, pp 65–82

Derré I, Rapoport G, Msadek T (1999) CtsR, a novel regulator of stress and heat shock response, controls *clp* and molecular chaperone gene expression in gram-positive bacteria. Mol Microbiol 31:117–131

Estrem ST, Gaal T, Ross W, Gourse RL (1998) Identification of an UP element consensus sequence for bacterial promoters. Proc Natl Acad Sci USA 95:9761–9766

Harley CB, Reynolds RP (1987) Analysis of *E.coli* promoter sequences. Nucleic Acids Res 15:2343–2361

Helmann JD (1995) Compilation and analysis of *Bacillus subtilis* σ^A-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. Nucleic Acids Res 23:2351–2360

Lazazzera BA, Grossman AD (1997) A regulatory switch involving a Clp ATPase. Bioessays 19:455–458

Lisser S, Margalit H (1993) Compilation of *E.coli* mRNA promoter sequences. Nucleic Acids Res 21:1507–1516

Lonetto M, Gribskov, M, Gross, CA (1992) The σ^{70} family: sequence conservation and evolutionary relationships. J Bacteriol 174:3842–3849

McCracken A, Timms P (1999) Efficiency of transcription from promoter sequence variants in *Lactobacillus* is both strain and context dependent. J Bacteriol 181:6569–6572

Mercenier A, Pouwels PH, Chassy BM (1994) Genetic engineering of lactobacilli, leuconostocs and *Streptococcus thermophilus*. In: Gasson MJ, de Vos WM (eds) Genetics and biotechnology of lactic acid bacteria. Blackie, Glasgow, pp 252–293

Natori Y, Kano Y, Imamoto F (1988) Characterization and promoter selectivity of *Lactobacillus acidophilus* RNA polymerase. Biochimie 70:1765–1774

O'Sullivan DJ, Klaenhammer TR (1993) Rapid mini-prep isolation of high-quality plasmid DNA from *Lactococcus* and *Lactobacillus* spp. Appl Environ Microbiol 59:2730–2733

Platteeuw C, Simons G, de Vos WM (1994) Use of the *Escherichia coli* β-glucuronidase (*gusA*) gene as a reporter gene for analyzing promoters in lactic acid bacteria. Appl Environ Microbiol 60:587–593

Pouwels PH, Leer RJ (1993) Genetics of lactobacilli: plasmids and gene expression. Antonie van Leeuwenhoek 64:85–107

Ross W, Gosink K, Salomon J, Igarashi K, Zhou C, Ishihama A, Severinov K, Gourse R (1993) A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase. Science 262:1407–1413

Ross W, Aiyar SE, Salomon J, Gourse RL (1998) *Escherichia coli* promoters with UP elements of different strengths: modular structure of bacterial promoters. J Bacteriol 180:5375–5383

Rush CM, Hafner, LM, Timms P (1994) Genetic modification of a vaginal strain of *Lactobacillus fermentum* and its maintenance within the reproductive tract after intravaginal administration. J Med Microbiol 41:272–278

Sambrook J, Fritsch EF, Maniatus J (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Scheirlinck T, Mahillon J, Joos H, Dhaese P, Michiels F (1989) Integration and expression of α-amylase and endoglucanase genes in the *Lactobacillus plantarum* chromosome. Appl Environ Microbiol 55:2130–2137

Turner MS, Timms P, Hafner LM, Giffard PM (1997) Identification and characterisation of a basic cell-surface located protein from *Lactobacillus fermentum* BR11. J Bacteriol 179:3310–3316

Van der Vossen JMBM, van der Lelie D, Venema G (1987) Isolation and characterization of *Streptococcus cremoris* Wg2-specific promoters. Appl Environ Microbiol 53:2452–2457

Van Rooijen RJ, Gasson MJ, de Vos WM (1992) Characterization of the *Lactococcus lactis* lactose operon promoter: contribution of flanking sequences and LacR repressor to promoter activity. J Bacteriol 174:2273–2280

Vidgrén G, Palva I, Pakkanen R, Lounatmaa K, Palva A (1992) S-layer protein gene of *Lactobacillus brevis*: cloning by polymerase chain reaction and determination of the nucleotide sequence. J Bacteriol 174:7419–7427

Voskuil MI, Voepel K, Chambliss GH (1995) The –16 region, a vital sequence for the utilization of a promoter in *Bacillus subtilis* and *Escherichia coli.* Mol Microbiol 17:271–279

Wei MQ, Rush CM, Norman JM, Hafner LM, Epping RJ, Timms P (1995) An improved method for the transformation of *Lactobacillus* strains using electroporation. J Microbiol Methods 21:97–109

Yamagata H, Adachi T, Tsuboi A, Takao M, Sasaki T, Tsukagoshi N, Udaka S (1987) Cloning and characterization of the 5' region of the cell wall protein gene operon in *Bacillus brevis* 47. J Bacteriol 169:1239–1245