Clustered organization and transcriptional analysis of a family of five csp genes of Lactococcus lactis MG1363

Wouters, Jeroen A.; Sanders, Jan-Willem; Kok, Jan; Vos, Willem M. de; Kuipers, Oscar P.; Abee, Tjakko; Wouters, J.W.

Published in:
Microbiology

DOI:
10.1099/00221287-144-10-2885

Publication date:
1998

Microbiology (1998), **144**, 2885–2893 Printed in Great Britain

Clustered organization and transcriptional analysis of a family of five *csp* genes of *Lactococcus lactis* MG1363

Jeroen A. Wouters,[1,2] Jan-Willem Sanders,[3] Jan Kok,[3] Willem M. de Vos,[1] Oscar P. Kuipers[1] and Tjakko Abee[2]

Author for correspondence: Oscar P. Kuipers. Tel: +31 318 659525. Fax: +31 318 650400. e-mail: kuipers@nizo.nl

[1] Microbial Ingredients Section, NIZO Food Research, PO Box 20, 6710 BA Ede, The Netherlands

[2] Laboratory for Food Microbiology, Food Science Group, Wageningen University and Research Center, Wageningen, The Netherlands

[3] Department of Genetics, Groningen University, Haren, The Netherlands

A family of genes encoding cold-shock proteins, named *cspA*, *cspB*, *cspC*, *cspD* and *cspE*, was cloned and sequenced from *Lactococcus lactis* MG1363. The genes *cspA* and *cspB* and the genes *cspC* and *cspD* are located in tandem repeats, an organization of *csp* genes that has never been encountered before. The five genes encode small (7·1–7·6 kDa) proteins with high mutual sequence identities (up to 85%) and high identities (about 45–65%) with the major cold-shock proteins from *Escherichia coli* (CspA) and *Bacillus subtilis* (CspB). Northern-blot analysis revealed single transcripts of about 300 nucleotides for each *csp* gene and showed that *cspA*, *cspB*, *cspC* and *cspD* mRNA levels were strongly increased upon cold shock to 10 °C (about 10-, 40-, 10- and 30-fold compared to 30 °C, respectively), whereas the *cspE* mRNA level was not increased. The expression of the cold-induced *csp* genes was highest in the 6–8 h lag phase after cold shock. A differential expression in time, in which *cspA* and *cspC* were maximally expressed at 2 h and *cspB* and *cspD* at 4 h after cold shock, was observed. The 35 and 10 regions of the five promoters were identified and transcriptional start sites were mapped in each case by primer extension at different temperatures which confirmed that regulation takes place at the transcriptional level. Significant differences were observed between the 5′-untranslated leader regions of the four cold-induced *csp* genes and the corresponding region of the non-cold-induced *cspE* gene.

Keywords: *csp* genes, low-temperature adaptation, transcription, *Lactococcus lactis*

INTRODUCTION

Lactococcus lactis plays an important role in many dairy fermentations. During processing and ripening of fermented dairy products these bacteria have to deal with different environmental stresses, such as low pH, high salt concentrations and temperature extremes (Rallu *et al*., 1996). Several stress responses of *L. lactis* have been studied and stress-induced genes could be identified (Rallu *et al*., 1996; Sanders *et al*., 1995; Van Asseldonk *et al*., 1993). However, low temperature stress has received less attention. Cold stress might be important for the survival of starter cultures after frozen storage and for fermentations taking place at low temperatures.

The response to cold shock has been extensively studied in *Escherichia coli* and has been shown to result in the induction of a specific set of 14 proteins. These proteins play a role in various cellular processes and include, among others, NusA, RecA, H-NS, GyrA, polynucleotide phosphorylase and CspA (further referred to as $CspA^{E}$) (Jones *et al*., 1987, 1996; Jones & Inouye, 1994, 1996). Maximal induction after cold shock was detected for $CspA^{E}$, which is transiently overexpressed (200-fold induction) and then represents 13% of the newly synthesized proteins (Goldstein *et al*., 1990; Jones *et al*., 1987). A highly similar protein, CspB (further referred to as $CspB^{B}$), has been described in *Bacillus subtilis* (Willimsky *et al*., 1992).

$CspA^{E}$ (Goldstein *et al*., 1990) and $CspB^{B}$ (Willimsky *et*

Abbreviations: CSP, cold-shock protein; UTR, untranslated leader region.

The EMBL accession numbers for the sequences reported in this paper are Y17215 (for *cspA* and *cspB*), Y17216 (for *cspC* and *cspD*) and Y17217 (for *cspE*).

al., 1992) are small proteins with a molecular mass of 7·4 kDa and a low isoelectric point (pI 5·9 and 4·3, respectively). $CspA^E$ acts as a transcriptional activator of at least two other genes encoding the cold-induced proteins GyrA (Jones *et al*., 1992) and H-NS (LaTeana *et al*., 1991), both involved in DNA supercoiling. The crystal structures of $CspA^E$ and $CspB^B$ have been resolved and both proteins are able to bind specifically to single-stranded DNA containing a Y-box motif (ATTGG) or its complementary sequence (CCAAT) (Graumann & Marahiel, 1994; Newkirk *et al*., 1994; Schindelin *et al*., 1994). Cold-shock proteins (CSPs) contain sequence regions highly homologous to the cold-shock domain of eukaryotic DNA-binding proteins, designated Y-box factors (Landsman, 1992). $CspA^E$ and $CspB^B$ are also considered to be RNA-binding proteins because they both possess highly conserved RNA-binding motifs, i.e. RNP-1 (ribonucleoprotein) and a rudimentary RNP-2 motif (Jones & Inouye, 1994; Schindelin *et al*., 1993), and it appears that $CspA^E$ can act as an RNA chaperone (Jiang *et al*., 1997). For $CspB^B$ a function as an anti-freeze protein has been suggested because a lower survival has been observed after freezing of cells in which the *cspB* gene was disrupted (Willimsky *et al*., 1992). The regulation of the synthesis of the major CSPs is still unclear but it seems to take place at the level of both transcription (Lee *et al*., 1994) and translation (Brandi *et al*., 1996). Recently, it was shown that the abundant presence of $CspA^E$ after cold shock is due to increased stability of its mRNA at low temperature (Fang *et al*., 1997).

In *E. coli*, *B. subtilis* and *Bacillus cereus*, families of *csp* genes of respectively nine, three and six members have been found (Graumann *et al*., 1996; Lee *et al*., 1994; Mayr *et al*., 1996; Nakashima *et al*., 1996; Yamanaka & Inouye, 1997). In *E. coli* at least three of the nine identified *csp* genes are cold induced (Lee *et al*., 1994; Nakashima *et al*., 1996). The *csp* genes of *E. coli* appeared to be scattered on the chromosome (Lee *et al*., 1994) and also for other bacteria only non-clustered *csp* genes have been reported (Graumann *et al*., 1996; Mayr *et al*., 1996). A recent study by Graumann *et al*. (1997) using a triple *csp* deletion mutant of *B. subtilis* revealed that CSPs are essential for cellular growth and for efficient protein synthesis at both optimal and low temperatures.

The cold-shock response of *L. lactis* IL1403 was studied by Panoff *et al*. (1994), revealing that 12 proteins were overexpressed after cold shock. Recently, one cold-induced *csp* gene was identified in *L. lactis* (Chapot-Chartier *et al*., 1997; Kim & Dunn, 1997) and two in another lactic acid bacterium, *Lactobacillus plantarum* (Mayo *et al*., 1997).

In this study, a family of five genes encoding CSPs of *L. lactis* MG1363 was characterized. A clustered organization of *csp* genes has been observed for the first time: two tandems of two *csp* genes. Transcriptional analysis of the *L. lactis csp* genes revealed cold induction for four of these genes and a differential expression of the respective genes during the adaptation phase after cold shock.

METHODS

Bacterial strains and growth conditions. *E. coli* MC1061 (Casadaban & Cohen, 1980) was used as a host strain in cloning experiments and was grown in Tryptone Yeast (TY) medium with aeration at 37 °C (Sambrook *et al*., 1989). Antibiotics were used in the following concentrations: ampicillin 50 μg ml^{-1}; chloramphenicol 10 μg ml^{-1}. *L. lactis* MG1363, a plasmid-free and prophage-cured derivative of *L. lactis* NCDO 712 (Gasson, 1983), was grown in M17 broth (Difco) supplemented with 0·5 % (w/v) glucose at 30 °C without aeration. Growth curves of *L. lactis* were obtained by measuring the OD_{600} at various time points by diluting the sample fourfold in M17 broth.

DNA techniques and sequencing. Chromosomal DNA of *L. lactis* was isolated as described previously (Vos *et al*., 1989). *L. lactis* cells were transformed by electroporation (Wells *et al*., 1993). *E. coli* cells were transformed by the $CaCl_2$ procedure and plasmid isolations were carried out according to established procedures (Sambrook *et al*., 1989). *E. coli* plasmid DNA was isolated on a large scale using Qiagen columns. Restriction enzymes, T4 DNA ligase and other DNA-modifying enzymes were purchased from Gibco-BRL Life Technologies, New England Biolabs or Promega and used as recommended by the manufacturers. Cloning procedures, radiolabelling of DNA fragments, agarose gel electrophoresis and Southern-blot hybridizations were performed according to established procedures (Sambrook *et al*., 1989). DNA fragments were isolated from agarose gels by using the GlassMAX DNA Isolation Matrix System (BRL Life Technologies). PCR was carried out according to conditions described previously (Kuipers *et al*., 1991). Nucleotide sequences of plasmid DNA were analysed with an ALF automatic sequencer (Pharmacia Biotech) in combination with an AutoRead sequencing kit (Pharmacia Biotech) with fluorescein-labelled primers. Oligonucleotides used as primers in sequencing reactions, primer extension experiments and PCR, were purchased from Pharmacia Biotech.

Cloning of *csp* genes. PCR with primers based on homologous regions of $CspA^E$ (Goldstein *et al*., 1990) and $CspB^B$ (Willimsky *et al*., 1992; Table 2) with chromosomal DNA of *L. lactis* MG1363 as a template resulted in the amplification of a fragment of about 200 bp (PCR1) with primers 1 and 2 (both containing an *Eco*RI site). When primers 3 and 4 were used, a fragment of about 550 bp (PCR2) was amplified. The fragments were cloned in pUC18 (pUC18PCR1) and pGEM-T (purchased from Promega; pGEM-TPCR2), respectively. The fragments were sequenced and appeared to contain parts of putative *csp* genes. By use of PCR1 as a probe in Southern hybridization, four hybridizing fragments (*Hin*dIII chromosomal DNA digest) were detected (Fig. 1a). The first hybridizing band was cloned as a 3·3 kb *Eco*RI–*Hpa*II fragment into the *Eco*RI and *Acc*I sites (after calf intestine alkaline phosphatase treatment) of pUC19, resulting in pUC19CspA/B (Table 1). The second hybridizing band was cloned as a *Hin*dIII–*Bgl*II fragment (2·1 kb) in the *Hin*dIII- and the *Bam*HI-sites of pUC19 (pUC19CspC/D; Table 1) and sequenced by primer walking. Attempts to clone the third hybridizing fragment either as a 3·5 kb *Hin*dIII fragment or as a 4·5 kb *Eco*RI–*Sac*I fragment in both a high-copy (pUC19) and a low-copy vector (pNZ84, a pACYC derivative; Van Alen-Boerrigter *et al*., 1991) failed. The fourth hybridizing

Table 1. Plasmids

Plasmid	Characteristics
pUC18PCR1	pUC18 containing a PCR fragment (PCR1) of about 200 bp obtained with primers 1 and 2
pGEM-TPCR2	pGEM-T containing a PCR fragment (PCR2) of about 550 bp obtained with primers 3 and 4
pUC19CspA/B	pUC19 containing a 3·3 kb *Eco*RI–*Hpa*II fragment including *cspA*/*cspB*
pUC19CspC/D	pUC19 containing a 2·1 kb *Hin*dIII–*Bgl*II fragment including *cspC*/*cspD*
pUC19CspE	pUC19 containing a 1·1 kb *Hin*dIII–*Pst*I fragment including *cspE* and its upstream region
pLEX	pUC18 containing a 0·95 kb PCR fragment which is cloned in the *Sau*3AI and *Hin*cII sites including *cspE* and its downstream region

Table 2. Oligonucleotides

Oligonucleotide	Sequence (5′–3′)
Primer 1	CGGAATTCGGIA(A/T)IGTIAA(A/G)TGGTT(T/C)AA
Primer 2	CGGAATTCGTIAC(A/G)TTIGCIGC(C/T)TGIGGICC
Primer 3	GGNANNGTNAA(A/G)TGGTT(C/T)AA
Primer 4	(G/A/T)AT(A/G)AANCC(A/G)AANCC(C/T)TT
pAMILEX	GAACGCAATGAGTCCTG
pAMI4	TGACAGCGGGCCTAACC
PE*cspA*	GCCATAGCCTTGTCCATATTG
PE*cspB*	GCCAAATCCTTTATCTGGA
PE*cspC*	CTTGCATATCATCTGCCA
PE*cspD*	ACCAAATCCTTTAGTAGC
PE*cspE*	TGTGCGAAAACGTCGTTT

band was cloned as a *Hin*dIII–*Pst*I fragment (1·1 kb) in the *Hin*dIII and *Pst*I sites of pUC19 (pUC19CspE; Table 1). On this fragment a putative *csp* gene was located; to obtain its downstream region the following inverse PCR strategy was used. Chromosomal DNA was digested with *Hpa*II and self-ligated. PCR was performed with this template and with pAMILEX and pAMI4 (Table 2) as primers. A 950 bp fragment was obtained which was cloned in the blunt *Hin*cII and the *Bam*HI site (compatible with *Sau*3AI) of pUC18 after digestion with *Sau*3AI (resulting in pLEX; Table 1).

DNA and deduced protein sequence analysis. Computer analysis of DNA sequences and the deduced amino acid sequences was performed with the programs PC/GENE (version 6.70; IntelliGenetics) and Clone (Version 4.0; Clone Manager). The EMBL/GenBank and SWISS-PROT/PIR databases were used to search for amino acid sequence similarities.

RNA techniques and primer extension experiments. RNA isolation, Northern blotting and subsequent hybridization with radiolabelled probes was performed as described previously (Kuipers *et al*., 1993). For cold-shock experiments, cultures were grown at 30 °C to mid-exponential phase, after which they were spun down by centrifugation and resuspended in medium precooled to 10 °C. After exposure to 10 °C for various time periods (0, 0·5, 1, 2, 4 and 24 h) total RNA was isolated. The same oligonucleotides were used as probes in Northern blotting and as primers in primer extension experiments (PE*cspA* to PE*cspE*; Table 2). Quantification of the *csp* transcripts in Northern blotting was performed using the Dynamics Phosphor Imaging System. Cross-hybridization of the probes to the other *csp* genes was checked using Southern blotting, quantified with the same system. As a control for the RNA quantity the *usp45* gene, which is constitutively expressed (Van Asseldonk *et al*., 1990), was used and correction factors were calculated by using the Phosphor Imaging System. Primer extension experiments on the *csp* genes were carried out as described previously (Kuipers *et al*., 1993), with the same RNA samples as used for Northern blotting. The resulting cDNA was subjected to electrophoresis alongside nucleotide sequencing ladders generated with the same primers using the dideoxy chain-termination method (Sanger *et al*., 1977) and [α-^{32}P]dATP as radiolabel.

RESULTS

Cloning of genes encoding putative CSPs

Using different primers based on the homologous sequences of CspA^{E} and CspB^{B} (Table 2) two PCR products of about 200 bp (PCR1) and 550 (PCR2) bp

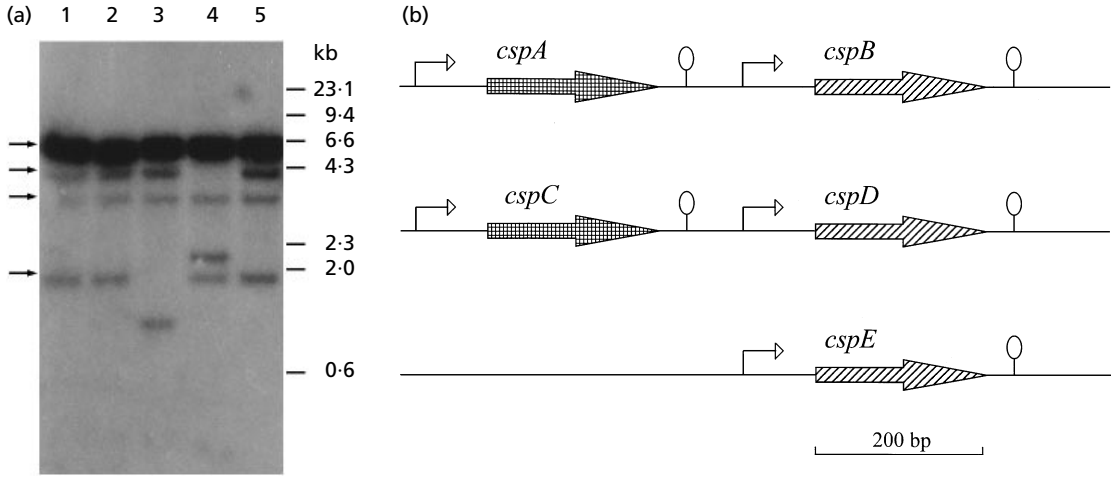

Fig. 1. (a) Southern hybridization, with PCR1 used as a probe, of chromosomal DNA of *L. lactis* MG1363 digested with *Hin*dIII (lane 1), *Hin*dIII and *Eco*RI (lane 2), *Hin*dIII and *Pst*I (lane 3), *Hin*dIII and *Bgl*II (lane 4), and *Hin*dIII and *Bam*HI (lane 5). Marker sizes are indicated on the right, and arrows indicate the hybridizing *Hin*dIII fragments. (b) Organization and nomenclature of the *csp* genes found in *L. lactis* MG1363. The large arrows indicate the ORFs, the smaller arrows indicate the transcription starts and the major terminators are indicated by a hairpin structure.

were amplified with *L. lactis* MG1363 chromosomal DNA as a template. After cloning and sequencing it appeared that these PCR products contained parts of genes homologous to the major *csp* genes. In a Southern-blotting experiment, using PCR1 as a probe, four hybridizing fragments were detected in different digests of *L. lactis* chromosomal DNA (Fig. 1a). Two *csp* genes, named *cspA* and *cspB*, are located on an *Eco*RI–*Hpa*II fragment (cloned in pUC19 resulting in pUC19CspA/B). Another fragment (cloned in pUC19 resulting in pUC19CspC/D) also contained two *csp* genes (named *cspC* and *cspD*), organized in a tandem repeat. A single *csp* gene, named *cspE* (cloned in pUC19 resulting in pUC19CspE), is located on a *Hin*dIII–*Pst*I fragment and its downstream region was cloned by an inverse PCR strategy (pLEX). The organization of the different *csp* genes is shown in Fig. 1(b). In Southern hybridization with PCR2 as a probe only two fragments, identical to fragments that hybridized with PCR1 as a probe, could be detected (data not shown). When the different *csp* genes were used as probes in Southern hybridization (different chromosomal-DNA digests) no extra hybridizing bands could be detected compared to the four bands obtained when using PCR1 as a probe. In an *Eco*RI digest all *csp* homologues were located on only two fragments, indicating a clustered organization on the *L. lactis* MG1363 chromosome. No hybridization was observed using plasmid DNA (isolated from several *L. lactis* strains) and PCR1 as a probe, indicating that these *csp* genes are chromosomally encoded and that no homologues are located on plasmids (data not shown).

A remarkably high nucleotide sequence identity was found for the two tandem repeats: 79% over 800 nt containing both ORFs. In the tandem repeats the first ORFs (*cspA* and *cspC*) and the second ORFs (*cspB* and *cspD*) are highly similar (81% and 82% identity, respectively). Also the spacing between the two adjacent ORFs is similar for both tandem repeats (268 nt for *cspA* and *cspB*, 277 nt for *cspC* and *cspD*).

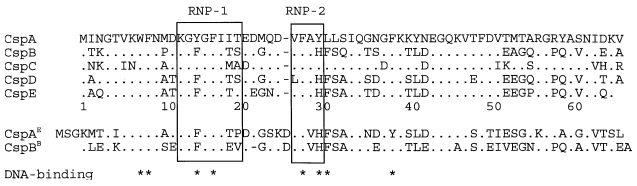

Fig. 2. Alignment of the deduced amino acid sequences of lactococcal CSPs (CspA–CspE) and the amino acid sequences of $CspA^E$ and $CspB^B$. Identical amino acids are indicated with dots; gaps are indicated with dashes. Important regions for DNA binding are indicated with asterisks, and RNA-binding motifs (RNP-1 and RNP-2) are boxed.

The five CSPs of *L. lactis* have a mutual identity of 52–85% at the amino acid level. The identity to the major CSPs, $CspA^E$ and $CspB^B$, is about 45–65% and is lowest for CspA and CspC (Fig. 2, Table 3). The calculated molecular masses of the *L. lactis* CSPs range from 7·1 kDa for CspE to 7·6 kDa for CspA and CspC (Table 3). CspA and CspC have an unusually high pI (approximately 9) compared to other CSPs (approximately 5).

Cold induction of *csp* genes

Cells of *L. lactis* were cultured to the mid-exponential phase at 30 °C, after which they were subjected to a cold shock by resuspending in precooled GM17 medium (10 °C). The growth characteristics of the cold-shocked

Table 3. Identity (%), size of ORF in amino acids, molecular mass and pI of the *L. lactis* CSPs, $CspA^E$ and $CspB^B$

	Identity (%)							Size (aa)	Mol. mass (kDa)	pI
	CspA	CspB	CspC	CspD	CspE	$CspA^E$	$CspB^B$			
CspA	·	62	76	59	60	45	50	66	7·6	9·2
CspB		·	56	80	82	59	62	66	7·3	4·9
CspC			·	52	52	48	47	66	7·6	9·6
CspD				·	85	64	65	66	7·2	4·4
CspE					·	61	63	65	7·1	4·6
$CspA^E$						·	61	70	7·4	5·9
$CspB^B$							·	67	7·4	4·3

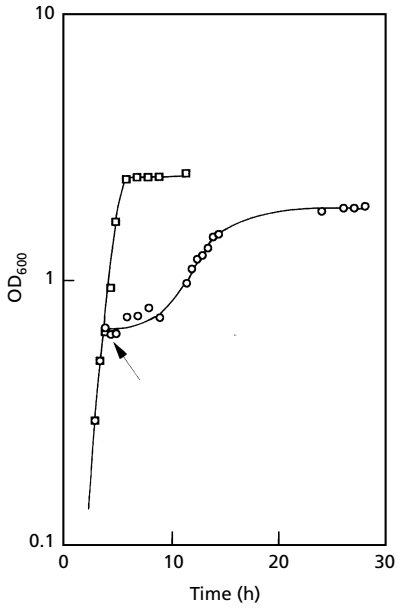

Fig. 3. Growth of *L. lactis* MG1363 at 30 °C (squares) and after cold shock to 10 °C (circles). The arrow indicates the time point of cold shock.

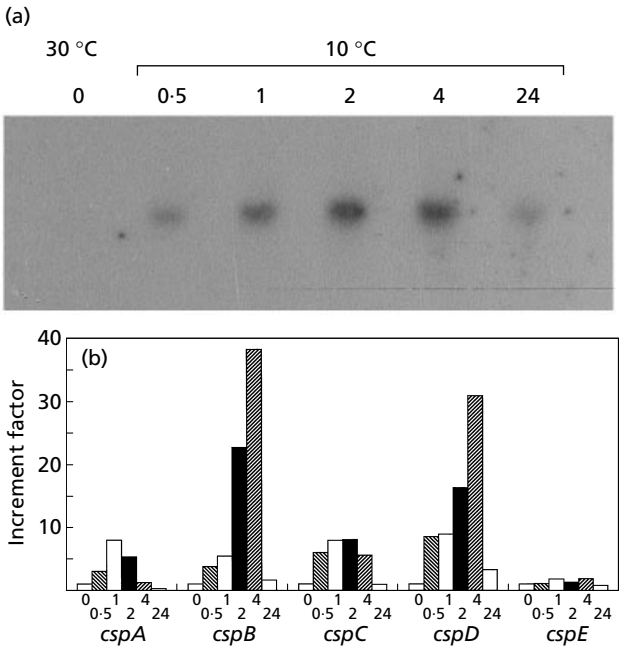

Fig. 4. (a) Northern blot of RNA extracted at 0, 0·5, 1, 2, 4 and 24 h after cold shock, hybridized with a probe specific for *cspB* (Table 2). The transcript size is about 300 nt. (b) Increase in mRNA levels at different times after cold shock relative to $t = 0$ (30 °C). Correction for mRNA amounts was performed using *usp45* (Van Asseldonk *et al.*, 1990) as a standard.

culture are shown in Fig. 3. A lag time of about 6–8 h after cold shock was observed, after which exponential growth was resumed with a lower growth rate (sixfold reduction) as compared to 30 °C. The amount of mRNA of the *csp* genes was monitored by Northern blotting at various times after cold shock (Fig. 4a; only shown for *cspB*). Probes specific for each cold-shock gene were used (Table 2); the cross-hybridization for all probes was calculated and appeared to be maximally 6% with primer PE*cspE* and *cspC* (data not shown). Transcripts of about 300 nt were detected for all *csp* genes, whereas for *cspA* and *cspC* larger transcripts (about 450 and 350 nt, respectively) were also detected in small amounts (<5%; see below).

The results of the Northern blotting of the *csp* genes, using *usp45* as an internal control, established that *cspB* and *cspD* were induced about 40- and 30-fold, respectively, at 10 °C, whereas c*spA* and *cspC* were induced about 10-fold compared to the level at 30 °C (Fig. 4b). At 30 °C ($t = 0$) a relatively high *cspE* mRNA level was detected compared to the other four *csp* genes, but *cspE* seemed not to be induced significantly at low temperature (Fig. 4b). Strikingly, the time at which maximal mRNA levels were found was different for the cold-induced *csp* genes. *cspA* and *cspC* reached maximal accumulation at 1–2 h after cold shock whereas for *cspB* and *cspD* maximal accumulation occurred at about 2–4 h after cold shock (Fig. 4b). The mRNA levels of *cspA*, *cspB*, *cspC* and *cspD* were decreased at 8 h after cold shock (data not shown), when exponential growth was resumed (Fig. 3). Other stress conditions, such as

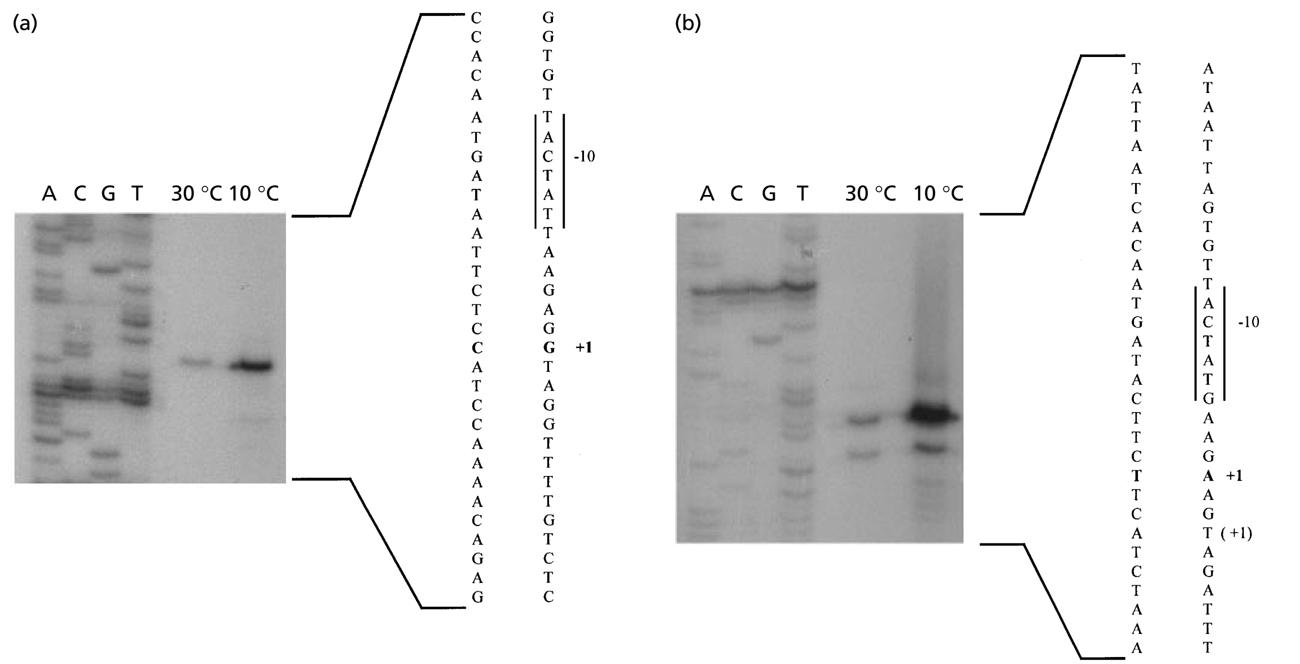

Fig. 5. Primer extension experiments for (a) *cspC* and (b) *cspD*. Sequence ladders are indicated on the left. RNA samples were taken at the mid-exponential growth phase at 30 °C and at 2 h after cold shock to 10 °C. The nucleotide sequences, the −10 promoter regions and the transcription starts are indicated on the right.

heat stress (10 min 42 °C), salt stress (10 min 0·5 M NaCl), low-pH stress (10 min pH 4·0, adjusted with lactic acid) or stationary-phase conditions (2 h after reaching stationary phase) did not result in increased mRNA levels of any of the *csp* genes (data not shown).

Identification of promoter regions

Using the primer extension technique, transcription start points of the *csp* genes were identified (Fig. 5; only shown for *cspC* and *cspD*); they are indicated in Fig. 6. For *cspD* a double transcription start was found: a major start at the indicated A-residue and a minor start at the T-residue three bases downstream. For each *csp* gene, transcripts were detectable at 30 °C, and for *cspA*, *cspB*, *cspC* and *cspD* increased amounts of transcript were found at 10 °C. The same transcription start points were identified at high and low temperature. Northern blotting showed that the mRNA size for the different *csp* genes is about 300 nt, which corresponds well with the detected transcription starts and the putative terminators [$\Delta G = -6$, -10, -8, -8 and -8 kcal mol^{-1} (-25, -42, -33, -33 and -33 kJ mol^{-1}) for *cspA*, *cspB*, *cspC*, *cspD* and *cspE*, respectively]. For *cspA* and *cspC* hairpin structures [$\Delta G = -10$ and -14 kcal mol^{-1} (-42 and -59 kJ mol^{-1}), respectively] were found further downstream the ORFs, for which the size of the mRNA corresponds with larger transcripts that were detected in small amounts (only detected after prolonged exposure of the blots to X-ray films). When DNA fragments containing parts of the *csp* genes and the region between the clustered *csp* genes were used as probes, only transcripts of about 300 nt were detected, indicating that the *csp* genes located in tandem repeats are monocistronic.

```
                       -35                    -10         ↓
                   +  **+** *             ** ** +**  *   ***+ **** **+  +*++
cspA  (-45) AAAGCATAATTTGACTTTCTATCAGAGAGGTGGTACTATCAAGAGGTAGGTTTT-GTCGTGAGT
cspB  (-46) TTTTTATAGGTTGACATGTAACCTATATAATGATATTATAGACGAGTAGATTTT-GTCATGAGT
cspC  (-45) ATAGACCAATTTGACTTTGGTAAGAATAGGTGTTACTATTAAGAGGTAGGTTTT-GTCTCGAGT
cspD  (-43) AGTGAGTAGTTTGACATTGAGTATAATTAGTGTTACTATGAAGAAGTAGATTTT-GTCATGAGT
cspE  (-45) GATTTTCTACTTTACTTTCCATATTAACTATGATAACATGAAAGTGTAAGTTTTTGTTTT-ACG

                                                                     SD
              +*++** +     ++ ++***** + +++*+  ++  **++**  +  *+  ++++++*  *
cspA  (+19) TTAAATGTTTCAGGAATGCACTTTTTCGCAAAAATAGGATTTACAATTAATGC-GAAAGGATGT
cspB  (+18) -TAAATGTATATTATATTCACTTTT-CACAAAATAAGG-TTTACAATTTTTGTTGAAAGGAAAT
cspC  (+19) TTAAATGTTTTAGGAATGCACTTTTTCGCAAAA-TAGGATTTACAGATGATGC-GAAAGGATCT
cspD  (+21) -AAAATGTTTTATAGATTCACTTTT-CGCAAAATGAG-ATTTACAATTTTTGTCGAAAGGAAAT
cspE  (+19) ATTA--GTTCTGTG------CTTTTGGGTTCATTCCT-ATT--CAAGAAATTAGATTTACAAAT

            +
cspA  (+81) ATTAAGATGATAAATGGAACAGTAAAATGGTTCAATATGG
cspB  (+79) AATTAATATGACAAAAGGAACTGTAAAATGGTTTAATCCA
cspC  (+80) ATTATATGAATAAAGGAACAATAAATTGGTTTAACGCCGA
cspD  (+82) AAATATTATGGCAAATGGAACAGTAAAATGGTTTAACGCT
cspE  (+73) TTATAAAATAAGGAAGTACCAAAAATGGCACAAGGAACTG
```

Fig. 6. Alignment of the nucleotide sequence of the *csp* promoters and the 5′ UTR sequences of the lactococcal *csp* genes. The translated regions are indicated in bold; SD-sequences are double-underlined; transcription starts are underlined and in bold (and indicated with an arrow); −35 and −10 regions are double-underlined and in bold; TG dinucleotides (−16 region) are underlined; start codons are underlined, and in bold italic; Y-box motifs are double underlined and in italic. Identical nucleotides in all five *csp* genes are indicated with *; identical nucleotides in the four cold-induced *csp* genes only are indicated with +. Nucleotides are numbered from their +1 transcription starts.

The detected transcription start sites allowed identification of −35 and −10 promoter regions of the *csp* genes (Fig. 6). The promoter regions are 67–92 %

identical to the established consensus sequences of *L. lactis* (De Vos & Simons, 1994). The consensus 17 bp spacing between the −35 and −10 regions is found for all lactococcal *csp* promoters (Fig. 6). The non-cold-induced *cspE* gene has the lowest similarity (4 nt mismatches) with the consensus promoters, whereas the promoter regions of the cold-induced *csp* genes are less different from the consensus promoter regions (3, 3, 2 and 1 nt mismatches for *cspA*, *cspC*, *cspD* and *cspB*, respectively). In the promoter regions of *cspC* and *cspB* complementary sequences (CCAAT) of the Y-box motifs (ATTGG) are present (Fig. 6). Several of these motifs were also found further up- and downstream of the promoter regions of the other lactococcal *csp* genes.

The 5′-untranslated leader regions (5′ UTRs) of the cold-induced *cspA*, *cspB*, *cspC* and *cspD* genes are highly identical (approximately 60 %) whereas the identity with this region of the non-cold-inducible *cspE* is much lower (about 30 % ; Fig. 6). Furthermore, the 5′ UTR of *cspE* (94 nt) is slightly longer than those of the other lactococcal *csp* genes (86, 84, 83 and 87 nt for *cspA*, *cspB*, *cspC* and *cspD*, respectively). The 5′ UTRs of all lactococcal *csp* genes appear to be rich in secondary structure, encompassing the entire region as calculated by the method of Zuker & Stiegler (1981).

DISCUSSION

A family of five genes, named *cspA*, *cspB*, *cspC*, *cspD* and *cspE*, encoding putative CSPs was cloned from *L. lactis* MG1363 and it appeared that these *csp* genes were organized in clusters. *cspA* and *cspB* as well as *cspC* and *cspD* are located in a tandem repeat whereas *cspE* was found as a single gene. To our knowledge, a clustered organization of *csp* genes has never been observed before (Graumann *et al*., 1996; Lee *et al*., 1994; Mayo *et al*., 1997; Mayr *et al*., 1996; Willimsky *et al*., 1992). *cspB* is identical to the *cspB* gene that was recently obtained from *L. lactis* AM2 using an inverse PCR strategy (Chapot-Chartier *et al*., 1997).

The five *csp* genes can be grouped based on sequence analysis: a group consisting of *cspA* and *cspC* (the first genes in the tandem repeats); and a group consisting of *cspB*, *cspD* and *cspE*. Members within these groups code for highly similar proteins (about 80 % identity) whereas the identity between these two groups is only about 55 %. High similarity (45–65 % identity) was also observed with the sequences of the major CSPs $CspA^{E}$ and $CspB^{B}$, and was lowest for CspA and CspC. The residues important for single-stranded DNA binding of $CspA^{E}$ and $CspB^{B}$ (Newkirk *et al*., 1994; Schröder *et al*., 1995) are highly conserved in CspB, CspD and CspE, whereas in CspA and CspC some additional residues are different from the $CspA^{E}$ and $CspB^{B}$ DNA-binding residues. The RNA-binding RNP-1 (consensus KGFGF) and RNP-2 (consensus VFVH) motifs (Jones & Inouye, 1994; Schindelin *et al*., 1993; Schröder *et al*., 1995) are also found in the *L. lactis* CSPs although some differences are observed. Interestingly, the pI values of CspA and CspC (9·2 and 9·6, respectively) are much higher than those of the other CSPs (approximately 4·5) due to the presence of more basic residues (8 and 11 for CspA and CspC, respectively, compared to 7 for CspB, CspD and CspE) and the presence of 4 tyrosine residues for CspA and CspC and no tyrosine residues in CspB, CspD and CspE. This high pI of CspA and CspC might result in an improved nucleic acid binding capacity since these proteins do not need to overcome charge repulsion when approaching nucleic acids (Schröder *et al*., 1995). Furthermore, protein 3-D modelling based on the crystal structure of $CspA^{E}$ and $CspB^{B}$ (Schindelin *et al*., 1993, 1994) revealed a similar β-barrel structure formed by five β-strands for all five lactococcal CSPs (J. A. Wouters, unpublished results).

For all *csp* genes transcripts of about 300 bp were found and no combined transcripts were found for the *csp* genes located in tandem repeats. Furthermore, Northern blotting revealed increased mRNA levels for the *csp* genes at different times after cold shock, indicating that regulation of these genes takes place at the transcriptional level. Maximal induction of mRNA was approximately 40- and 30-fold for *cspB* and *cspD*, respectively, whereas the mRNA level of *cspA* and *cspC* increased approximately 10-fold. *cspE* was not induced at 10 °C. A differential expression in time after cold shock was observed for the different *csp* genes. mRNA levels of *cspA* and *cspC* increase shortly after cold shock (in the first 2 h) whereas *cspB* and *cspD* mRNA levels are highest at 4 h after cold shock. Possibly the more basic CSPs, CspA and CspC, are involved in the regulation of the expression of their counterparts CspB and CspD located further downstream. Since no mRNA induction was observed upon exposure to other stress conditions, such as heat, salt, low pH and stationary phase, it is concluded that these *csp* genes, with the exception of *cspE*, might play a specific role in low-temperature adaptation. Recently, it was shown that the non-cold-induced *cspD* gene of the *E. coli* CspA family is in fact induced under stationary-phase conditions (Yamanaka & Inouye, 1997).

Recent studies indicate that the 5′ UTR plays an important role in the stability of the *E. coli cspA* transcript (Fang *et al*., 1997) and the regulation of $CspA^{E}$ expression after cold shock (Jiang *et al*., 1996). Although the 5′ UTRs of the lactococcal *csp* genes are not as exceptionally long (83–94 nt) as this region of the *E. coli cspA* (159 nt; Goldstein *et al*., 1990), they might play a similar role. Most intriguing in this respect is the finding that the 5′ UTRs of the four cold-induced lactococcal *csp* genes are highly similar but that clear differences are observed in this sequence of the non-cold-induced *cspE* gene, indeed suggesting a regulatory function of this leader.

Future research will focus on the differential expression, the clustered organization and the regulation of the newly described *csp* genes in *L. lactis*. The physiological role of the *L. lactis* CSPs will be studied using single and multiple overexpression constructs and using strains with disrupted *csp* genes.

ACKNOWLEDGEMENTS

We thank Tomasz Oczkowski and Girbe Buist for their technical assistance with cloning of *cspE*.

REFERENCES

Brandi, A., Pietroni, P., Gualerzi, C. O. & Pon, C. L. (1996). Post-transcriptional regulation of CspA expression in *Escherichia coli*. *Mol Microbiol* **19**, 231–240.

Casadaban, M. J. & Cohen, S. N. (1980). Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J Mol Biol* **138**, 179–207.

Chapot-Chartier, M. P., Schouler, C., Lepeuple, A. S., Gripon, J. C. & Chopin, M. C. (1997). Characterization of *cspB*, a cold-shock-inducible gene from *Lactococcus lactis*, and evidence for a family of genes homologous to the *Escherichia coli cspA* major cold shock gene. *J Bacteriol* **179**, 5589–5593.

De Vos, W. M. & Simons, G. F. M. (1994). Gene cloning and expression systems in lactococci. In *Genetics and Biotechnology of Lactic Acid Bacteria*, pp. 52–105. Edited by M. J. Gasson and W. M. de Vos. London: Blackie Academic & Professional.

Fang, L., Jiang, W., Bae, W. & Inouye, M. (1997). Promoter-independent cold-shock induction of *cspA* and its derepression at 37 °C by mRNA stabilization. *Mol Microbiol* **23**, 355–364.

Gasson, M. J. (1983). Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic streptococci after protoplast-induced curing. *J Bacteriol* **154**, 1–9.

Goldstein, J., Pollitt, N. S. & Inouye, M. (1990). Major cold shock protein of *Escherichia coli*. *Proc Natl Acad Sci USA* **87**, 283–287.

Graumann, P. & Marahiel, M. A. (1994). The major cold shock protein of *Bacillus subtilis* CspB binds with high affinity to the ATTGG- and CCAAT-sequences in single stranded oligonucleotides. *FEBS Lett* **338**, 157–160.

Graumann, P., Schröder, K., Schmid, R. & Marahiel, M. A. (1996). Cold shock stress-induced proteins in *Bacillus subtilis*. *J Bacteriol* **178**, 4611–4619.

Graumann, P., Wendrich, T. M., Weber, M. H. W., Schröder, K. & Marahiel, M. A. (1997). A family of cold shock proteins in *Bacillus subtilis* is essential for cellular growth and for efficient protein synthesis at optimal and low temperatures. *Mol Microbiol* **25**, 741–756.

Jiang, W., Fang, L. & Inouye, M. (1996). The role of the 5′-end untranslated region of the mRNA for CspA, the major cold-shock protein of *Escherichia coli*, in cold-shock adaptation. *J Bacteriol* **178**, 4919–4925.

Jiang, W., Hou, Y. & Inouye, M. (1997). CspA, the major cold-shock protein of *Escherichia coli*, is an RNA chaperone. *J Biol Chem* **272**, 196–202.

Jones, P. G. & Inouye, M. (1994). The cold shock response – a hot topic. *Mol Microbiol* **11**, 811–818.

Jones, P. G. & Inouye, M. (1996). RbfA, a 30S ribosomal binding factor, is a cold-shock protein whose absence triggers the cold-shock response. *Mol Microbiol* **21**, 1207–1218.

Jones, P. G., VanBogelen, R. & Neidhardt, F. C. (1987). Induction of proteins in response to low temperature in *Escherichia coli*. *J Bacteriol* **169**, 2092–2095.

Jones, P. G., Krah, R., Tafuri, S. R. & Wolffe, A. P. (1992). DNA gyrase, CS7.4, and the cold shock response in *Escherichia coli*. *J Bacteriol* **174**, 5798–5802.

Jones, P. G., Mitta, M., Kim, Y., Jiang, W. & Inouye, M. (1996). Cold shock induces a new major ribosomal-associated protein which unwinds double-stranded RNA in *Escherichia coli*. *Proc Natl Acad Sci USA* **93**, 76–80.

Kim, W. S. & Dunn, N. W. (1997). Identification of a cold shock gene in lactic acid bacteria and the effect of cold shock on cryotolerance. *Curr Microbiol* **35**, 59–63.

Kuipers, O. P., Boot, H. J. & De Vos, W. M. (1991). Improved site-directed mutagenesis method using PCR. *Nucleic Acids Res* **19**, 4558.

Kuipers, O. P., Beerthuyzen, M. M., Siezen, R. J. & De Vos, W. M. (1993). Characterization of the nisin gene cluster *nisABTCIPR* of *Lactococcus lactis*: requirement of expression of the *nisA* and *nisI* genes for development of immunity. *Eur J Biochem* **216**, 281–291.

Landsman, D. (1992). RNP-1, an RNA-binding motif is conserved in the DNA-binding cold shock domain. *Nucleic Acids Res* **20**, 2861–2864.

LaTeana, A., Brandi, A., Falconi, M., Spurio, R., Pon, C. L. & Gualerzi, C. O. (1991). Identification of a cold shock transcriptional enhancer of the *Escherichia coli* major cold shock gene encoding nucleoid protein H-NS. *Proc Natl Acad Sci USA* **88**, 10907–10911.

Lee, S. J., Xie, A., Jiang, W., Etchegaray, J., Jones, P. G. & Inouye, M. (1994). Family of the major cold-shock protein, CspA (CS7.4), of *Escherichia coli*, whose members show a high sequence similarity with the eukaryotic Y-box binding proteins. *Mol Microbiol* **11**, 833–839.

Mayo, B., Derzelle, S., Fernandez, M., Leonard, C., Ferain, T., Hols, P., Suarez, J. E. & Delcour, J. (1997). Cloning and characterization of *cspL* and *cspP*, two cold-inducible genes from *Lactobacillus plantarum*. *J Bacteriol* **179**, 3039–3042.

Mayr, B., Kaplan, T., Lechner, S. & Scherer, S. (1996). Identification and purification of a family of dimeric major cold shock protein homologs from the psychrotrophic *Bacillus cereus* WSBC10201. *J Bacteriol* **178**, 2916–2925.

Nakashima, K., Kanamaru, K., Mizuno, T. & Horikoshi, K. (1996). A novel member of the *cspA* family of genes that is induced by cold shock in *Escherichia coli*. *J Bacteriol* **178**, 2994–2997.

Newkirk, K., Feng, W., Jiang, W., Tejero, R., Emerson, S. D., Inouye, M. & Montelione, G. T. (1994). Solution of NMR structure of the major cold shock protein (CspA) from *Escherichia coli*: identification of a binding epitope for DNA. *Proc Natl Acad Sci USA* **91**, 5114–5118.

Panoff, J.-M., Legrand, S., Thammavongs, B. & Boutibonnes, P. (1994). The cold shock response in *Lactococcus lactis* subsp. *lactis*. *Curr Microbiol* **29**, 213–216.

Rallu, F., Gruss, A. & Maguin, E. (1996). *Lactococcus lactis* and stress. *Antonie Leeuwenhoek* **70**, 243–251.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sanders, J. W., Leenhouts, K., Haandrikman, A. J., Venema, G. & Kok, J. (1995). Stress response in *Lactococcus lactis*: cloning, expression analysis and mutation of the superoxide dismutase gene. *J Bacteriol* **177**, 5254–5260.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463–5467.

Schindelin, H., Marahiel, M. A. & Heinemann, U. (1993). Universal nucleic acid-binding domain revealed by crystal structure of the *B. subtilis* major cold-shock protein. *Nature* **364**, 164–168.

Schindelin, H., Jiang, W., Inouye, M. & Heinemann, U. (1994). Crystal structure of CspA, the major cold shock protein of *Escherichia coli*. *Proc Natl Acad Sci USA* **91**, 5119–5123.

Schröder, K., Graumann, P., Schnuchel, A., Holak, T. A. & Marahiel, M. A. (1995). Mutational analysis of the putative nucleic acid-binding surface of the cold-shock domain, CspB, revealed an essential role of aromatic and basic residues in binding of single-stranded DNA containing the Y-box motif. *Mol Microbiol* **16**, 699–708.

Van Alen-Boerrigter, I. J., Baankreis, R. & De Vos, W. M. (1991). Characterization and overexpression of the *Lactococcus lactis pepN* gene and localization of its product, aminopeptidase N. *Appl Environ Microbiol* **57**, 2555–2561.

Van Asseldonk, M., Rutten, G., Oteman, M., Siezen, R. J., De Vos, W. M. & Simons, G. (1990). Cloning, expression in *Escherichia coli* and characterization of *USP45*, a gene encoding a highly secreted protein from *Lactococcus lactis* MG1363. *Gene* **95**, 155–160.

Van Asseldonk, M., Simons, A., Visser, H., De Vos, W. M. & Simons, G. (1993). Cloning, nucleotide sequence and regulatory analysis of the *Lactococcus lactis dnaJ* gene. *J Bacteriol* **175**, 1637–1644.

Vos, P., Van Asseldonk, M., Van Jeveren, F., Siezen, R. J., Simons, G. & De Vos, W. M. (1989). A maturation protein is essential for the production of active forms of *Lactococcus lactis* SK11 serine proteinase located in or secreted from the cell envelope. *J Bacteriol* **171**, 2795–2802.

Wells, J. M., Wilson, P. W. & Le Page, R. W. F. (1993). Improved cloning vectors and transformation procedure for *Lactococcus lactis*. *J Appl Bacteriol* **74**, 629–636.

Willimsky, G., Bang, H., Fischer, G. & Marahiel, M. A. (1992). Characterization of *cspB*, a *Bacillus subtilis* inducible cold shock gene affecting cell viability at low temperatures. *J Bacteriol* **174**, 6326–6335.

Yamanaka, K. & Inouye, M. (1997). Growth-phase-dependent expression of *cspD*, encoding a member of the CspA family in *Escherichia coli*. *J Bacteriol* **179**, 5126–5130.

Zuker, M. & Stiegler, P. (1981). Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res* **9**, 133–148.

Received 16 January 1998; revised 15 June 1998; accepted 7 July 1998.