Journal of Industrial Microbiology & Biotechnology (1997) 19, 286–293

Detection and speciation of bacteria through PCR using universal major cold-shock protein primer oligomers

KP Francis and GSAB Stewart

Department of Applied Biochemistry and Food Science, University of Nottingham, Leicestershire, LE12 5RD, UK

The detection of bacteria using PCR is a well-established diagnostic technique. However, conventional PCR requires the use of DNA primer oligomers that are specific to the target organism and, as a consequence, a sample can only be tested for the presence of that specific target. A significant advantage would be to probe a sample for the presence of any bacteria, followed by identification. To achieve this it is necessary to identify a DNA sequence common to all bacteria. Here we demonstrate that such a sequence may be that encoding the major cold-shock proteins. Using two universal PCR primer oligomers from conserved regions of these gene homologues, we have amplified a 200 base-pair DNA sequence from more than 30 diverse Gram-positive and Gram-negative bacteria, including representatives from the genera *Aeromonas, Bacillus, Citrobacter, Enterobacter, Enterococcus, Escherichia, Klebsiella, Lactobacillus, Lactococcus, Listeria, Pediococcus, Photobacterium, Proteus, Salmonella, Shigella, Staphylococcus, Streptococcus,* and *Yersinia*. Sequence analysis of the amplified products confirmed a high level of DNA homology. Significantly, however, there are sufficient nucleotide variations to allow the unique allocation of each amplified sequence to its parental bacterium.

Keywords: major cold-shock protein; PCR; universal primer oligomers; bacterial detection

Introduction

Exposure of bacterial cells to an environmental stress often leads to the synthesis of proteins that are essential for that organism's survival. Such adaptive responses have been well documented in *Escherichia coli* and include a number of studies on low temperature adaptation and the cold-shock response [reviewed in Ref. 20]. When an exponential culture of *E. coli* is shifted from 37°C to 10°C, a novel set of proteins is induced. Identification of a number of these cold-shock proteins, of which there are at least thirteen, revealed several previously characterized polypeptides that are involved in a range of diverse cellular processes, plus one protein that is unique to cold-shock [19]. This novel 70 amino acid polypeptide was shown to be induced 300-fold (13% of new protein synthesized) following a shift from 37°C to 10°C and was subsequently termed the major cold-shock protein (MCSP), or CspA [10].

CpsA is part of a family of proteins consisting of nine members [5,10,13,20,23,28], at least three of which are cold-inducible [13,23,28]. All nine proteins are relatively small (69–74 amino acids) and are highly similar to each other. MCSP homologues have also been identified in a number of other bacteria, including psychrotrophic, mesophilic and thermophilic *Bacillus* [16,26,33,39], *Salmonella typhimurium* [22,35], *Streptomyces clavuligerus* [3], *Listeria monocytogenes* [11], *Haemophilus influenzae* [9], *Stigmatella aurantiaca* [36], *Lactobacillus plantarum* [25], *Arthrobacter globiformis* [4] and *Mycobacterium tuberculosis* [7], as well as in species of *Pseudomonas* and *Micrococcus* [17,27,30]. Furthermore, as in *E. coli* MCSP families have recently been identified in *Bacillus cereus*, *Bacillus subtilis* and *Pseudomonas fragi* [16,26,27]. Two of the six identified MCSPs of *B. cereus* were found to be cold-inducible [26], with at least three cold-inducible MCSPs identified in *B. subtilis* [16]. In *P. fragi*, four MCSPs have been identified, two that are cold-inducible and two that are heat-inducible [27].

Database analysis [1], has shown that these highly conserved bacterial MCSPs share regions of striking similarity (over 40% identity in most cases) with the nucleic acid-binding domain of the eukaryotic Y-box proteins [39,40]. This domain, designated as the cold-shock domain, comprises the full length of bacterial MCSPs, but forms less than a third of Y-box proteins (approximately 80 amino acids, usually located at the N-terminal region of the 300–400 residues). Both sets of prokaryotic and eukaryotic proteins have been shown to preferentially bind to single-stranded DNA sequences containing an ATTGG/CCAAT motif [8,26,31,37,42] and contain two RNA-binding motifs, RNP-1 (KGFGF) and a partial RNP-2 (VFVH) [21,32]. Two-dimensional nuclear magnetic resonance spectroscopy of the MCSP CspB from *B. subtilis* has shown that these motifs are close to the proposed DNA-binding surface and play a key role in nucleic acid recognition and binding of the cold-shock domain [15]. The fact that proteins containing RNP motifs are not only found in bacteria, but in nearly all the organelles of animals, plants and fungi where RNA is present [6], suggests that these motifs are ancient protein structures of fundamental importance.

In this study we show, using a single pair of degenerate PCR primer oligomers that were derived from the cold-shock domains of both prokaryotic and eukaryotic organisms [11,13,23,37,39], that it is possible to amplify MCSP gene sequences from a large number of diverse Gram-posi-

Correspondence: GSAB Stewart, Department of Applied Biochemistry and Food Science, University of Nottingham, Leicestershire, LE12 5RD, UK
Received 14 April 1997; accepted 25 August 1997

tive and Gram-negative bacteria. Moreover, we demonstrate that although these DNA sequences are highly homologous there is sufficient nucleotide variation to allow the unique allocation of each amplified product to its parental bacterium. It is proposed that the MCSP genes provide a ubiquitous PCR target for the rapid detection and identification of bacteria, a finding that might prove useful for the screening of environmental, food and clinical samples for bacterial contamination.

Materials and methods

Bacterial strains

All bacteria used in this study were pure strains that had been previously characterized. Each strain was cultured from a frozen stock onto selective media (Table 1) from which a sample was taken to perform PCR directly from bacterial cells.

Oligonucleotide primers and PCR conditions

Degenerate oligonucleotide primers which were used for the amplification of partial MCSP gene sequences (Figure 1), were derived from the cold-inducible MCSPs [1] of *E. coli cspA* and *B* [13,23], *B. subtilis cspB* [39], *L. monocytogenes cspL* [11] and the eukaryotic gene sequence encoding the Y-box protein, FRG Y1, from *Xenopus laevis* [37]. Composite oligonucleotide sequences were as follows: CSPU5 – 5′CCC GAA TTC GGT A(ATC)A GTA AAA TGG TT(TC) AAC (GT)C 3′ and CSPU3 – 5′ CCC GGA TCC GGT TAC GTT A(GC)C (AT)GC T(TG)(CG) (ACT)GG (TGA)CC 3′, where bracketed nucleotides show degeneracies used and the underlined sequences correspond to restriction enzyme sites for *Eco*RI (CSPU5) and *Bam*HI (CSPU3). PCR was performed with a Techne Progene automated thermocycler with 0.5-ml thin-walled PCR tubes (Anachem, Luton, UK). Reactions were carried out in 50-μl volumes containing 5 μl of 10× PCR buffer (supplied wth *Taq* DNA polymerase; Advanced Biotechnologies, Epsom, UK), 1.8 mM $MgCl_2$, 100 pmol of the oligonucleotide primers CSPU5 and CSPU3, 0.2 mM of each deoxynucleotide triphosphate (dATP, dCTP, dGTP, dTTP; Pharmacia, St Albans, UK), 1 U of *Taq* DNA polymerase (Advanced Biotechnologies), and a pin head-sized aliquot of bacteria picked from an agar plate of selective media

Table 1 Bacterial strains used to perform PCR with the universal primer oligomers CSPU3 and CSPU5 for the detection of MCSP gene sequences

Bacterial strain	Isolate	GS	Selective agar (Oxoid Ref. No.)
Aeromonas hydrophila	EI	Neg	*Aeromonas* (CM833+SR136)
Aeromonas salmonicida	NCIMB1102	Neg	*Aeromonas* (CM833+SR136)
Bacillus cereus	EI	Pos	*B. cereus* (CM617+SR99)
Bacillus subtilis	NCTC3610	Pos	*B. cereus*-Neg (CM617+SR99)
Bacillus subtilis	EI	Pos	*B. cereus*-Neg (CM617+SR99)
Bacillus megaterium	NCTC	Pos	*B. cereus*-Neg (CM617+SR99)
Citrobacter freundii	NCTC6071	Neg	MacConkey No. 3 (CM115)
Citrobacter ballerapensis	NCTC6081	Neg	MacConkey No. 3 (CM115)
Enterobacter aerogenes	NCTC10006	Neg	MacConkey No. 3 (CM115)
Enterococcus faecalis	NCTC775	Pos	KF *Streptococcus* (CM701)
Escherichia coli	MC4100	Neg	MacConkey No. 3 (CM115)
Klebsiella pneumoniae	NCTC9633	Neg	MacConkey No. 3 (CM115)
Lactobacillus casei	NCIB6375	Pos	Rogosa (CM627)
Lactococcus cremoris	NCIB8662	Pos	M17 (CM785)
Listeria innocua	NCTC11288	Pos	*Listeria*, Oxford (CM856+SR140)
Listeria monocytogenes	ATCC23074	Pos	*Listeria*, Oxford (CM856+SR140)
Listeria grayi	NCTC10815	Pos	*Listeria*, Oxford (CM856+SR140)
Pediococcus pentosaceus	EI	Pos	M17 (CM785)
Photobacterium mondopomensis	NCIMB841	Neg	Instant Ocean formulation
Photobacterium phosphoreum	NCIMB844	Neg	Instant Ocean formulation
Proteus vulgaris	NCTC4175	Neg	MacConkey No. 3 (CM115)
Salmonella typhimurium	LT2	Neg	MacConkey No. 3 (CM115), XLD (CM469)
Salmonella virchow	NCTC5742	Neg	MacConkey No. 3 (CM115), XLD (CM469)
Shigella boydii	NCTC9327	Neg	MacConkey No. 3 (CM115), XLD (CM469)
Shigella flexneri	NCTC0002	Neg	MacConkey No. 3 (CM115), XLD (CM469)
Staphylococcus aureus	RN4220	Pos	Baird Parker (CM275)
Staphylococcus epidermidis	NCTC11047	Pos	Baird Parker (CM275)
Staphylococcus intermedius	EI	Pos	Baird Parker (CM275)
Streptococcus dysgalactiae	NCTC4335	Pos	Blood agar No. 2 (CM271)
Streptococcus pyogenes	NCTC8198	Pos	Blood agar No. 2 (CM271)
Yersinia enterocolitica	NCTC10460	Neg	*Yersinia* (CM653+SR109)

Relative isolate indexes for bacteria obtained from National Collections of Type Cultures (NCTC), American Type Culture Collection (ATCC) and National Collections of Industrial and Marine Bacteria (NCIMB) are listed, as are Gram-stain type and selective growth media (Oxoid) used for bacterial culture. Instant Ocean formulation consisted of 33 g Instant Ocean (Aquarium Systems), 5 g peptone (Difco), 3 g yeast extract (Difco) and 15 g bacteriological agar (Oxoid), made-up to 1 L with deionized water. Environmental isolates (EI) of bacterial strains (routinely used as type strains in our laboratory) were gained from a number of different sources: *A. hydrophila* was gifted by Juan Tomas, University of Barcelona, Spain; *B. cereus* was isolated from dried egg powder derived from an egg processing plant; *B. subtilis* was isolated from flour derived from a bakery; *P. pentosaceus* was a gift of Mike Gasson, IFR, Norwich; and *S. intermedius* was isolated from a dried baby milk formulation. The identification of each bacterial strain was confirmed prior to PCR by the use of an appropriate API test strip (bioMerieux; data not shown). GS, Gram stain negative (neg) or positive (pos).

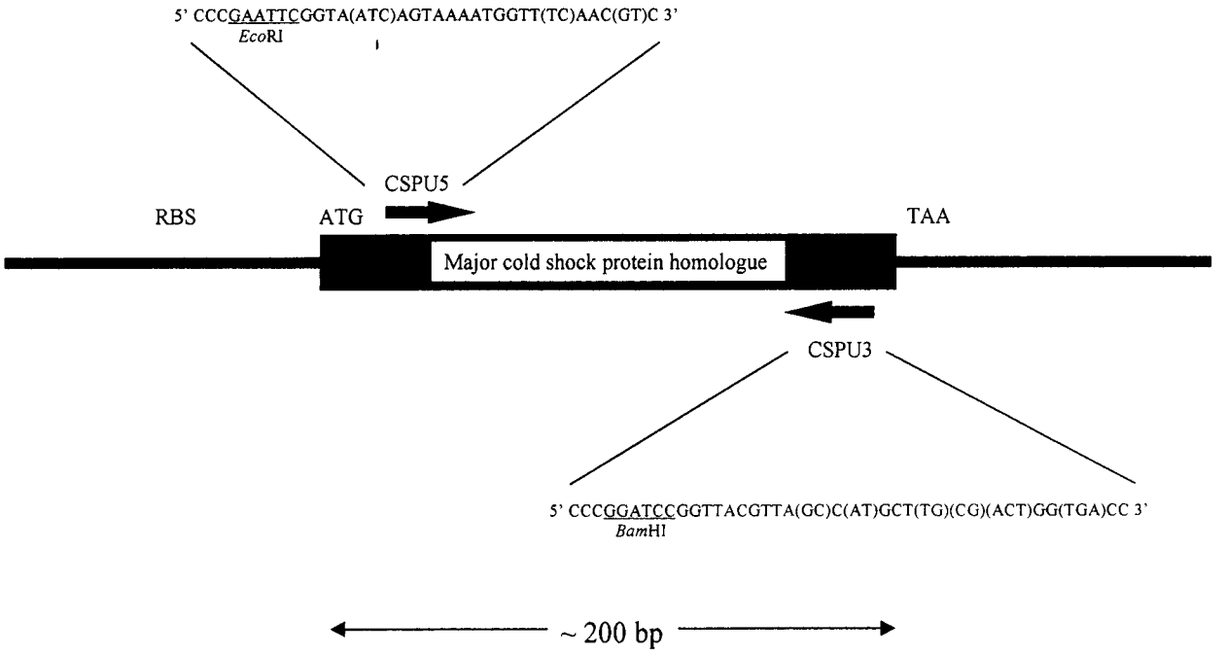

Figure 1 Schematic representation of the region of DNA amplified from bacterial MCSP genes using the universal PCR primer oligomers CSPU3 and CSPU5. Arrows indicate the approximate positions at which each primer hybridizes (approximately 5–15 nucleotides inside of either ATG or TAA [10,13,39]). Nucleotide sequences ATG and TAA indicate the positions of translational start and stop codons, respectively. RBS indicates the position of the ribosome-binding site.

(Table 1). Bacterial cells were lysed by heating the mixture for 5 min at 95°C. Amplification of MCSP DNA was then attempted with 25 cycles at 95°C for 15 s, 50°C for 30 s and 72°C for 30 s, followed by a final extension step at 72°C for 4 min.

Cloning and sequencing of PCR products

PCR products were analysed on a 1.6% agarose gel (NuSieve, FMC BioProducts, Rockland, USA). An appropriate volume of each was then run on a low melting point agarose gel (SeaPlaque GTG, FMC BioProducts) and 200-bp DNA fragments were band extracted using a freeze-thaw procedure [29]. Suspended DNA was digested with *Eco*RI and *Bam*HI (1 h at 37°C with 1 U of each enzyme; Boehringer Mannheim, Lewes, UK), the restriction enzyme sites tailing one or other oligonucleotide, followed by heat inactivation of these enzymes at 85°C for 30 min. Restriction enzyme-digested DNA was ligated overnight at room temperature (1 U T4 DNA ligase; Promega, Southampton, UK) with an equimolar ratio of similarly cut plasmid pT7 blue DNA (Novagen, Cambridge, UK). The mixture was dialysed and then electroporated into competent *E. coli* JM109 cells [2]. Transformed bacteria were selected on LB plates containing 50 μg ml^{-1} ampicillin, 0.004% 5′-bromo-4-chloro-3-indoyl-β-d-galactoside (X-gal; Boehringer Mannheim) and 0.004% isopropyl-β-d-thiogalactoside (IPTG; Boehringer Mannheim). Colonies appearing white were then screened using the PCR procedure described above, but using 100 pmol of the CSPU5 oligonucleotide primer in combination with 40 pmol of a universal M13 reverse oligonucleotide primer (GAC CAT GAT TAC GCC AAG CTT GC; PE Applied Biosystems, Manchester, UK). Bacterial colonies giving PCR products of approximately 250 bp (indicating ligation of the PCR product with the plasmid) were used to inoculate 10-ml volumes of LB. Plasmid DNA derived from 3 ml of an overnight culture of these *E. coli* clones was obtained using an alkaline lysis procedure (QIAprep Spin Miniprep Kit, Qiagen, Crawley, UK). Dialysed DNA was then sequenced with both universal M13 forward (CCC AGT CAC GAC GTT GTA AAA CGA C; PE Applied Biosystems) and reverse (see above) oligonucleotide primers using an ABI 373a sequencer (PE Applied Biosystems).

Phylogenetic analysis of data

Bacterial MCSP DNA sequences and their deduced amino acid sequences were compared using the Genetics Computer Group program PileUp (version 8, Wisconsin, USA; accessed via the BBSRC facility at Daresbury, UK). This program creates a multiple sequence alignment from a group of related sequences using progressive pairwise alignments. Dendrographic representations of PileUp data were created using the Genetics Computer Group programs Distances and Growtree (version 8, Wisconsin, USA; accessed via the BBSRC facility at Daresbury, UK). Growtree creates a phylogenetic tree from a distance matrix created by Distances, which used an unweighted pair group method based on arithmetic averages.

Results

Identification of MCSP sequences from 18 genera of bacteria

Using the universal degenerate oligonucleotides CSPU5 and CSPU3 (Figure 1), partial MCSP gene sequences (designated *cspA* homologues) were amplified directly in

less than 1 h from more than 30 diverse Gram-positive and Gram-negative bacteria. Database analysis [1] confirmed both the DNA sequences (Figure 2) and deduced amino acid sequences (Figure 3) amplified from each of the latter bacteria to be highly homologous to known MCSP sequences [3–5,7,9–11,13,16,23,25–28,33,35,36,39].

Although a range of annealing temperatures (40–60°C) were shown to amplify MCSP sequences from the majority of bacteria investigated, an annealing temperature of 50°C was found to be optimal since this allowed only a single 200-bp DNA band to be amplified from all bacteria tested. Moreover, in each case this amplification was achieved using one reaction mixture formulation and a single set of PCR conditions (see materials and methods). Direct sequencing of the latter PCR products gave clear sequence data in the majority of cases. However, approximately 20% of PCR products were found to give sequence chromatographs with a number of double peaks, indicating more than one partial MCSP gene sequence to be present. Hence, all MCSP sequences data shown in Figures 2 and 3 were derived from cloned PCR products. The number of thermal cycles for each of the PCR reactions was restricted to 25 to ensure that there was no spurious amplification of low level contamination. Increasing the number of cycles to 35 or greater, usually resulted in the amplification of a single faint product of approximately 200 bp from a control PCR reaction with no added bacteria. In most cases, sequencing of the amplified DNA fragment from this control identified *E. coli cspA* [13] (data not shown).

Nucleotide variation between MCSPs allows the unique allocation of each amplified sequence to its parental bacterium

Alignment of the partial MCSP DNA sequences amplified from the 31 species of bacteria using CSPU5 and CSPU3 (Figure 2), shows these genes to have considerable homology. In the majority of cases, however, there is sufficient nucleotide variation to allow the unique allocation of each amplified product to its parental bacterium. This variation between partial MCSP sequences can be most clearly seen when the above data are compared dendrographically and is apparent at both the DNA (Figure 4) and amino acid (data not shown) levels. Using this type of analysis, a clear distinction can be drawn between Gram-negative and Gram-positive bacteria, with strains that have been documented as being closely related (eg, bacterial strains from the genera *Escherichia*, *Shigella*, *Citrobacter*, *Enterobacter* and *Salmonella* [18]) adjacent in both dendrograms.

Most variation between MCSP sequence is seen in Gram-positive bacteria. Here, not only is it possible to distinguish between bacterial genera, but also between different bacterial species (eg, *S. aureus*, *S. intermedius*, and *S. epidermidis*). Moreover, comparison of the amplified MCSP sequences gained from two *B. subtilis* strains (NCTC3610 and an environmental isolate from bread) show a total of six differences in 136 nucleotides, demonstrating that it is even feasible to distinguish between strains, within species. These latter *Bacillus* sequences are almost identical to the published *B. subtilis cspB* sequence from strain JH642 [39], and show only three and five differences in 136 nucleotides to NCTC3610 and the above environmental isolate, respectively. Furthermore, since all three of the above *B. subtilis* nucleotide sequences encode protein sequences that differ by a single amino acid, it is almost certain that each encodes the equivalent protein (ie, *B. subtilis* CspB [39]).

At present it is not known whether each of the partial MCSP sequences is of the equivalent gene homologue (ie, encoding a protein with the exact same function) in each group of bacteria. However, since both universal MCSP oligonucleotides, CSPU5 and CSPU3, were derived from the bacterial sequences of cold-inducible homologues [10,13,23,39], there is a strong probability that these particular MCSPs were preferentially amplified. The latter hypothesis is strengthened by the fact that partial DNA sequences amplified from *E. coli*, *B. subtilis* and *L. monocytogenes* are all of cold-inducible MCSPs [10,13,39]. Possible exceptions to the above (ie, where a second MCSP homologue has been amplified in preference to *cspA*) are *B. megaterium* and *L. casei*, both of which are further displaced from established phylogenetic groupings than expected [18]. As more bacterial MCSP families are characterized it is likely that a clearer picture will be gained.

RNP motifs are highly conserved in bacterial MCSPs

Comparison of the deduced partial MCSP amino acid sequences gained from the 31 bacterial species (Figure 3), shows that the two RNA-binding motifs RNP-1 and RNP-2 [21,32] are present in the MCSPs of every bacteria tested. As with previously published MCSPs, the predominant amino acid sequences for these motifs are KGFGF for RNP-1 and VFVH for RNP-2. Exceptions to this are *L. casei* which has the RNP-1 motif KGYGF, and *S. disgalactii* and *S. pyogenes* which both have the RNP-2 motif VFAH. Schroder *et al* [34], have found that substitution of the first phenylalanine by tyrosine in the RNP-1 motif of *B. subtilis cspB* (ie, giving the same sequence as that found in *L. casei*), results in an increase in the nucleic acid-binding capacity of CspB. This same RNP-1 motif, KGYGF, is also the amino acid sequence predominantly found in eukaryotic Y-box proteins [21]. The conservation of these two RNP motifs suggests their structural importance for the binding of nucleic acids on the surface of MCSPs. Other aromatic and basic amino acids considered to be important for nucleic acid-binding [34] are also highly conserved in all of the MCSP sequences in this study.

Discussion

Data gained in this study strongly indicate that MCSPs are ubiquitous in eubacteria, as might be deduced from the prevalence of such cold-shock-domain-containing proteins throughout nature [6,21,40,42]. Although this information should prove useful from a taxonomic standpoint, we believe that the implication of MCSP ubiquity is especially relevant to bacterial diagnostics. Conventional PCR requires the use of DNA oligomers that are specific to the target organism and, as a consequence, a sample can only be tested for the presence of that specific target. Universal PCR primers, such as those described in this study, allow a sample to be probed for the presence of any bacteria,

```
S. dysgalactiae             TGAAAAAGGTTTCGGTTTTATTTCAACTGAGAACGG......TCAAGATGTCTTCGCGCATTTTTCAGCAA
S. pyogenes                 -----------T-----------------A-----.....-----------T--A----------C-
S. aureus                   ---------A-----C-----CGA-GT---AGGA-A......AA-T--C--A----TA----------
S. intermedius              ---------A-----C-----CGA-GT---AGGA-A......AA-T--C--A----TA----------
L. cremoris                 C----------T-----C--CGA-GT---AGGA-A......AA-C--C--A----TA--C--C-------
S. epidermidis              ----------T---------CGA-GT---AGGA-A......AA-C--C--A----TA--C--C-------
P. pentosaceus              ----------T---------CGA-GT---AGGT-A......AA-C--C--A----TA--CG---------
B. subtilis (NCTC)          ---------------A--C--CGA-GTA--AGGTCA......AG-C-----A----TT-----C--T--T-
B. subtilis (EI)            A--------A-----A--C--CGA-GTA--AGGTCA......AG-C-----A----TT--C--C--T--T-
B. cereus                   A----------------C--CGA-GT---AGGC--......AG----C--A----TT-----C-----T-
E. faecalis                 A----------T--G-----C---CGC--AG----......AAGC--C--A----TA--C--C-----T-
L. grayi                    ---------A--T--A-----C---CG---AG-T--......AAGC--C--A----TT--C--C-----T-
B. megaterium               A--------A--T--A-----CGA-CGC--AGCT--......AG-C-----A----TT-----C-----T-
L.monocytogenes/innocua     A--------A--T---------CGA-CGC--A-----......-G-C-----A----TA-----CAGC--T-
L. casei                    ---T--G----A---C--C--CA-TGG---AG----......C-----C--T----TA--C--CAGC--C-
P. mondopomensis            A--G---------T------C-AA-TCAAA-C--T--TGGT...GCT----------TT--C--CCGT----
P. phosphoreum              A--G---------T---------A-TCAA--C-----CGGT...GCT--C--A----TT--C--CCGT--T-
A. hydrophila               C--G---------T--C--C--C--CC-GACCG----CAGCAAA...--C--G----TT--C--C--C--C-
A. salmonicida              C--G---------T-----C--C--CC-G-CTG----CAGCAAA...--C--G----TA--C--C--T----
E. coli                     ---C-----C-----C--C--CA-TC----CG-T--CTCTAAA...-----G----TA--C--C--T--T-
S. boydii/flexneri          ---C-----C-----C--C--CA-TC----CG-T--CTCTAAA...-----G----TA--C--C--T--T-
C. freundii/ballerapensis   ---C----------C--C--CA-TC----CG-T--CTCTAAA...--C--G----TA--C--C--T--T-
E. aerogenes                ---C-----C-----C--C--CA-TC----CG-T--CTCTAAA...--C--G----TA--C--C--T--T-
S. typhimurium/virchow      ---T-----C-----C------A-TC----TG----TTCTAAA...--C--G----TA--C--C--C--T-
P. vulgaris                 ---T---------T--------CA-TC--A-AG----CAGCAAAG-T-----A----TT--C--C--T----
Y. enterocolitica           ---T--------T-----C---A-GC---CTG----CAGCAAA...--C--G----TT--C--C--T--T-
K. pneumoniae               ATCT-----C---------C--C--TC-AA-AG-T--TAGCAAA...-----T--T-TT--T--C--T----

S. dysgalactiae             TCCAAACTAATGGTTTCAAAACATTAGAAGAAGGACAAAAAGTGGAATTTGACGTTGAAGAAGGTCAACGT
S. pyogenes                 ----------------T-----G----------------------C---------C---------------
S. aureus                   -TA-CCAAG------A----T----------G--T---GCT--T--G-----A--A-TT-----CG-C--C
S. intermedius              -TA-CCAAG------A----T-------G-----T---GCT--T--G-----A--A-TT-----CG-C--C
L. cremoris                 -TA-CCAAG-A--A-A----T-------------T---TC---T--------A--A-TT-----CG-T--C
S. epidermidis              -TA-CCAAG-A----AT---T-------------T---TC---T--------A--A-TT-----CG-C--C
P. pentosaceus              -TA-CCAAG------A----T-------------T---TC---T--------A--A-TT-----CG-T--C
B. subtilis (NCTC)          -T---GGCG-A--C--------T----------C---GCT--TTCC--C--AA-C-TT------A-C--C
B. subtilis (EI)            -T---GGCG-A--C--------T----------C---TCT--TTCT--C--AA-C-TT-----AA-C--C
B. cereus                   -----GGCG-A--C--------C----------T---G----TACT--C--A--A--GC-----A-C---
E. faecalis                 -----GG-G------------T----------T---GC----ACT-----T--------TTCAG-C---
L. grayi                    -----GG-G-C--A-------------------T---GC----ACT--C------------TCAG-C--C
B. megaterium               -T---GGAG-A--A-----GT-T----------CG--TCT--TTC------T------C----C------
L.monocytogenes/innocua     -----GGCG-C--A------T-T-----C-----T---GC---AACT--C--------------C-----C
L. casei                    --A-CGG-G-A----A---G-GTC-T--C-----T--GGCT--TTCC-AC--T--A---C--TC-G-T---
P. mondopomensis            --GCTT-AG-G-----T-----C---ACT-----T--G--G--CTC------T-----GC----C---AAA
P. phosphoreum              --GCTT--G-A----------TC-T-CT-----C---------TCT-----A--A--GC----C---AAA
A. hydrophila               ----G--CCCG--C-----G--CC-G--C-----T--GCGC--T--G--CAC-A-C---C----C--GAAA
A. salmonicida              ----GT-C-CCA-C---------TC-G--C-----T--GCGC--T--G--CAC-A-C--GC----C--GAAA
E. coli                     ----G-ACG------A----T-TC-G--C-----T--G------TCC--CAC-A-C---AGC--CGCTAAA
S. boydii/flexneri          ----G-ACG------A----T-TC-G--C-----T--G------TCC--CAC-A-C---AGC--CGCTAAA
C. freundii/ballerapensis   ----G-ACG----C-A----T-TC-G--C-----T--G-----TTCC--CAC-A-C---AGC--CGCTAAA
E. aerogenes                ----G-ACG----C-A----T-CC-G--C-----T--G-----TTCC--CAC-A-C---AGC--CGCTAAA
S. typhimurium/virchow      -T--G-ACG------A----T-TC-G--C-----T--G-----TTCC--CAC-A-C---AGC--CGCTAAG
P. vulgaris                 ----G-GCG--A-C------T--C-GA-------T---G----TTCA---TCTA-C---A-T--CGCTAAA
Y. enterocolitica           ----G-GC----A---------CC-G--T-----T--------T--G--CTCTA-C--GA-T---GCTAAA
K. pneumoniae               ----G-GCG--A-C--------C---A-T-----T---G----TAGC--CACTA-C---A-C---GC-AAA
```

Figure 2 Alignment of partial MCSP gene sequences amplified from the bacteria shown in Table 1 using the universal PCR primers CSPU3 and CSPU5 (see text for PCR conditions). Sequences were aligned using the Genetics Computer Group program PileUp (version 8, Wisconsin, USA; accessed via the BBSRC facility at Daresbury, UK). This program creates a multiple sequence alignment from a group of related sequences using progressive pairwise alignments. Nucleotides identical to those aligned in the top sequence (*S. dysgalactiae*) are shown as dashes (–), with dots (.) indicating spaces within the alignments.

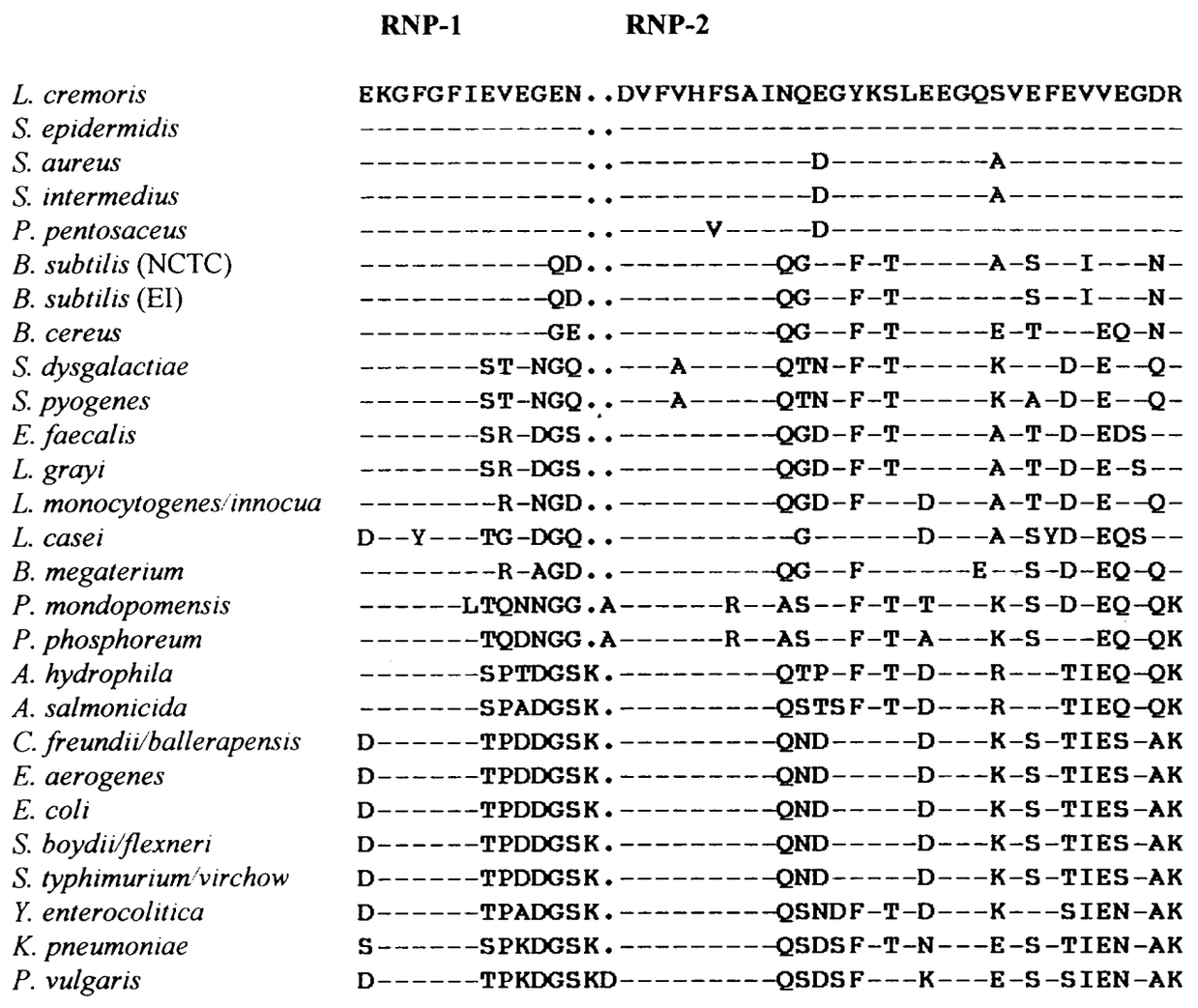

Figure 3 Alignment of partial MCSP amino acid sequences as deduced from their respective amplified gene sequences (Figure 2). Sequence alignments were achieved as for Figure 2. Amino acids identical to those aligned in the top sequence (*L. cremoris*) are shown as dashes (–), with dots (.) indicating spaces within the alignments. Positions of the nucleic acid-binding motifs RNP-1 and RNP-2 are also shown.

followed by identification based on nucleotide variations between strains. Although the partial MCSP gene sequences are highly homologous (Figure 2), in the majority of cases there are sufficient nucleotide variations to allow the unique allocation of each amplified sequence to its parental bacterium (Figure 4).

The above findings reveal the diagnostic potential of a MCSP PCR-based assay for the detection/discrimination of bacteria. Cloning and sequencing PCR products generated using this procedure, however, is not feasible as a method of bacterial discrimination due to time and expense. Hence, in order for this assay to be of benefit to industry it would be necessary to incorporate alternative methodologies that are capable of discriminating between two nucleotide sequences on the basis of a single base mismatch. Techniques such as single strand conformation polymorphism analysis, fluorescence-assisted mismatch analysis and fluorogenic 5′ nuclease PCR assays, are all capable of detecting a single base mismatch and have been successfully used for allelic discrimination of human diseases [14,24,38]. Furthermore, each of the latter techniques works optimally with DNA fragments of 200 bp or less, which is optimal for the analysis of PCR fragment gained during this study (larger gene sequences would be unable to give additional information without including extra primer oligomers). Fluorogenic 5′ nuclease PCR assays, which have also been used for the specific detection of pathogenic bacteria in food [41] and for the detection/discrimination of psychrotolerant and mesophilic *B. cereus* strains gained from milk (our unpublished data), can be used to monitor the amplification of a specific target sequence during (or immediately after) PCR, adding little or no extra time. Moreover, since this assay is specific for the detection of a particular amplified target sequence, the amplification of non-specific DNA sequences (eg, from contaminating *E. coli* DNA which may be present in certain PCR reagents) does not cause interference.

Although the above assay could equally be developed for 16S rDNA, we believe that targeting MCSP nucleic acid sequences has more potential for future development,

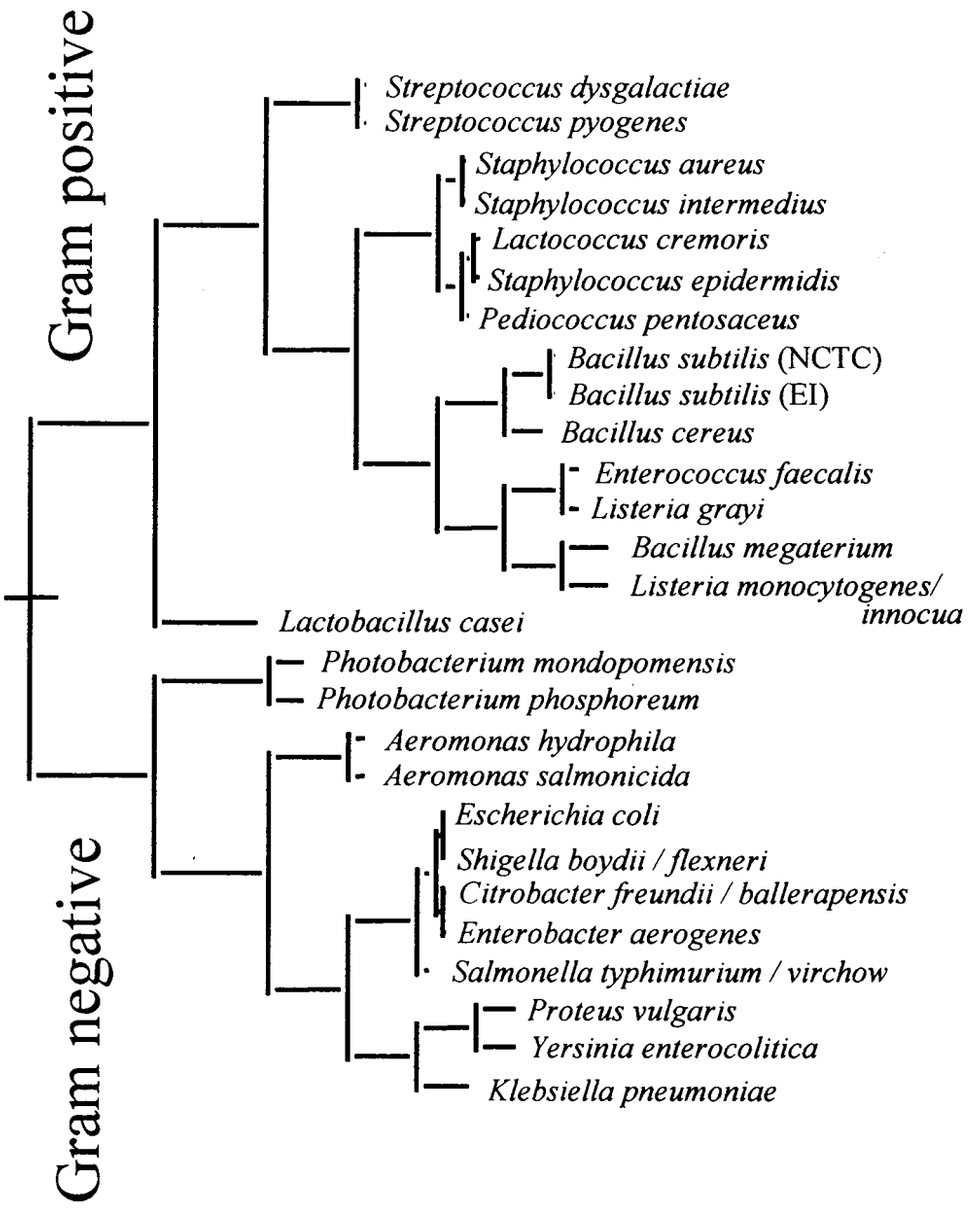

Figure 4 Dendrographic representation of PileUp (Genetics Computer Group programs, version 8, Wisconsin, USA; accessed via the BBSRC facility at Daresbury, UK) derived from the alignment of partial MCSP gene sequences shown in Figure 2. The dendrogram was created using the Genetics Computer Group programs Distances and Growtree (version 8, Wisconsin, USA; accessed via the BBSRC facility at Daresbury, UK). Growtree creates a phylogenetic tree from a distance matrix created by Distances which used an unweighted pair group method based on arithmetic averages. Vertical line distances indicate closeness of sequence homologies.

especially with regard to the discrimination of live bacteria from those that are non-viable or dead. Recently, Goldenberg *et al* [12] have shown that *E. coli cspA* mRNA is extremely unstable at 37°C with a half-life of approximately 10 s, but that this same transcript is highly stable at 15°C. We have unpublished data to indicate that similar cold-inducible MCSP mRNA stability is found in a number of other bacteria, both Gram-positive and Gram-negative. Furthermore, the fact that cold-inducible MCSP mRNA is induced both rapidly and substantially post-cold shock [12,39, unpublished data], heightens the advantage of using this particular template for RT-PCR in preference to other universal mRNA templates or 16s rRNA.

Acknowledgements

We thank W Fielder, J Throup and J Keyte at the University of Nottingham for their help in preparing bacterial cultures and with the computer analysis of sequence data. We also thank C Rees, F Barnard and J Dickinson at the University of Nottingham, S Picton from PE Applied Biosystems, UK, and B Mayr, S Lechner and S Scherer from the Technsche

Universitat Munchen, Germany for their helpful comments during the course of this work. KPF was funded via a BBSRC RoPA award.

References

1 Altschull SF, W Gish, W Miller, EW Myers and DJ Lipman. 1990. Basic local alignment search tool. J Mol Biol 215: 403–410. Program utilized via GenBank at National Center for Biotechnology Information, Bethesda, MD, USA.
2 Ausubel FM, R Brent, RE Kingston, DD Moore, JG Seidman, JA Smith and K Struhl. 1995. Current Protocols in Molecular Biology. Section 1: 1.8.4. Wiley, USA.
3 Av-Gay Y, Y Aharonowitz and G Cohen. 1992. *Streptomyces* contain a 7.0 kDa cold shock like protein. Nucl Acids Res 20: 5478.
4 Berger F, N Morellet, F Menu and P Potier. 1996. Cold shock and cold acclimation proteins in the psychrotrophic bacterium *Arthrobacter globiformis* S155. J Bacteriol 178: 2999–3007.
5 Blattner FR, G Plunkett, GF Meyhew, NT Perna and FD Glasner. 1997. *E. coli* genome project. GenBank/EMBL/DDBJ accession No. AE000252(g1787834).
6 Burd CG and G Dreyfuss 1994 Conserved structures and diversity of functions of RNA-binding proteins. Science 265: 615–621.
7 Devlin K and CM Churcher. 1997. *Mycobacterium tuberculosis* sequencing project. Gen/Bank/EMBL/DDBJ accession no. Z95436.
8 Didier DK, J Schiffenbauer, SL Woulfe, M Zachris and BD Schwartz. 1988. Characterization of the cDNA encoding a protein binding to the major histocompatibility complex class II Y box. Proc Natl Acad Sci USA 85: 7322–7326.
9 Fleischmann RD, MD Adams, O White, RA Clayton, EF Kirkness, AR Kerlavage, CJ Bult, J-F Tomb, BA Dougherty, JM Merrick, K McKenney, G Sutton, W FitzHugh, C Fields, JD Gocayne, J Scott, R Shirley, L Liu, A Glodek, JM Kelley, JF Weidman, CA Phillips, T Spriggs, E Hedblom, MD Cotton, TR Utterback, MC Hanna, DT Nguyen, DM Saudek, RC Brandon, LD Fine, JL Fritchman, JL Fuhrmann, NSM Geoghagen, CL Gnehm, LA McDonald, KV Small, CM Fraser, HO Smith and JC Venter. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269: 496–512.
10 Francis KP, CED Rees and GSAB Stewart. 1995. Identification of a major cold-shock protein homologue in *Listeria monocytogenes*. GenBank/EMBL/DDBJ accession No. X91789.
11 Francis KP and GSAB Stewart. 1997. Gene duplication: a mechanism for the evolution of major cold-shock protein families? GenBank/EMBL/DDBJ accession No. AF003591.
12 Goldenberg D, I Azar and AB Oppenheim. 1996. Different mRNA stability of the *cspA* gene in the cold-shock response of *Escherichia coli*. Mol Microbiol 19: 241–248.
13 Goldstein J, NS Pollitt and M Inouye. 1990. Major cold shock protein of *Escherichia coli*. Proc Natl Acad Sci USA 87: 283–287.
14 Glavac D and M Dean. 1993. Optimization of the single strand-conformational polymorphism (SSCP) technique for detection of point mutations. Hum Mutat 2: 404–414.
15 Graumann P and MA Marahiel. 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.
16 Graumann P, K Schroder, R Schmid and MA Marahiel. 1996. Cold shock stress-induced proteins in *Bacillus subtilis*. J Bacteriol 178: 4611–4619.
17 Hebraud M, E Dubois, P Potier and J Labadie. 1994. Effect of growth temperature on the protein levels in the pychrotrophic bacterium, *Pseudomonas fragi*. J Bacteriol 176: 4017–4024.
18 Holt JG, NR Krieg, PHA Sneath, JT Staley and ST Williams (eds). 1994. Bergey's Manual of Determinative Bacteriology (9th edn). Williams and Wilkins, Baltimore.
19 Jones PG, RA VanBogelen and FC Neidhardt. 1987. Induction of proteins in response to low temperature in *Escherichia coli*. J Bacteriol 169: 2092–2095.
20 Jones PG and M Inouye. 1994. The cold-shock response—a hot topic. Mol Microbiol 11: 811–818.
21 Landsman D. 1992. RNP-1, an RNA-binding motif is conserved in the DNA-binding cold shock domain. Nucl Acids Res 20: 2861–2864.
22 Law RM and AK Bej. 1993. Identification and characterization of *cspS*, a cold shock related gene from *Salmonella typhimurium*. GenBank/EMBL/DDBJ accession No. L23115.
23 Lee SJ, A Xie, W Jiang, J-P Etchegaray, PG Jones and M Inouye. 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.
24 Livak KJ, J Marmaro and JA Todd. 1995. Towards fully automated genome-wide polymorphism screening. Nature Genet 9: 341–342.
25 Mayo B, S Derzell, M Fernandez, C Leonard, T Ferain, P Hols, J Suarez and J Delcour. 1996. Cloning and characterization of *cspL* and *cspP*, two cold-inducible genes from *Lactobacillus plantarum*. J Bacteriol 179: 3039–3042.
26 Mayr B, T Kaplan, S Lechner and S Scherer. 1996. Identification and purification of a family of dimeric major cold shock protein homologs from the psychrotrophic *Bacillus cereus* WSBC 10201. J Bacteriol 178: 2916–2925.
27 Michel V, J Labadie and M Hebraud. 1996. Effect of different temperature upshifts on protein synthesis by the psychrotrophic bacterium *Pseudomonas fragi*. Curr Microbiol 33: 16–25.
28 Nakashima K, K Kanamaru, T Mizuno and K Horikoshi. 1996. A novel member of the *cspA* family of genes that is induced by cold shock in *Escherichia coli*. J Bacteriol 178: 2994–2997.
29 Qian L and M Wilkinson. 1991. DNA fragment purification—removal of agarose 10 minutes after electrophoresis. Biotechniques 10: 736.
30 Ray MK, T Sitaramamma, S Ghandhi and S Shivaji. 1994. Occurrence and expression of *cspA*, a cold shock gene, in Antarctic psychrotrophic bacteria. FEMS Microbiol Lett 116: 55–60.
31 Schindelin H, MA Marahiel and U Heinemann. 1993. Universal nucleic acid-binding domain revealed by crystal structure of the *B. subtilis* major cold-shock protein. Nature 364: 164–168.
32 Schnuchel A, R Wiltscheck, M Czisch, M Herrier, G Willimsky, P Graumann, MA Marahiel and TA Holak. 1993. Structure in solution of the major cold-shock protein from *Bacillus subtilis*. Nature 364: 169–171.
33 Schroder K, P Zuber, G Willimsky, B Wagner and M Marahiel. 1993. Mapping of the *Bacillus subtilis cspB* gene and cloning of its homologs in thermophilic, mesophilic and psychrotrophic bacilli. Gene 136: 277–280.
34 Schroder K, P Graumann, A Schnuchel, TA Holak and MA Marahiel. 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.
35 Smith CM, WH Koch, SB Franklin, PL Foster, TA Cebula and E Eisenstadt. 1990. Sequence analysis and mapping of the *Salmonella typhimurium* LT2 *umuDC* operon. J Bacteriol 172: 4964–4978.
36 Stamm I, A Leclerque and W Plaga. 1996. A gene encoding a cold shock-like protein (Csp) from *Stigmatella aurantiaca*. GenBank/EMBL/DDBJ accession No. U70990.
37 Tafuri SR and AP Wolffe. 1990. *Xenopus* Y-box transcription factors molecular cloning, functional analysis, and developmental regulation. Proc Natl Acad Sci USA 87: 9028–9032.
38 Verpy E, M Biasotto, T Meo and M Tosi. 1994. Efficient detection of point mutations on color-coded strands of target DNA. Proc Natl Acad Sci USA 91: 1873–1877.
39 Willimsky G, H Bang, G Fischer and MA Marahiel. 1992. Characterization of *cspB*, a *Bacillus subtilis* inducible cold shock gene affecting cell viability at low temperatures. J Bacteriol 174: 6326–6335.
40 Wistow G. 1990. Cold shock and DNA binding. Nature 344: 823–824.
41 Witham PK, CT Yamashiro, KJ Livak and CA Batt. 1996. A PCR-based assay for the detection of *Escherichia coli* shiga-like toxin genes in ground beef. Appl Environ Microbiol 62: 1347–1353.
42 Wolffe AP. 1994. Structural and functional properties of the evolutionarily ancient Y-box family of nucleic acid binding proteins. Bioessays 16: 245–251.