Implication of CcpN in the regulation of a novel untranslated RNA (SR1) in *Bacillus subtilis*

Andreas Licht, Sven Preis† and Sabine Brantl*
Institut für Molekularbiologie, Friedrich-Schiller-Universität Jena, Hans-Knöll-Str. 2, Jena D-07745, Germany.

Summary

Antisense-RNAs have been investigated in detail over the past 20 years as the principal regulators in accessory DNA elements such as plasmids, phages and transposons. However, only a few examples of chromosomally encoded bacterial antisense RNAs were known. Meanwhile, >70 small non-coding RNAs from the *Escherichia coli* genome have been found, the functions of the majority of which remain to be elucidated. Only one systematic search has been performed for Gram-positive bacteria, so far. Here, we report the identification of a novel small (205 nt) non-translated RNA – SR1 – encoded in the *Bacillus subtilis* genome. SR1 was predicted by a computational approach and verified by Northern blotting. Knockout or overexpression of SR1 did not affect growth. SR1 was derepressed under conditions of gluconeogenesis, but repressed under glycolytic conditions. Two regulatory levels could be identified, one involving CcpA, the second, more important, involving the recently identified regulator CcpN.

Introduction

Over the past three years, novel small non-coding RNAs, found in both prokaryotes and eukaryotes, have been the focus of many articles. In prokaryotes, antisense RNAs have been known and studied for over 20 years as principal regulators in prokaryotic accessory DNA elements like plasmids, phages and transposons where they inhibit gene expression by a variety of mechanisms (for reviews see Brantl, 2002; Wagner et al., 2002). It became clear that non-coding RNAs are more abundant and important than initially imagined. Alone in 2001, three groups have translated RNA – SR1 – encoded in the *Escherichia coli* genome (Argaman et al., 2001; Rivas et al., 2001; Wassarman et al., 2001) and it was suggested that >100 of such small chromosomally encoded RNAs might exist in this species. However, the function of only a few chromosomally encoded RNAs is known so far such as MicF (for a review see Delihas and Forst, 2001), OxyS RNA (Altuvia et al., 1997; Altuvia et al., 1998), DsrA (Majdalani et al., 1998; Lease and Belfort, 2000), as well as the recently clarified Spot42 RNA (Møller et al., 2002) and RyhB-RNA (Massé and Gottesman, 2002). In 2004, the function of four novel, recently identified non-coding RNAs of the *E. coli* chromosome were elucidated, namely GadY (Opdyke et al., 2004), MicC (Chen et al., 2004), SgrS (Vanderpool and Gottesman, 2004) and IstR1 (Vogel et al., 2004). In contrast to plasmid- or phage-encoded regulatory RNAs, the majority of chromosomally encoded RNAs known so far are expressed in *trans* which correlates with being only partially complementary to their targets. Not all of these newly discovered RNAs act as *bona fide* antisense regulators by complementary base pairing with their target RNAs. For example, 6S RNA has been shown to mediate growth phase specific changes in RNA polymerase by directly binding to *E. coli* RNA polymerase σ70 and β/β’ subunits (Wassarman and Storz, 2000).

Whereas a variety of small regulatory RNAs has been already identified and studied in *E. coli*, relatively little is known about chromosomally encoded regulatory RNAs from Gram-positive bacteria. One of the few examples is a cis-encoded antisense RNA of *Clostridium acetobutylicum*, whose target is the glutamine synthetase gene *glnA* RBS (Fierro-Monti et al., 1992). Another one is the multifunctional 514 nt RNAIII encoded in the *agr* locus of *Staphylococcus aureus* (Morfeldt et al., 1995; Huntzinger et al., 2005). Recently, two RNAs encoded in the *yocJ* intergenic region (BS203, Ando et al., 2002) and in the *asp-yrvM* intergenic region (BS190, Suzuma et al., 2002) of the *Bacillus subtilis* genome were found. Whereas BS203 seems to be unique for *B. subtilis*, BS190 is suggested to be also encoded in the genomes of *Bacillus halodurans* and *Listeria monocytogenes*. In both cases, these RNAs seem to be abundant, because they were discovered in ethidium bromide-stained polyacrylamide gels as significant bands. Neither of these RNAs is essential, because knockout strains could be constructed and failed to show severe phenotypes except for a reduced growth rate in the case of BS190. So far,
the functions of these RNAs remain elusive. A recent RNA expression analysis using an antisense B. subtilis genome array revealed a few non-coding transcripts in intergenic regions, the majority of which are probably representing processing or read-through products from upstream or downstream transcripts (Lee et al., 2001). No detailed investigation of these transcripts was reported.

The majority of chromosomally encoded ncRNAs are expressed only under certain environmental conditions, e.g. solely in stationary phase or depending on the composition of the media. One of the best studied examples is MicF that is only expressed at high temperatures and high salt (for a review on the global regulatory network involved in MicF regulation see Delilhas and Forst, 2001). Other examples include OxyS, which is induced by oxidative stress (Altuvia et al., 1997), RyhB responding to the iron concentration (Massé and Gottesman, 2002), the recently identified sugar-regulated SgrS (Vanderpool and Gottesman, 2004) or IstR, which is SOS-induced (Vogel et al., 2004). In several cases, regulatory DNA regions have been identified upstream of the ncRNA promoters, like the FurB box upstream of pRnfB or the SgrR box upstream of pSgrR.

Recently, a novel regulator for CcpA-independent catabolite repression in B. subtilis was identified and designated CcpN (Servant et al., 2005). The authors showed that two genes, gapB and pckA, necessary for efficient gluconeogenesis from Krebs cycle intermediates, were repressed in the presence of glucose but not via CcpA, the major transcriptional regulator for catabolite repression in B. subtilis. They found that CcpN was responsible for this glucose repression.

Here, we describe the identification of a small RNA termed SR1 encoded in the B. subtilis genome and the implication of CcpN in the regulation of SR1 expression. The 205 nt long RNA SR1 was found using computer predictions for potential promoters and terminators in intergenic regions of the B. subtilis genome and verified by Northern blotting. SR1 expression was maximal under gluconeogenic conditions. A sr1 knockout strain grew equally well as the wild-type strain indicating that SR1 is not essential for B. subtilis. Overexpression of sr1 either from its own or from the heterologous tet promoter was not detrimental for growth. SR1 synthesis is repressed under glycolytic conditions. We identified two DNA regions upstream of the sr1 promoter responsible for sugar-mediated repression of SR1 transcription, namely a cre site located 275 bp upstream and the 17 bp sugar box – binding site for CcpN – located 58 bp upstream of the SR1 transcriptional start site. The major part of sugar-mediated repression is accomplished by CcpN, whereas the catabolite control system involving CcpA and the cre box play only a minor role.

**Results**

**Computer prediction of small non-coding RNAs in the intergenic regions of the B. subtilis genome**

A computational approach was devised for predicting small non-coding RNA genes in the intergenic regions of the B. subtilis genome. First, the B. subtilis genome was searched for ≥5 adjacent T residues, an important property of rho-independent transcriptional terminators. The resulting polyT-database, containing 15232 and 15003 stretches of polyT on the + and − strands, respectively, was used to predict rho-independent transcriptional terminators by *in silico* folding of stretches of 50 bases upstream of these 5 T residues. The predicted secondary structures were analysed for features characteristic of rho-independent terminators (folding into a stem-loop or hairpin with a 3–8 nt loop and a GC rich 4–16 bp stem). Putative terminators with a predicted free energy of folding of more than −10 kcal mol⁻¹ were discarded, reducing the number of terminator sequences in the database to 320 on the + strand and 551 on the − strand. The average folding energy for these terminators was ≈−15 kcal mol⁻¹. This terminator database was used to predict potential promoters 50–450 bp upstream of the terminator sequences. The criteria for a promoter prediction programme were generated using the DDBTB database of B. subtilis promoters and transcription factors (Ishii et al., 2001). From this database, sequences of the −35 region, of the −10 region and the length of the spacer region were deduced as search criteria.

First, we chose to search for sequences similar to those of σ⁺ dependent promoters by scoring similarity to known −35 boxes and −10 boxes. This gave 113 putative promoters on the + strand and 74 on the − strand that showed a high score in both −35 and −10 boxes and had 16–18 nt spacers. Additionally, the terminator database was searched for promoters whose −35 and −10 box sequences were identical to promoters from the DDBTB database, and for promoters with all possible combinations of −35 and −10 box DDBTB-based sequences, resulting in 22 and 16 or 112 and 74 promoter sequences in both instances respectively. Furthermore, we searched for SigB, SigE and SigK promoters, however, only two putative SigB dependent promoters were found.

In the last step, the predicted new genes for small RNAs were subjected to an analysis for the following criteria: (i) length (putative RNA should be between 80 and 350 nt long); (ii) overlaps with other genes; (iii) overlaps with genes on the complementary strand (i.e. not being completely located in intergenic regions); (iv) GC content; and (v) homology to sequences in related organisms like B. halodurans. Both of the latter criteria were the least useful, because (i) the comparison of the GC content did not
result in altered predictions and (ii) \(B.\ subtilis\) and \(B.\ halodurans\) are only 43% identical.

We found in total 25 putative genes, 16 on the + strand and nine putative genes on the – strand of the \(B.\ subtilis\) genome, 19 of which had no open reading frames (ORFs) on the complementary strand.

**Northern blotting identified a ~200 nt long RNA – SR1 – encoded in the pdhD-yktA intergenic region of the \(B.\ subtilis\) genome**

To verify the computer predicted small RNAs, \(B.\ subtilis\) strain DB104 was grown at three different temperatures (30\(^\circ\)C, 37\(^\circ\)C and 48\(^\circ\)C) in complex and minimal medium, 0.5 ml aliquots were taken at different growth phases, frozen in liquid nitrogen and, subsequently, used for the isolation of total RNA as described (Experimental procedures). Five to 10 \(\mu\)g of each RNA sample were separated on 6% polyacrylamide gels along with the \([\gamma^{32P}]\)-ATP-labelled pBR322 MspI marker. Northern blotting was performed as described using single-stranded oligonucleotides complementary to the putative small RNAs. With this approach, none of the predicted RNAs could be found. As controls, oligonucleotides complementary to the previously described small RNAs BS203 and BS190 (Ando et al., 2002; Suzuma et al., 2002) were used successfully. To increase the sensitivity, we decided to use intrinsically labelled double-stranded DNA obtained by a polymerase chain reaction (PCR) upon inclusion of \([\alpha^{32P}]\)-dATP. With these probes we were able to detect one small RNA designated SR1 (small RNA 1) with a size of about 200 nt. This result was verified with two \([\alpha^{32P}]\)-UTP-labelled riboprobes, one for the small RNA, the other for a hypothetical RNA encoded by the complementary strand. As shown in Fig. 1A, SR1 is encoded in the intergenic region between the pyruvate dehydrogenase D (\(pdhD\)) gene and an unidentified gene \(yktA\). In initial experiments, SR1 was found only in TY medium at 30\(^\circ\)C, 37\(^\circ\)C and 48\(^\circ\)C, and exclusively in the stationary phase.

**Mapping of the 5’ and 3’ ends of SR1**

To map the 5’ and the 3’ end of SR1, primer extension and 3’ RACE were used. As shown in Fig. 2, the 5’ end of SR1 is located at position 1036 (see Fig. 1B), exactly 10 nt downstream from the centre of the putative –10 box of the \(sr1\) promoter. This position coincided well with the result of an in vitro transcription experiment with \(B.\ subtilis\) \(\sigma^A\) RNA polymerase using a DNA fragment spanning nt 901–1247 (Fig. 1B) as template (not shown). Thus, the predicted \(\sigma^A\) promoter of the \(sr1\) gene with TTGACA as –35 box and TAATAT as –10 box could be verified both in vitro and in vivo. The 3’ end of SR1 was mapped with 3’ RACE at position 1241 within the predicted bidirectional terminator (gel not shown). From these data we can conclude that SR1 is 205 nt long.

**Expression of SR1 in \(B.\ subtilis\) DB104 is different in complex and minimal medium and at different temperatures**

To investigate the expression pattern of SR1 in \(B.\ subtilis\), DB104 was grown in TY medium at 30\(^\circ\)C, 37\(^\circ\)C and 48\(^\circ\)C, time samples were taken at different optical densities in

---

**Fig. 1. Location of the \(sr1\) gene.**

A. Schematic representation of the location of the \(sr1\) gene on the \(B.\ subtilis\) chromosome. The direction of transcription is indicated by arrows. Above the line, genes transcribed from the + strand are shown.

B. Sequence of the \(sr1\) gene. –35 and –10 boxes of the \(sr1\) promoter are indicated. Nucleotide numbers have been altered in comparison with the chromosome to facilitate the location of sequences analysed in this report as follows: nt 1533439 on the \(B.\ subtilis\) chromosome (+ strand) has been altered into nt 1000, and nt 1534439 on the \(B.\ subtilis\) chromosome has been altered into nt 2000. Arrows indicate the bidirectional transcription terminator shared with the \(speA\) gene on the – strand. Start and direction of transcription are indicated by an arrow, and transcription termination is indicated by a black dot. The ATG start codon and the TAA stop codon of the open reading frame are shown by grey boxes.
early and late log phase and in early and late stationary phase. Total RNA was prepared and analysed by Northern blotting as described above. Interestingly, expression of \textit{sr1} was detectable at all temperatures. In all cases, the highest expression was observed upon entry of the cells into stationary phase. Afterwards, it ceased, but remained constant till late stationary phase (Fig. 3A). In contrast, upon cultivation in glucose minimal medium (Spizizen medium), \textit{SR1} was weakly detectable only at 48°C and only in late stationary phase, but neither at 37°C or below nor upon entry into stationary phase (Fig. 3B).

\textit{SR1} expression can be repressed by various sugars

To find out, whether the exhaustion of certain nutrients like carbohydrates or phosphates triggers the increase of \textit{sr1} expression upon onset of stationary phase, we added either glucose or arabinose, sucrose, fructose or glycerol to a final concentration of 2% to growing cultures shortly before onset of stationary phase. Alternatively, phosphates were added to a final concentration of 100 mM.

No alteration of \textit{SR1} expression was observed with phosphates, however, all different sugars were able to repress \textit{SR1} expression in less than 30 min (in the case of glycerol 60 min) after addition (Fig. 3C). Furthermore, a comparison of \textit{SR1} expression in minimal medium in the absence and presence of 0.5% glucose revealed that the failure to detect \textit{SR1} in this medium was solely due to the presence of glucose (Fig. 3B).

The ORF of \textit{SR1} is not expressed in \textit{B. subtilis}

Inspection of the \textit{SR1} sequence (Fig. 1B) revealed the presence of an ORF starting with an ATG codon that is preceded by a very weak putative SD sequence. This ORF encodes a 39 aa peptide. To find out, whether or not this ORF is translated in \textit{B. subtilis}, one transcriptional and three translational fusions of the \textit{sr1} gene pACT87 and pACP87/pACP422/pACP421, respectively, were constructed as described in \textit{Experimental procedures} and integrated into the \textit{amyE} locus of \textit{B. subtilis} strain DB104. As a control for the translational fusions, the empty vector pAC5 was integrated into the \textit{amyE} locus.

The measurements of \(\beta\)-galactosidase activities (Table 1) from the corresponding \textit{B. subtilis} strains grown in TY or in Spizizen minimal medium with or without glucose (derepression of \textit{SR1} expression) revealed that the

<table>
<thead>
<tr>
<th></th>
<th>(\beta)-Galactosidase activities (Miller units)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Transcriptional fusion</td>
</tr>
<tr>
<td></td>
<td>(pACT87)</td>
</tr>
<tr>
<td>TY – glc</td>
<td>580.0</td>
</tr>
<tr>
<td>TY + glc</td>
<td>49.9</td>
</tr>
<tr>
<td>Spiz – glc</td>
<td>1130.0</td>
</tr>
<tr>
<td>Spiz + glc</td>
<td>16.9</td>
</tr>
</tbody>
</table>

Cells were grown in TY or Spizizen medium with (+ glc) or without (– glc) 2% or 0.5% glucose, respectively, till entry into stationary phase, samples were taken and \(\beta\)-galactosidase activities were determined as described previously (Brantl, 1994). Average activities obtained from three independent measurements are shown.
SR1 promoter was with approximately 580 and 1130 Miller units in TY and minimal medium without glucose, respectively, very strong, when 87 bp upstream sequence were present (pACT87). However, no peptide was translated from a single copy chromosomal sr1 gene, neither in pACP87, carrying 87 bp upstream of pSR1 (Table 1), nor in pACP422 or pACP421, carrying 150 or 250 bp upstream of pSR1 (latter data not shown). The difference in the β-galactosidase activities between transcriptional and translational fusions was about 500- to 1130-fold. In fact, the activities measured from the translational lacZ fusions were in the same range as those obtained with the empty integrated pAC5 vector sequence (Table 1). As in cases where translation of an ORF is known, the difference in β-galactosidase activities between transcriptional and translational fusions were only two- to threefold (e.g. Stülke et al., 1997), we can safely assume that SR1 is a non-translated RNA.

Neither SR1 knockout nor overexpression is detrimental for growth

To find out whether SR1 is essential in B. subtilis, a sr1 knockout strain was constructed. For this purpose, vector pINT1 carrying the CAT gene flanked by the upstream and downstream regions of the sr1 wild-type gene, was constructed as described in Experimental procedures, and integrated into the chromosome of B. subtilis DB104 or B. subtilis 168. Successful integration by double cross over was confirmed by PCR and by Northern blotting. Growth curves of the Δsr1 strain grown in TY medium and in minimal medium with or without glucose at 37°C or 48°C proved to be identical to that of the wild-type strain.

To analyse, whether overexpression of sr1 causes significant physiological effects, two overexpression strains were constructed, one for inducible and the other for constitutive expression of the sr1 gene under control of its own promoter in B. subtilis. For inducible overexpression, the promoterless sr1 gene was inserted into the expression vector pWH353 (Geissendörfer and Hillen, 1989) resulting in plasmid pWSR1. For constitutive overexpression, the sr1 gene including its own promoter was inserted into the broad host range lactococcal vector pGK14 yielding plasmid pGKSR1. B. subtilis Δsr1 strains were transformed with either pWSR1 or pGKSR1. A significant expression of sr1 from pGKSR1 in TY medium and in minimal medium without glucose was visible in logarithmic growth phase and increased upon entry into stationary
phase about ninefold (not shown). Similarly, a ~10-fold increase in SR1 expression was observed upon induction of pWSR1 with sublethal tetracycline concentrations. Growth curves of *B. subtilis* strains containing the overexpression plasmids pGKSR1 or pWSR1 in TY or in minimal medium with or without glucose were identical to those obtained with *B. subtilis* containing the empty vectors.

In summary, these data indicate that neither knockout nor overexpression of SR1 is detrimental to growth of *B. subtilis*.

The cis-acting target of sugar repression can be narrowed down to 17 bp located upstream of the SR1 promoter

The fact that sugar-mediated repression was also observed when plasmid pGKSR1 carrying the sr1 gene with 100 bp upstream of the sr1 promoter was analysed in a chromosomal sr1 knockout strain indicated that the DNA region responsible for the repression effect resided within these 100 bp. This cis-acting target of glucose repression was designated ‘sugar box’.

To narrow down the sequence responsible for sugar-mediated repression, first, the region upstream of the sr1 promoter pSR1 was shortened progressively. For this purpose, several pGK14 derivatives were constructed as described in *Experimental procedures* for pGKSR50 using the primers shown in Table S1 (*Supplementary material*) and analysed for sugar-mediated SR1 repression in Northern blots in the presence and absence of 2% glucose. As shown in Fig. 4A, pGKSR30 carrying 30 bp upstream of the −35 box of pSR1 still showed a 20-fold repression, whereas pGKSR20 only showed a 2.5-fold repression. Plasmids pGKSR23 and pGKSR22 still revealed a 18- to 19-fold repression (not shown), suggesting that the upstream border of the sugar box is located about 22 nt upstream of the −35 box of pSR1.

To narrow down the downstream border of the sugar box, pGKSR1 derivatives carrying 30 bp upstream of pSR1 were constructed containing exchanges of 4, 5, 6 or 8 bp, respectively, upstream of the −35 box of pSR1. The corresponding plasmids were designated pGKSR30/4, pGKSR30/5, pGKSR30/6, etc. Northern analyses revealed that 5 bp could be altered without impairing the sugar repression, whereas substitution of 6 bp (pGKSR30/6) led to a decrease of the repression effect from 14-fold (pGKSR30/5) to fivefold (Fig. 4B). Figure 4C summarizes the upstream and downstream borders of the sugar box. The results summarized here indicate that the sugar box is an element necessary for glucose mediated repression, but do not prove that this element is sufficient, i.e. that no other regulatory elements are required.

A provisional consensus sequence for the sugar box could be derived by gel mobility shift assays

To find out whether a protein binding to the sugar box could be identified by EMSA (electrophoretic mobility shift assay), a radioactively labelled PCR fragment containing pSR1 with the sugar box was incubated with crude extracts of *B. subtilis* strain DB104 prepared after growth in the

![Fig. 4. Localization of the upper and lower border of the sugar box. Autoradiographs of the Northern blots are shown.](image-url)
presence or absence of glucose in either logarithmic or stationary phase and analysed in a native PAA gel. After overnight exposure of the dried gel in the PhosphorImager, one retarded band could be observed in all cases, i.e. with extracts from all growth phases and independent of the presence or absence of glucose in the medium (data not shown). This was unexpected, because we suspected a dependence of protein expression on the presence of glucose and/or on the growth phase.

To derive a provisional consensus sequence for the sugar box, 23 bp long radioactively labelled oligodeoxynucleotides containing the sugar box flanked by 2 GC bp on either side were used to obtain double-stranded DNA fragments for EMSAs as described in Experimental procedures. Thereby, each single nucleotide was replaced by one of the three others. Table S2 in Supplementary material summarizes the oligonucleotides used for these EMSA studies. Figure 5A shows an example for such an EMSA. From all these EMSAs, the following provisional consensus sequence was derived:

\[
\begin{align*}
5' & \quad \text{A} \quad \text{A} \quad \text{A} \quad \text{T} \quad \text{T} \\
& \quad \text{G} \quad \text{G} \quad \text{G} \quad \text{Y} \quad \text{Y} \\
& \quad \text{T} \quad \text{T} \quad \text{T} \quad \text{Y} \quad \text{A} \quad \text{T} \quad \text{A} \quad \text{T} \quad \text{A} \\
& \quad \text{Y} \quad \text{A} \quad \text{G} \quad \text{G} \quad \text{G} \quad \text{G} \\
& \quad \text{a} \quad \text{a} \quad \text{a} \quad \text{a} \\
& \quad \text{c} \quad \text{c} \quad \text{c} \quad \text{c} \\
\end{align*}
\]

Bases drawn with small letters yielded a weaker binding of the protein. All other capitalized bases could stand at the position without altering the binding efficiency.

Identification of other sugar box regulated genes in the B. subtilis genome

Using the programmes Sequence Search and Gene Finder (S. Preis, unpublished) the B. subtilis genome was searched for further sugar-box regulated genes. The search parameters were defined such that all sequences could be found which deviated in either none or only one position from the consensus sequence. The search resulted in 1076 sequences in the entire B. subtilis genome, among which 76 of these sequences were located upstream of promoters. Ten of these sequences were chosen for which the distance of the sugar box to the transcriptional start site was similar to that of SR1 and for which a regulation of the downstream genes by glucose or other nutrients was already published or assumed (Table 2). The corresponding labelled oligonucleotides were tested in EMSAs. Only two of the analysed oligonucleotides, gapB and pckA, showed a binding activity in EMSA (Fig. 5B). Interestingly, the sugar box variant upstream of the gapB gene showed a 20-fold higher binding efficiency compared with the SR1 sugar box.

Identification of the regulatory protein binding to the sugar box

To identify the protein binding to the sugar box, a 3’ biotinylated oligonucleotide containing two copies of the gapB sugar box coupled to streptavidine magnetic bead conjugates was used. With the help of these immobilized oligonucleotides the target protein was isolated from a B. subtilis protein crude extract as described in Experimental procedures. About 1.4 µg of the eluted protein was separated on a 12% SDS PAA gel. The only visible band was cut out and subjected to tryptic digestion followed by N-terminal sequence determination. With this approach, the protein was identified as YqzB, which was recently discovered by another approach and termed CcpN (Servant et al., 2005).

Analysis of sugar-mediated repression of SR1 in a ccpN knockout strain

To analyse whether CcpN is indeed responsible for the observed sugar-mediated repression effect of SR1 transcription, ccpN knockout strains of B. subtilis DB104 were constructed by replacing ccpN by either a chloramphenicol or a phleomycin cassette (see Experimental procedures). To analyse the repression of SR1 transcription in the ccpN knockout strain, wild-type and knockout strains were grown in parallel in TY medium with or without 2% glucose, RNA was prepared and Northern blots performed. As shown in Fig. 6A, glucose-mediated repres-

<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene product</th>
<th>‘Sugar box’ sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>gapB</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase (B)</td>
<td>TTAATGTGTTATCTAATT</td>
</tr>
<tr>
<td>citB</td>
<td>Acotinate</td>
<td>TTAATGTGTTATGAGAG</td>
</tr>
<tr>
<td>pckA</td>
<td>Phosphoenolpyruvate carboxykinase</td>
<td>AATATATGTTATACTAATT</td>
</tr>
<tr>
<td>ssnA</td>
<td>tmRNA</td>
<td>GGCCTGTGCTATATTTAAT</td>
</tr>
<tr>
<td>nadE</td>
<td>NH3+-dependent NAD+ synthetase</td>
<td>GCTATGTGACAGACTATGT</td>
</tr>
<tr>
<td>thyA</td>
<td>Thymidylate synthase A</td>
<td>TTATGTTGACAGACTATGT</td>
</tr>
<tr>
<td>thyB</td>
<td>Thymidylate synthase B</td>
<td>TTAATGTGTTACACTACTA</td>
</tr>
<tr>
<td>clpP</td>
<td>ATP dependent Clp protease (proteolytic SU)</td>
<td>GCAATCTGACATACGAATGG</td>
</tr>
<tr>
<td>clpX</td>
<td>ATP dependent Clp protease (ATP binding SU)</td>
<td>GTTTTTGTGTCATACATA</td>
</tr>
<tr>
<td>lctE</td>
<td>Lactate dehydrogenase</td>
<td>TAAATGTGAAATACTTCCAC</td>
</tr>
</tbody>
</table>

© 2005 Blackwell Publishing Ltd, Molecular Microbiology, 58, 189–206
sion was significantly impaired in the ccpN knockout strain compared with the wild-type strain. However, a residual repression could still be observed in the ccpN knockout strain, indicating that CcpN is not the only protein responsible for the sugar-mediated repression. The sr1 expression level in the absence of glucose was not altered in the knockout strain compared with the wild-type.

To find out whether the residual repression was caused by a region upstream of the sugar box, a double knockout strain (Δsr1, ΔccpN) was transformed with pGKSR22 comprising the sr1 gene with a 22 bp upstream region of pSR1. Figure 6B shows a comparison of sugar-mediated SR1 repression between the Δsr1 strain and the Δsr1, ΔccpN double knockout strain in the presence of pGKSR22. In the double-knockout strain no glucose repression was observed anymore indicating that a second regulatory element upstream of the sugar box must be present in the B. subtilis chromosome.

Furthermore, SR1 proved to be transcribed in the logarithmic phase in TY medium when no CcpN was present, whereas no SR1 was transcribed in the wild-type strain under the same conditions (not shown). This result shows that CcpN represses SR1 transcription before onset of the stationary phase, most likely via the small amounts of sugar present in yeast extract which is one component of the TY medium.

CcpA-dependent catabolite repression is the second level of SR1 regulation by various sugars

Inspection of the sequence further upstream of the sr1 promoter revealed a putative cre element with the
sequence 5'-TGAAACCGTTTTTA 239 bp upstream of the −35 box of psr1. To find out whether or not this putative cre element which deviates in two positions (6th and 13th) from the consensus cre site (5'-TGWAARCGYTWNC; Stülke and Hillen, 2000) contributes to sugar-mediated repression of SR1 transcription, both single and double knockout strains for ccpA and ccpN were constructed and glucose-dependent repression of SR1 was analysed.

Glucose-dependent repression could not be observed anymore in the double knockout strain (Fig. 6C). The same result was obtained for a ccpN/ptsH double knockout strain (data not shown).

To support these data about the involvement of CcpA-dependent catabolite repression, a mutated or a wild-type cre element were placed upstream of pSR1. As the sr1 gene could not be cloned in E. coli when more than 100 bp upstream of the −35 box of pSR1 were present, the sequence of a heterologous RNA (RNAIII of plasmid pIP501) downstream from the sr1 promoter was used as a reporter. The reporter constructs with sr1 wild-type (5'-TGA AAC CGT TTT TA) or mutated cre element (sequence of the mutated cre element: 5'-GTT GCT ATG GCG AC) upstream of wild-type pSR1 were integrated into the amyE locus of the B. subtilis chromosome of a ccpN knockout strain. Northern analysis with a probe against RNAIII revealed a small but measurable 1.7-fold contribution of the cre element to sugar-mediated repression of pSR1 (Fig. 6D).

As sr1 is not expressed in Spizizen minimal medium containing 0.5% glucose, we asked, whether glucose present in this medium represses transcription of Psr1 via ccpN or ccpA or both. To answer this question, a wild-type, a ccpN-knockout, a ccpA knockout and a ccpN/ccpA double knockout strain were grown in Spizizen minimal medium containing 2% glucose. The repression factor above each pair of bands indicates the expression level without glucose compared with the expression level with glucose at the corresponding OD. Here, only the chromosomal copy of SR1 is analysed. Expression of sr1 in DB104 and DB104 (ΔccpN::cat) with and without 2% glucose at OD560 = 4.0 or 5.0. The repression factor above each pair of bands indicates the expression level without glucose compared with the expression level with glucose at the corresponding OD. Here, only the chromosomal copy of SR1 is analysed. Expression of sr1 in DB104 (Δsr1::cat) (pGKSR22) and DB104 (Δsr1::cat, ΔccpN::phleo) (pGKSR22). Here, the chromosomal copy of sr1 is missing, and expression of the plasmid copy of sr1 is analysed. Expression of sr1 in DB104(ΔccpN::cat), in QB5407 (ΔccpA::spe) and in QB5407(ΔccpN::cat, ΔccpA::spe). Here, expression of the chromosomal copy of sr1 in the ccpN or the ccpN/ccpA background is monitored. Reprobing was done as described in Fig. 3.

Expression of RNAIII under control of pSR1 with wild-type or mutated cre element in a ccpN knockout strain.

Fig. 6. Analysis of sugar-mediated repression of plasmid or chromosome-encoded SR1 in ccpN, ccpA and ccpN/ccpA knockout strains. Autoradiographs of the Northern blots are shown.

A. Expression of sr1 in DB104 and DB104 (ΔccpN::cat) with and without 2% glucose at OD560 = 4.0 or 5.0. The repression factor above each pair of bands indicates the expression level without glucose compared with the expression level with glucose at the corresponding OD. Here, only the chromosomal copy of SR1 is analysed.

B. Expression of sr1 in DB104 (Δsr1::cat) (pGKSR22) and DB104 (Δsr1::cat, ΔccpN::phleo) (pGKSR22). Here, the chromosomal copy of sr1 is missing, and expression of the plasmid copy of sr1 is analysed.

C. Expression of sr1 in DB104(ΔccpN::cat), in QB5407 (ΔccpA::spe) and in QB5407(ΔccpN::cat, ΔccpA::spe). Here, expression of the chromosomal copy of sr1 in the ccpN or the ccpN/ccpA background is monitored. Reprobing was done as described in Fig. 3.

D. Expression of RNAIII under control of pSR1 with wild-type or mutated cre element in a ccpN knockout strain.

© 2005 Blackwell Publishing Ltd, Molecular Microbiology, 58, 189–206
medium with or without glucose, RNA was prepared and subjected to Northern blotting. The very same picture as with TY medium (Fig. 6C) was obtained (data not shown). This result indicates that the repression of SR1 transcription in Spizizen minimal medium is solely due to glucose and that the products of the same genes, ccpN and ccpA mediate this effect.

In summary, sugar-mediated repression of SR1 transcription is accomplished by two regulatory systems: the most important system is CcpN contributing to the majority of sr1 repression, the second, less important system is the CcpA dependent catabolite repression system comprising CcpA/HPr and the distant cre-element.

**DNase I footprinting experiments identify two sites protected by CcpN**

DNase I footprinting experiments with the sr1 promoter region were performed as described in Experimental procedures. The PCR fragment for the coding strand was generated with labelled primer SB753 and unlabelled primer SB346, whereas the PCR fragment for the non-coding strand was obtained with labelled primer SB826 and unlabelled primer SB753. As shown in Fig. 7, two protected sites were visible for both strands, out of which one (site I) covered the previously identified sugar box (Fig. 7C), whereas the other one (site II) overlapped part of the –35 box, the spacer and the entire –10 box of the sr1 promoter, but not the transcription start site. A similar increase of the protection of both sites upon increasing CcpN concentration indicated a similar affinity of the protein for both sites. The induction of one DNase I hypersensitive site in both the coding and the non-coding strand in the centre of the sugar box may suggest a conformational change in the DNA upon CcpN binding.

**Discussion**

Here, we report the detection of a novel small non-translated RNA designated SR1 (small RNA 1) encoded in the pdhD-yktA-intergenic region of the B. subtilis chromosome and its regulation by various sugars accomplished...
by two components, the recently identified CcpN and the catabolite control protein CcpA.

A computer based approach was used to predict its RNA, and subsequent Northern blotting confirmed its expression in vivo. SR1 is 205 nt long and expressed under control of a typical ς^A promoter. It contains an ORF for a putative 39 aa peptide with a non-typical SD sequence. The comparison of translational and transcriptional lacZ fusions of the SR1-encoded peptide (Table 1) clearly demonstrated that this peptide is not translated in B. subtilis under single-copy conditions from the chromosome. Therefore, SR1 is a small non-translated RNA. Inspection of all published genomes for homologues of SR1 or the SR1 ORF revealed sequences of high similarity in Bacillus licheniformis, Geobacillus kaustophilus and at least sequences coding for homologous peptides in Bacillus cereus and Bacillus anthracis (Fig. 8). Whereas in all cases, the ORF with its putative SD sequence is present, only in two other cases, B. licheniformis and G. kaustophilus, this sequence is preceded by a typical ς^A promoter. DNA sequence homology to the sr1 gene was only significant in the peptide encoding stretch in all compared cases. No homologues for SR1 or the SR1 ORF were found in the other sequenced Bacillus genomes like B. halodurans or B. thuringiensis.

SR1 shows an interesting expression profile. Expression of SR1 proved to be maximal under conditions of gluconeogenesis, e.g. in minimal medium without glucose. In TY complex medium that contains small amounts of glucose present in the yeast extract, SR1 is expressed maximally when cells enter stationary phase, i.e. after exhaustion of the sugar. An induction of SR1 expression by increasing cell density or oxygen starvation, as characteristic for stationary phase stress, too, was not observed (data not shown). In minimal medium without glucose, SR1 is expressed constitutively, whereas no expression was found in the presence of glucose and only a weak expression in the presence of glucose at 48°C in late stationary phase (overnight). This expression pattern can be explained by glucose-mediated repression of SR1 transcription. The expression of SR1 at 48°C might result from weaker glucose transport at this temperature, which, in turn, is probably due to the intrinsic thermal instability of the RNA switch element (RAT) that is required to form to allow expression of the glucose transporter gene, ptsG. The fact that SR1 is only expressed under certain conditions is in good agreement with the expression pattern of the majority of the recently detected small RNAs from the E. coli genome, which are also found mainly in stationary phase (Argaman et al., 2001). Apparently, SR1 has a function under conditions of gluconeogenesis, perhaps in fine-tuning of metabolism.

SR1 transcription was found to be repressed by the addition of glucose or various other sugars, like fructose, sucrose, arabinose and glycerol, i.e. under glycolytic conditions. The region responsible for sugar-mediated repression was narrowed down to 17 bp (Fig. 4C) and, provisionally, designated sugar box. Gel shift assays with B. subtilis crude protein extracts showed that a protein bound to this box. Using mutated 23 bp DNA targets in an EMSA, a provisional consensus sequence for a binding factor was derived. Searching the B. subtilis genome with this consensus motif, 10 putative sugar-box regulated genes were found, out of which only sr1, gapB and pckA seemed to be regulated by this box (Fig. 5B). With the help of a 3′ biotinylated oligonucleotide containing two copies of the gapB motif and streptavidine magnetic beads, the protein bound to the sugar box could be purified and – using tryptic digestion and mass spectrometry – identified as the recently published CcpN (Servant et al., 2005). Interestingly, Servant et al. found a CcpN regulation of gapB and pckA, too, but predicted, by a bioinformatics approach and a transcriptome analysis (which were not described in the article) that these are the only genes regulated by CcpN. It cannot be excluded that other still unidentified genes encoding small RNAs might be regulated by CcpN, too, because such genes would escape the conventional transcriptome or bioinformatics based searches.

A comparison of the CcpN DNase I footprints at the gapB, pckA and the sr1 promoters shows some differences: In the case of gapB, both binding sites are located downstream from the –35 box, OP_{gap} overlapping the –10 box and OP_{gap} being located downstream from the transcription start site (Servant et al., 2005). In the pckA case, only one large binding site was found, which, however, overlaps both –35 and –10 box and only the first 3 nt of the pckA-mRNA. Only 6 nt of the region upstream of the –35 box were protected in the coding and 9 nt in the non-coding strand. By contrast, footprinting of the sr1 promoter region revealed two CcpN binding sites, out of which site I is almost identical with the previously identified sugar box and located clearly upstream of the –35 box, whereas site II overlaps the 3′ part of the –35 box, the spacer and the –10 box, but not the transcriptional start site. Six nucleotides of the strongly conserved sequence 5′-TATACTA covering the –10 box in gapB and pckA are part of the sugar box that we identified for sr1, namely 5′-TATACA, but this sequence does not occur in binding site II. However, as shown by a combination of EMSA and Northern blot experiments, the minimal box required to be bound comprises at least 17 bp. Furthermore, a sequence comparison between the CcpN binding sites of gapB, pckA and sr1 shows that no inverted repeat is present in the latter case. To unequivocally identify the bases contacted by CcpN, we will perform methylation, ethylation and missing base interference experiments. A possible explanation for the difference in the location of the CcpN
Fig. 8. Alignments of the sr1 DNA and SR1 ORF sequence with similar sequences from other bacteria. A. DNA sequence alignments. Above B. subtilis sequence, below the sequences from Bacillus licheniformis, Geobacillus kaustophilus, Bacillus cereus and Bacillus anthracis. The ccpN operator present in all genomes is underlined in the sequence, below the sequences from B. subtilis, B. anthracis and in the other bacteria. The following symbols indicate: asterisk (*), the residues or nucleotides in that column are identical in all sequences in alignment; colon (:), conserved substitutions; full point (.), semi-conserved substitutions. The alignments were based on the sequences published by Veith et al. (2004) (Bacillus licheniformis); Ivanova et al. (2003) (Bacillus cereus and Bacillus anthracis), Takami et al. (2004) (Geobacillus kaustophilus) and Ravel et al. (2004) (Bacillus anthracis).

binding sites in the three known cases might be the deviation of the gapB and pckA−35 boxes from the consensus −35 box TTGACA, which is, in contrast, perfectly conserved in pSR1. One could imagine that the repression mechanism used by CcpN is different in the gapBiPckA promoter compared with the sr1 promoter. Experiments are under way to elucidate the repression mechanism of CcpN at the sr1 promoter.

Analysis of a ccpN knockout strain, a ccpA knockout strain and a double knockout strain (Fig. 6) revealed that,
although CcpN is responsible for the majority of sugar-mediated repression, CcpA as a second factor is involved in sugar-mediated repression in the absence of CcpN. Indeed, a cre site located 275 bp upstream of the SR1 transcription start site was found. Although in the majority of cases studied so far (Stülke and Hillen, 2000), cre elements are located either within the promoter region (e.g. bglPH, acu, amyE), where they hinder transcription initiation or downstream from the transcription start site (e.g. acsA, xyl, hut, xynPB) where they might block elongation or interact with the RNA polymerase by repressor loop formation, one case, the lev operon, was published were the cre element is located between −50 and −36 upstream of the transcription start site. Furthermore, the sr1 cre element deviates in two positions from the consensus sequence, although the mutation in position 13 (C to T) has been observed in one other case (xynCB; Galinier et al., 1999) where nevertheless an efficient glucose repression was reported. The mutation in position 6 was, however, rarely, found in functional cre elements, too (dra-nupC-pdp-operon; Zeng et al., 2000). The combination of a long distance between cre site and sr1 promoter and the deviation from the consensus sequence might be responsible for the relatively small effect of the catabolite repression system on the sugar-mediated repression of the sr1 gene. One could speculate that loop formation in the pSR1 upstream region promoted by a still unidentified DNA-binding and bending protein might enable a direct contact between CcpA/HPr−Ser-P bound to the cre element and the RNA polymerase bound to the sr1 promoter, thus inhibiting transcription initiation or promoter escape. Such a direct contact might be prevented when CcpN is bound at its operator site I upstream of the sr1 promoter explaining why the CcpA effect is only observed in the absence of CcpN. Interestingly, repressor loop formation and a direct interaction between CcpA bound to a downstream cre site and the RNA polymerase has been proposed recently for the xyl promoter (Kim et al., 2005).

CcpN proved to be present in crude extracts throughout growth in TY medium indicating that this protein is expressed constitutively. As the repressor CcpN is always present, but repression is only observed under conditions of glycolysis, there must be a process by which CcpN senses the metabolic situation in the cell and – most probably by a ligand available only when sugars are present – represses SR1 transcription. The facts that (i) the expression level of SR1 in a ccppN:cat strain was comparable to that of the wild-type strain and (ii) CcpN bound to the DNA in vitro in the absence of a ligand (EMSA experiments) argue for a ligand that increases the DNA-binding ability of CcpN or, alternatively, a ligand required for a conformational change that results in an efficient contact with RNA polymerase. Similar hypotheses about the involvement of a positive cofactor or corepressor have been suggested by Servant et al. (2005).

CcpN was found to have two CBS domains (aa 86–134 and aa 153–204), present in all three kingdoms of life. Recently, it could be shown that CBS domains have an affinity to adenosine phosphates and thereby, a specificity for AMP (Scott et al., 2004). This prompted us to look for effects of ATP, ADP, AMP and other analogues or purines on the ability of CcpN to bind DNA in EMSAs and by protein UV crosslinking or chamber dialysis. However, until now, our various attempts to identify this ligand failed. Inspection of sugar-mediated sr1 expression in a ΔrelA strain showed that ppGpp is not involved. Also without success, we tested various intermediates of the sugar metabolism such as PEP, pyruvate, citrate, glyceraldehyde-3-phosphate for their ability to alter the binding behaviour of CcpN in EMSAs.

Interestingly, recently, a novel small RNA, SgrS, induced by phosphosugar stress from the E. coli chromosome, was published that was found to be regulated also by a novel transcriptional regulator, SgrR (Vanderpool and Gottesman, 2004). In contrast to CcpN, SgrF is an activator of SgrS transcription and the sgrR gene is located adjacent to the sgrS gene and is transcribed divergently to sgrS. SgrR was suggested to be a sensor of phosphosugars and has a solute binding domain. However, it has still to be shown if phosphosugars or other ligands bind to this domain. As in the case of SgrS, SR1 or RyhB one could imagine a lot of other still unidentified regulatory circuits that might involve the post-transcriptional action of a small RNA and a transcriptional regulator that turns on or off the expression of this RNA in response to changes in the environmental conditions.

What is the function of SR1 in B. subtilis? Because a knockout strain demonstrated that SR1 is not essential for B. subtilis, we propose a role for this RNA under conditions of gluconeogenesis, perhaps in the fine-tuning of metabolism. Using two-dimensional protein gel electrophoresis, we could identify three putative targets of SR1, which proved, in further investigations, to be downstream targets. Further analyses which will be published elsewhere will show that SR1 is a regulatory RNA acting by direct base pairing with its target RNA.

Experimental procedures

Enzymes and chemicals

Chemicals used were of the highest purity available. Taq DNA polymerase was purchased from Roche or SphaeroQ, the Netherlands, RNA ligase from New England Biolabs and Thermoscript reverse transcriptase from Invitrogen. Sequencing reactions were performed according to Sanger et al. (1977) using a Sequenase kit from Amersham Bioscience.
**Strains, media and growth conditions**

*Escherichia coli* strain TG1 was used for cloning and *B. subtilis* strain DB104 (Kawamura and Doi, 1984) was used for the isolation of total RNA under the corresponding conditions. *B. subtilis* strain QB5407 (ccpA::spc) (Faires et al., 1999) was used for the construction of a double knockout strain. TY medium (16 g Bacto tryptone, 10 g Yeast extract, 5 g NaCl in 1 l) was used as a complex medium and Spizizen minimal medium was used as a minimal medium.

**Construction of plasmids for transcriptional and translational lacZ fusions**

Plasmid pAC6 (Table 3) was used to insert a EcoRI-BamHI fragment obtained by PCR from chromosomal DNA of *B. subtilis* DB104 with oligodeoxyribonucleotides SB827 (Table S1) and SB831 to obtain a transcriptional fusion of the sr1-promoter with 87 bp upstream of the −35 box and the promoterless lacZ gene. The resulting plasmid pACT87 was integrated into the amyE locus of strain DB104 and double crossing over was confirmed by streaking the chloramphenicol resistant transformants on agar with 0.5% starch and subsequent overlay with iodine/potassium-iodide solution.

For the construction of translational sr1 lacZ fusions, three different fragments were obtained by PCR from chromosomal DNA of *B. subtilis* DB104 with primer pairs SB827/SB828, SB421/SB828 or SB422/SB828 and inserted into the BamHI/EcoRI (pACP87) or BamHI (pACP421 and pACP422) vector of pAC5 respectively. In all cases, the promoterless, SD-less lacZ gene was fused in frame with the 3rd codon of the putative sr1-encoded peptide. The resulting vectors pACP421, pACP422 and pACP87 contain 250, 150 and 87 bp, respectively, upstream of the −35 box of pPan1. All three plasmids were integrated into the amyE locus in parallel with the empty vector pAC5, and double crossing over was confirmed as above.

**Construction of plasmids for sr1 knockout and overexpression strains**

To obtain a sr1 knockout strain, plasmid pINT1 was constructed in the following way. First, DNA fragments upstream of and downstream from the sr1 gene were amplified by PCR using chromosomal DNA of *B. subtilis* DB104 and oligodeoxyribonucleotides SB324 and SB325 as well as oligodeoxyribonucleotides SB326 and SB327 and − after digestion with the appropriate restriction enzymes, inserted as 800 bp BamHI/EcoRI or 820 bp SalGI/PstI fragments into the corresponding pUC19 vectors cleaved with the same pair of enzymes. This approach resulted in pFRONT1 and pBACK1. The chloramphenicol acetyl transferase gene was obtained from the previously constructed vector pCOP7C (S. Brantl, unpublished) as 1.4 kb EcoRI/Sal fragment and inserted into the pUC19 EcoRI/SalI vector resulting in plasmid pCAT. Next, plasmid pBACK was constructed by jointly cloning the CAT-

---

**Table 3. Plasmids used in this study.**

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pUC19</td>
<td><em>E. coli</em> cloning vector, Ap&lt;sup&gt;n&lt;/sup&gt;, MCS</td>
<td>Sambrook et al. (1989)</td>
</tr>
<tr>
<td>pGK14</td>
<td>Broad host range lactococcal vector, Em&lt;sup&gt;n&lt;/sup&gt;</td>
<td>Brantl (1994)</td>
</tr>
<tr>
<td>pPR1</td>
<td><em>E. coli</em>/<em>B. subtilis</em> shuttle vector, Ap&lt;sup&gt;h&lt;/sup&gt;, phleo&lt;sup&gt;h&lt;/sup&gt;</td>
<td>Brantl and Behnke (1992)</td>
</tr>
<tr>
<td>pWH353</td>
<td><em>E. coli</em>/<em>B. subtilis</em> shuttle vector, Tet system, Km&lt;sup&gt;n&lt;/sup&gt;</td>
<td>Geissendörfer and Hillen (1989)</td>
</tr>
<tr>
<td>pAC6</td>
<td>pBR322 based vector for integration of transcriptional lacZ fusions into amyE locus of <em>B. subtilis</em>, Ap&lt;sup&gt;h&lt;/sup&gt;, Cm&lt;sup&gt;n&lt;/sup&gt;</td>
<td>Stülke et al. (1997)</td>
</tr>
<tr>
<td>pAC5</td>
<td>pBR322 based vector for integration of translational lacZ fusions in to amyE locus of B. subtilis, Ap&lt;sup&gt;h&lt;/sup&gt;, Cm&lt;sup&gt;n&lt;/sup&gt;</td>
<td>Stülke et al. (1997)</td>
</tr>
<tr>
<td>pCOP7C</td>
<td>pCOP7 with CAT gene as EcoRI/Sal fragment</td>
<td>S. Brantl, unpublished</td>
</tr>
<tr>
<td>pCAT</td>
<td>pUC19 with CAT gene</td>
<td>This study</td>
</tr>
<tr>
<td>pFRONT1</td>
<td>pUC19 with 800 bp upstream of sr1</td>
<td>This study</td>
</tr>
<tr>
<td>pBACK1</td>
<td>pUC19 with 820 bp downstream of sr1</td>
<td>This study</td>
</tr>
<tr>
<td>pBACK2</td>
<td>pBACK1 with CAT gene from pCAT</td>
<td>This study</td>
</tr>
<tr>
<td>pINT1</td>
<td>Vector for substitution of sr1 by CAT gene</td>
<td>This study</td>
</tr>
<tr>
<td>pINT1P</td>
<td>As pINT1, but with Phleo&lt;sup&gt;h&lt;/sup&gt; instead of CAT gene</td>
<td>This study</td>
</tr>
<tr>
<td>pUCSR1</td>
<td>pUC19 with promoterless sr1 gene (nt 1037–1246)</td>
<td>This study</td>
</tr>
<tr>
<td>pWSR1</td>
<td>pWH353 with sr1-HindIII-fragment from pUCSR1</td>
<td>This study</td>
</tr>
<tr>
<td>pUCSR1/2</td>
<td>pUC19 with sr1 gene (nt 901–1246)</td>
<td>This study</td>
</tr>
<tr>
<td>pGKS1</td>
<td>pGK14 with pUCSR1/2 BamHI/HindIII fragment</td>
<td>This study</td>
</tr>
<tr>
<td>pGKS1X</td>
<td>pGK14 with sr1 BamHI/HindIII fragments, where X denotes the number of bases upstream of pPan1</td>
<td>This study</td>
</tr>
<tr>
<td>pGKS30/Y</td>
<td>pGKS30 with mutated pPan1 upstream regions, Y denotes the number of exchanged bases upstream of pPan1</td>
<td>This study</td>
</tr>
<tr>
<td>pINT4C</td>
<td>Vector for substitution of ccppN gene by CAT gene</td>
<td>This study</td>
</tr>
<tr>
<td>pINT4P</td>
<td>Vector for substitution of ccppN gene by Phleo&lt;sup&gt;h&lt;/sup&gt; gene</td>
<td>This study</td>
</tr>
<tr>
<td>pQGDR</td>
<td>Vector for overexpression of C-terminally His-tagged CcppN in <em>E. coli</em></td>
<td>This study</td>
</tr>
<tr>
<td>pACT87</td>
<td>pAC6 with pPan1 and 87 bp upstream of −35 box</td>
<td>This study</td>
</tr>
<tr>
<td>pACP87</td>
<td>pAC5 with pPan1 and 87 bp upstream of −35 box and first 3 codons of putative SR1 ORF fused in frame with lacZ</td>
<td>This study</td>
</tr>
<tr>
<td>pACR1</td>
<td>pAC6 with pPan1 and wild-type cre fused to rmalIII</td>
<td>This study</td>
</tr>
<tr>
<td>pACR2</td>
<td>As pACR1, but with mutated cre site</td>
<td>This study</td>
</tr>
</tbody>
</table>

Plasmid pWSR1 used for inducible overexpression of sr1 in *B. subtilis* was constructed as follows: the promoterless sr1 gene was amplified by oligodeoxynucleotides SB316 and SB317 using 1 μl of chromosomal DNA of *B. subtilis* 168 as template, digested with HindIII and inserted into the pUC19 HindIII-vector resulting in plasmid pUCSR1. The insert was confirmed by sequencing. Subsequently, the pUCSR1 HindIII fragment was inserted into the unique HindIII site of vector pWH353 (Geissendörfer and Hillen, 1989) downstream of the Tet promoter yielding plasmid pWSR1, and the correct insert orientation was verified by sequencing. Plasmid pGKSR1 used for constitutive overexpression of sr1 in *B. subtilis* under its own promoter was obtained in the following way. First, pUCSR1/2 was constructed by cloning a PCR fragment obtained on chromosomal DNA from *B. subtilis* 168 as template with primers SB350 and SB317 into the pUC19 BamHI/HindIII vector. The insert sequence was confirmed. Afterwards, the BamHI/HindIII fragment of pUCSR1/2 was inserted into the BamHI/HindIII vector of pGK14 (Brantl, 1994) resulting in plasmid pGKSR1.

### Construction of plasmids for the analysis of sr1 genes with mutated upstream regions

Plasmid pGKSR50 containing the sr1 gene with 50 bp upstream of the −35 box was constructed in the following way: A BamHI/HindIII fragment obtained by PCR with primers SB317 and matutagenic primer SB423 on chromosomal DNA of *B. subtilis* DB104 as template was inserted into the pGK14 BamHI/HindIII vector resulting in pGKSR50. The insert was confirmed by sequencing.

All other pGKSRX vectors were obtained in the same way, except that primer SB317 in combination with the corresponding mutagenic primer listed in Table S1 was used. For plasmids used to map the upstream border of the sugar box, X designates the number of base pairs upstream of the −35 box of ρSR. Plasmids constructed to map the downstream border of the sugar box were termed pGKSR30/γ, where Y is the number of altered bases directly upstream of the −35 box of ρSR. To construct plasmids pACR1 and pACR2 used for the analysis of the putative cre element, first, the promoterless *rnall* gene was cloned as a XbaI/EcoRI fragment into the pUC19 vector using SB783 and SB784 and plasmid pPR1 as a template. In parallel, the sr1 wild-type and a mutated sr1 promoter regions were cloned as BamHI/XbaI fragments with chromosomal DNA of *B. subtilis* DB104 as template and primer pairs SB421/SB785 (wild-type cre) and SB751/SB785 (mutated cre), respectively, into pUC19. Subsequently, the corresponding BamHI/XbaI sr1-promoter fragments and the XbaI/EcoRI *rnall* fragment were jointly inserted into the pAC6 BamHI/EcoRI vector resulting in pACR1 (wild-type cre) and pACR2 (mutated cre).

### Construction of plasmids for ccpN knockout and overexpression strains

The construction of plasmids pINT4C and pINT4P used to obtain *ccpN* knockout strains was performed as follows: First, the fragments FRONT4 (upstream of *ccpN*) flanked by BamHI and EcoRI and BACK4 (downstream from *ccpN*) flanked by SalI and PstI were obtained by PCR on chromosomal DNA of *B. subtilis* DB104 using the primer pairs SB660/SB661 and SB662/SB663, respectively, were cloned separately into the corresponding pUC19 vectors and confirmed by sequencing. Afterwards, the BamHI/EcoRI fragment FRONT4 and the EcoRI/SalI fragment with the chloramphenicol cassette (see above) were jointly cloned into the BamHI/SalI vector of pUC19. Subsequently, the BamHI/SalI fragment of this pUC19 vector was inserted together with the SalI/PstI fragment BACK4 into the BamHI/PstI vector of pUC19 resulting in pINT4C. In the case of integration vector pINT4P, the EcoRI/SalI fragment encoding the phleomycin resistance gene that was previously obtained by PCR from PUB110 was used instead of the chloramphenicol resistance cassette.

A *ccpN* overexpression strain was constructed by cloning a Ncol/BgIII digested PCR fragment obtained with primers SB673 and SB674 on chromosomal DNA of *B. subtilis* DB104 into the pQE60 Ncol/BgIII vector (Qiagen). The resulting vector was designated pQGDR. For cloning and subsequent purification of the his-tagged protein, *E. coli* strain TG1 (REP4) was used. The sequence was confirmed.

### Isolation of chromosomal DNA from *B. subtilis*

Two millilitres of an overnight culture of *B. subtilis* or 5 ml of a logarithmically grown culture were centrifuged and pellets suspended in 0.5 ml TEG buffer (Tris-HCl, pH 7.5, glucose, 25 mM EDTA) containing 2 mg ml⁻¹ lysozyme. After incubation at 37°C for 30 min, 2 μl of RNase A (10 mg ml⁻¹) were added and incubation was continued for 10 min at 37°C. Cells were lysed with 25 μl 10% SDS, 26 μl pronase solution (20 mg ml⁻¹, 1 h predigested at 37°C) were added and the mixture was incubated for 20–50 min at 37°C. Subsequently, two to five (depending on the amount of interphase) extractions with 400 μl phenol/chloroform (1:1) were performed to remove proteins and cellular debris followed by one extraction with chloroform only. The supernatant of the last chloroform extraction was precipitated with ethanol, the pellet was washed with 200 μl 80% ethanol and suspended in 40–50 μl bidistilled water. Two microlitres were loaded onto an 1% agarose gel to estimate the concentration.

### In vitro transcription

A riboprobe for SR1 was synthesized *in vitro* by T7 RNA polymerase from a DNA template generated by PCR with oligonucleotides SB315 and downstream primer SB303 in the presence of 10 μCi of [α³²P]-UTP as described previously (Brantl and Wagner, 1996). A riboprobe for the complementary strand was generated in the same way, but using a PCR fragment obtained with SB314 combined with downstream primer SB309.
Isolation of total RNA for Northern blot analyses

Overnight cultures of *B. subtilis* strains were diluted 100-fold and grown in TY or minimal medium. At different optical densities (OD_{600} of 0.2, 0.5, 1, 2, 4, 5, 6 or 12) between 0.5 (OD 5–12) and 2 ml (OD 0.2) of each culture were immediately frozen in liquid nitrogen. Frozen samples were stored at –20°C for later preparation of total RNA. Centrifuged cells were suspended in 100 µl lysis buffer 1 (100 mM NaCl, 50 mM EDTA, 10% sucrose, 1 mg ml\(^{-1}\) lysozyme). After 5 min at 37°C, 300 µl lysis buffer 2 (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM NaOAc, 2% SDS) were added and one hot phenol extraction (65°C) and two or three cold phenol/chloroform extractions were performed. Supernatant was precipitated with 1/10 volume 3 M NaOAc pH 7.0 and two volumes ethanol. After centrifugation, the RNA was dissolved in 20 µl DEPC treated water and stored at –20°C.

Northern blot analysis

For the analysis of putative small RNA species, 5 µl total RNA were denatured at 95°C with 5 µl formamide loading dye, FD (90% formamide, 15 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol), and subsequently separated on a 6% denaturing polyacrylamide gel. The RNA was electrophertransferred onto Biodyne A transfer membrane (0.2 µm) from Pall corporation at 30 V for 15 h. Transfer was performed in a BioRad blotting chamber in 1× TBE buffer (89 mM Tris, 89 mM boric acid, 0.49 mM EDTA, pH 8.3) at 4°C.

Subsequently, the RNA was dried and bound to the membrane by UV-crosslinking. Prehybridization was carried out for 2–4 h at 62°C in 10 ml prehybridization buffer (6× SSC, 5× Denhardt, and 0.5% SDS, 0.3 mg ml\(^{-1}\) salmon sperm DNA). Hybridization was performed overnight at 62°C in the same buffer lacking salmon sperm DNA but containing 1–2 × 10⁶ cpm ml\(^{-1}\) of labelled probe ([\(\gamma^{32}\)P]-dATP-labelled double-stranded DNA-fragment generated by PCR or [\(\gamma^{32}\)P]-ATP-labelled oligonucleotides). Membranes were washed once for 30 min at 62°C in 2× SSC, 0.5% SDS and once for 30 min in 0.5× SSC, 0.5% SDS. Signals were quantified in a Fuji PhosphorImager. Removal of labelled probe was done by boiling (10 min in 1% SDS) followed by prehybridization. To correct for loading errors, reprobing was carried out with [\(\gamma^{32}\)P]-ATP-labelled oligonucleotide C767-27 (5′GGG TGT GAC CTC TTC GCT ATC GCC ACC) that is complementary to *B. subtilis* SS rRNA; both prehybridization and hybridization were performed at 62°C in 6× SSC, 5× Denhardt, and 0.5% SDS (probe: 3 × 10⁵ cpm ml\(^{-1}\) of C767-27). Filters were washed and quantified as described above. All Northern blot analyses and subsequent calculations were performed at least twice on total RNA isolated independently from different cultures.

Primer extension

Oligodeoxyribonucleotide SB346 was 5′ end-labelled with T4 polynucleotide kinase and hybridized to total RNA prepared from *B. subtilis* strain 168 (pGKSR1) by incubation at 65°C for 5 min followed by quick cooling on ice. Reverse transcription was performed using the Thermoscript RT system (Invitrogen) in three subsequent 20 min steps at 55°C, 60°C and 65°C. The products were heat-denatured with formamide loading dye for 5 min at 95°C, quickly cooled on ice and loaded onto a 6% denaturing polyacrylamide gel along with a sequencing reaction with the same oligonucleotide.

3′ RACE

Total RNA was prepared as described above, treated with DNase and dephosphorylated with calf intestinal phosphatase. ~15 µg of dephosphorylated RNA were used in an *in vitro* ligation reaction with 500 pmol of the 3′ RNA adapter (E1, 5′-phosphate-UUC ACU GUU CUU AGC GGC CGC AUG CUC-idT-3′ [Dharmacon Research]; 3′idT, 3′ inverted deoxythymidine). After 5 min at 95°C followed by quick cooling on ice (5 µg RNA and 500 pmol Adapter), the reaction was performed for 12 h at 17°C in a total volume of 50 µl containing 50 U of T4 RNA ligase (New England Biolabs) and NEB ligase buffer supplemented with 10% DMSO and ATP (total concentration of 150 µM). After phenol and chloroform extractions and ethanol precipitation, 5 µg RNA were reverse transcribed in a total volume of 20 µl with 100 pmol of primer SB400 complementary to the 3′ adapter in three subsequent 20 min steps at 50°C, 55°C and 60°C with Thermoscript RT system (Invitrogen). The products of reverse transcription were amplified by the use of 1 µl of the 20 µl reaction with 25 pmol of each of the adapter specific primer SB400 and mRNA specific primer SB316 (sr1) or SB361 (uppr-mRNA), respectively, for 30 cycles (94°C/30 s, at 52°C/30 s and 72°C/30 s) and native Taq polymerase (SphaeroQ, Holland). The products were purified from 3% agarose gels, and 1/10 was used for a second PCR reaction with SB400 and internal oligonucleotide SB401 (sr1). Fragments were isolated from 3% agarose gels, digested with EcoRI and HindIII and inserted into the pUC19 EcoRI/HindIII vector. Colony PCR of white colonies was used to identify clones with the appropriate insert size. The corresponding recombinant plasmids were isolated and sequenced using the reverse sequencing primer. All enzymatic treatments of RNA were performed in the presence of 1 µl of the RNase inhibitor RNAGuard (Amersham Bioscience).

Electrophoretic mobility shift assay

Double-stranded DNA targets for the wild-type and mutated sugar box sequences were obtained in the following way: Each single oligonucleotide containing a mutation in the sugar box sequence (listed in Table S2, Supplementary material) was 5′ end-labelled, purified from a 8% denaturing gel and, subsequently, annealed with the same amount of the complementary non-labelled oligonucleotide. Binding reactions were performed in a final volume of 20 µl containing 1 nM labelled DNA target and either *B. subtilis* DB104 crude extracts prepared by sonication as published (Brantl, 1994) or 20–7000 nM purified Hissr1-tagged protein. After incubation at 30°C for 30 min in electrophoresis buffer (0.5× TBE), aliquots of the reaction mixtures were separated on native 8% polyacrylamide gels run at 4°C for 2 h at 230 V. Gels containing labelled DNA fragments were dried and, subsequently, visualized and quantified on a Fuji PhosphorImager.

**Isolation of CcpN using biotin-conjugated oligonucleotides**

A biotinylated oligodeoxyribonucleotide containing two copies of the gapB sugar box was annealed with a non-biotinylated complementary oligodeoxyribonucleotide in equimolar amounts in 0.5× TBE at a temperature gradient from 65 to 37°C. A corresponding amount of streptavidine magnetic beads (2 mg beads bind to 300 pmol ds oligonucleotides) was washed for three times in the initial bead volume with TEN100 buffer. Subsequently, the biotinylated oligonucleotide was incubated with the washed beads in the twofold initial volume of TEN100 at 20°C for 30 min. After incubation, the beads were separated, washed twice with TEN100 and equilibrated twice with 0.5× TBE. After confirmation of quantitative binding to the beads, a protein crude extract prepared from B. subtilis DB104 by sonication was incubated with the oligomagnetic bead conjugates for 1 h at 4°C under slight agitation. Subsequently, the beads were separated and washed twice in binding buffer. The bound protein was eluted with glycerol-free 0.5× TBE and high NaCl concentrations. The eluate was stored in 50% glycerol at −20°C.

**Purification of CcpN_{His} from an E. coli overexpression strain**

CcpN_{His} was purified from TG1(REP4, pQGDR) using the Nickel-NT agarose kit from Qiagen. The elution buffer was 50 mM Na-phosphate pH 8.0, 300 mM NaCl, 250 mM imidazole. Fractions were stored with 50% glycerol at −20°C.

**DNase I footprinting**

A DNA fragment was generated by PCR with a 5’ end-labelled primer and an unlabelled primer on plasmid pGKSR1 as a template, purified from a 6% polyacrylamide gel and dissolved in water at 10,000 cpm µl⁻¹. Footprinting was performed in 0.5× TBE containing 6.25 mM MgCl₂, 25 ng µl⁻¹ herring sperm DNA and 10 000–15 000 cpm labelled PCR fragment as well as increasing amounts of purified CcpN-His₅ protein (from 60 to 480 ng). After binding of CcpN-His₅ for 30 min at 30°C, the samples were treated with 1 µl of DNase I (Roche, 0.05 U µl⁻¹) for 2 min at 37°C. Two control samples, one without protein, one without DNase I, were treated in parallel. The reaction was stopped by phenol extraction and subsequent ethanol precipitation. The pellets were dissolved in 4 µl water, denatured with 4 µl formamide loading dye for 5 min at 75°C and separated on a 6% denaturing polyacrylamide gel along with a Sanger sequencing reaction obtained from the primer used for labelling as a marker. The dried gel was analysed by PhosphorImaging.

**Measurement of β-galactosidase activities**

β-Galactosidase activities were measured as described previously (Brantl, 1994).

**Acknowledgements**

The authors are grateful to Dr B. Scholl, I. M. B. Jena for performing the tryptic digestion and mass spectrometry and to Eckhard Birch-Hirschfeld, Institute for Virology, Friedrich-Schiller-University Jena, for synthesis of all oligodeoxyribonucleotides. We thank Alberto Chinali for Fig. 3C and for the alignment between the B. subtilis and the SR1 ORF regions of other bacilli resulting in Fig. 8. This work was supported by Grant Br1552/6-1 from the Deutsche Forschungsgemeinschaft to S. Brantl. A. Licht is financed by a fellowship from the Fonds der Chemischen Industrie.

**References**


Staphylococcus aureus RNAIII and the endoribonuclease III coordinately regulate spa gene expression. 
*EMBO J* 24: 824–835.

DBTS: a database of *Bacillus subtilis* promoters and transcription factors. 

Ivanova, N., Sorokin, A., Anderson, I., Gallner, N., Candelon, B., Kapatrul, V., et al. (2003) Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*. 


*Proc Natl Acad Sci USA* 97: 9919–9924.


*Proc Natl Acad Sci USA* 95: 12462–12467.

*Proc Natl Acad Sci USA* 99: 4620–4625.


*EMBO J* 14: 4569–4577.


*Proc Natl Acad Sci USA* 74: 5463–5467.


Stülke, J., Martin-Verstraete, I., Zagorec, M., Rose, M., Klier, A., and Rapoport, G. (1997) Induction of the *Bacillus subtilis* ptsGHI operon by glucose is controlled by a novel antiterminator, GlcT. 


*Genes Dev* 15: 1637–1651.


**Supplementary material**

The following supplementary material is available for this article online:

**Table S1.** Oligodeoxyribonucleotides used for the construction of plasmids, for the purification of native CcpN, for riboprobes and for footprinting studies.

**Table S2.** Oligodeoxyribonucleotides used in EMSAs to derive a provisional consensus sequence of the sugar box.