Keeping signals straight in transcription regulation: specificity determinants for the interaction of a family of conserved bacterial RNA–protein couples

Oliver Schilling, Christina Herzberg, Tina Hertrich, Hanna Vörmann, Dirk Jessen, Sebastian Hübner, Fritz Titgemeyer and Jörg Stülke*

Department of General Microbiology, Institute of Microbiology and Genetics, Georg-August University Göttingen, Germany and Lehrstuhl für Mikrobiologie, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Received August 3, 2006; Revised September 20, 2006; Accepted September 21, 2006

ABSTRACT
Regulatory systems often evolve by duplication of ancestral systems and subsequent specialization of the components of the novel signal transduction systems. In the Gram-positive soil bacterium Bacillus subtilis, four homologous antitermination systems control the expression of genes involved in the metabolism of glucose, sucrose and β-glucosides. Each of these systems is made up of a sensory sugar permease that also acts as phosphotransferase, an antitermination protein, and a RNA switch that is composed of two mutually exclusive structures, a RNA antiterminator (RAT) and a transcriptional terminator. We have studied the contributions of sugar specificity of the permeases, carbon catabolite repression, and protein–RAT recognition for the straightness of the signalling chains. We found that the β-glucoside permease BglP does also have a minor activity in glucose transport. However, this activity is irrelevant under physiological conditions since carbon catabolite repression in the presence of glucose prevents the synthesis of the β-glucoside permease. Reporter gene studies, in vitro RNA–protein interaction analyzes and northern blot transcript analyzes revealed that the interactions between the antiterminator proteins and their RNA targets are the major factor contributing to regulatory specificity. Both structural features in the RATs and individual bases are important specificity determinants. Our study revealed that the specificity of protein–RNA interactions, substrate specificity of the permeases as well as the general mechanism of carbon catabolite repression together allow to keep the signalling chains straight and to avoid excessive cross-talk between the systems.

INTRODUCTION
To sense their environment and to adapt to changing conditions, all organisms possess signal transduction systems which are composed of a sensor that perceives the signal, a regulator that can modify its activity in response to the signal, and a target of regulation. This general scheme can be modified in many ways: The sensor and the regulator are often combined in the same molecule as in the Lac repressor. The sensor and the regulator are usually proteins, but regulatory RNAs continue to be uncovered. The target of the regulation may be a protein, i.e. an enzyme, but for the control of gene expression, specific DNA or RNA sequences are the most common targets.

In bacteria, the number of environmental or internal signals that need to be sensed is much higher than the number of non-related regulatory systems. Thus, large families of regulation systems are present in bacteria. Among the most common families are the two-component regulatory systems, sigma factors with their anti-sigma factors as well as several families of repressor and activator proteins (1–5). All these families can be divided to sub-families that do often respond to similar signals. The evolution of signalling families is still in progress and can be observed in the transcriptional regulation of biodegradation pathways. Even more, new regulatory systems can be generated artificially (6,7). The similarity of the components of many families of signal transduction systems raises the question how the bacteria avoid excessive cross-talk, i.e. the activation of a regulatory protein by gratuitous inducers or the induction of a gene by a non-cognate regulator protein that recognizes a similar DNA sequence. This problem was the subject of extensive analyzes for the
two-component regulatory systems in the gram-positive soil bacterium *Bacillus subtilis* (8).

We are interested in the control of glucose utilization in *B. subtilis*. This sugar is transported by a specific permease of the phosphotransferase system (PTS) encoded by *ptsG* and is subsequently catabolized via the glycolytic pathway (9). The expression of the *ptsG* gene and of several glycolytic genes is inducible by glucose, however, the mechanisms differ. While *ptsG* expression is induced by transcriptional antitermination, the glycolytic gapA operon is controlled by the repressor CggR (10–13). Induction of *ptsG* expression involves a RNA switch which is the target of the antitermination protein GlcT, and the sensory glucose permease, PtsG. As part of the PTS, the glucose permease possesses two soluble domains that are involved in the phosphate transfer from phosphoenolpyruvate to the incoming sugar, the domains IIA and IIB (14). If glucose is present, the phosphate groups are immediately transferred to the sugar, whereas they accumulate on the glucose permease as well as on the two general proteins of the PTS, enzyme I and HPr, in the absence of glucose. Under these conditions, the glucose permease can transfer a phosphate residue to GlcT thereby inactivating the antitermination protein (15,16). GlcT is made up of three domains, an N-terminal RNA-binding domain, and two homologous PTS-regulation domains called PRD-I and PRD-II (15,17,18). Phosphorylation of a conserved histidine residue in PRD-I by the glucose permease results in GlcT inactivation in the absence of glucose. Biochemical studies revealed that PRD-II of GlcT can also be phosphorylated on a conserved histidine residue, however this phosphorylation is catalyzed by the HPr protein of the PTS and has only a very minor impact on the activity of GlcT (16). If in the right phosphorylation state, i.e. if non-phosphorylated in PRD-I, GlcT can bind its target site on the *ptsG* mRNA called RNA antiterminator (RAT, 15,19). The RAT overlaps a transcrip- tional terminator located in the leader region of the *ptsG* mRNA and the two structures form a RNA switch since they are mutually exclusive. Binding of GlcT to the RAT is thought to prevent the formation of the terminator and to allow transcription elongation into the *ptsG* structural gene. This regulatory system couples the availability of the inducer glucose to the phosphorylation state of the sensor permease and the antitermination protein GlcT resulting in either of two states of the *ptsG* RNA switch and subsequently in *ptsG* gene expression.

The regulatory system controlling *ptsG* expression is part of a family made up of highly conserved components, i.e. sensor permeases, antitermination proteins and RAT targets for the regulatory proteins (see Figure 1A). The additional permeases transport sucrose and the β-glucoside salicin. Two antitermination proteins, SacT and SacY, regulate expression of sucrose catabolic genes. While SacT is thought to be active at low sucrose concentrations, high concentrations of sucrose are required to activate SacY since its cognate permease SacX has a very weak transport activity (9). LicT controls the expression of the *bglIP* operon and the *licS* gene in the presence of salicin (20). SacT and LicT are only active if (i) their inducers are present and if (ii) no glucose is present in the medium. This allows their phosphorylation by HPr in the PRD-II leading to activation of the antitermination proteins. In contrast, SacY and GlcT, which are active in the presence of high sugar concentrations (or with the preferred sugar), are independent on a HPr-dependent activation even though HPr can phosphorylate these proteins. The glucose-dependent control of the antitermination proteins’ activity by HPr is part of the phenomenon of carbon catabolite repression which results in the preferential utilization of easily metabolizable carbon sources (18,21,22).

The RAT targets of the four antitermination proteins are all similar to each other (Figure 1B). Some determinants causing specificity of protein–RNA interaction have been identified in a pioneering work (19). The determination of the structure of the complex between the RNA-binding domain of LicT and its cognate RAT-RNA suggested that the essential contacts between the protein and the RNA are made in the loop regions of the RAT (23). The *ptsG* RAT recognized by GlcT is most different from all other RAT structures, and neither this RAT nor GlcT are involved in any cross-talk (24). In this work, we identified determinants that result in the regulatory specificity of the four distinct antitermination systems.

| Sugar       | Permease | Antitermination | Target gene (RAT) |
|-------------|----------|-----------------|-------------------|
| Glucose     | PtsG     | GlcT            | ptsG              |
| Salicin     | BglP     | LicT            | bglIP, licS       |
| Sucrose     | SacP     | SacT            | sacPA             |
| Sucrose     | SacX     | SacY            | sacXY, sacB       |

Figure 1. A family of antitermination systems controls sugar uptake and metabolism in *B. subtilis*. (A) The design of the four signalling systems. The target genes encode the following proteins: *ptsG*, glucose permease of the PTS; *bglP*, β-glucoside permease of the PTS; *bgIH*, phospho-β-glucoside hydrolase; *licS*, β-1,3-1,4-glucanase; *sacP*, sucrose permease of the PTS; *sacA*, sacrose; *sacX*, sucrose permease of the PTS (low affinity); *sacB*, extracellular levansucrase. (B) Comparison of the secondary structures of the related RAT-RNAs of *B. subtilis* (24,43).
MATERIALS AND METHODS

Bacterial strains and growth conditions

The *B. subtilis* strains used in this study are shown in Table 1. All *B. subtilis* strains are derivatives of the wild type strain MG168. Strains used in the cause of site-directed mutagenesis studies are listed in Tables 2 and 3. These strains were all derived by transformation from the basal mutant strains listed in Table 1. *Escherichia coli* DH5α and BL21(DE3) (25) were used for cloning experiments and for expression of recombinant proteins, respectively.

*B. subtilis* was grown in SP medium or in CSE minimal medium (26). The media were supplemented with auxotrophic requirements (at 50 mg/l), carbon sources and inducers as indicated. *E. coli* was grown in Luria–Bertani medium (LB medium) and transformants were selected on plates containing ampicillin (100 μg/ml). LB and SP plates were prepared by the addition of 17 g Bacto agar/l (Difco) to LB or SP medium, respectively.

Transformation and characterization of the phenotype

*B. subtilis* was transformed with plasmid DNA according to the two-step protocol described previously (27). Transformants were selected on SP plates containing kanamycin (Km 5 μg/ml), chloramphenicol (Cm 5 μg/ml), spectinomycin (Spc 100 μg/ml), or erythromycin plus lincomycin (Em 1 μg/ml and Lin 10 μg/ml).

In *B. subtilis*, amylase activity was detected after growth on SP medium supplemented with 10 g hydrolyzed starch/l (Connaught). Starch degradation was detected by sublimating iodine onto the plates.

Quantitative studies of *lacZ* expression in *B. subtilis* in liquid medium were performed as follows: cells were grown in CSE medium supplemented with the carbon source (Km 5 μg/ml). After growth to mid-log phase, samples were taken and assayed for β-galactosidase and amylase activities.

Table 2. Effect of mutations in the *ptsG* RAT on recognition by the different antiterminator proteins

| Strain   | RAT   | Relevant genotype | β-Galactosidase activity (U/mg protein) | CSE Glc | CSE Suc (0.1%) | CSE Suc (2%) | CSE Sal |
|----------|-------|-------------------|----------------------------------------|---------|----------------|---------------|---------|
| QB5448   | *ptsG*| wild type         | 9                                      | 548     | 231            | 329           | 267     |
| GP109    | *ptsG*| *gltC*            | 4                                      | 5       | 2              | 8             | 9       |
| GP387    | *ptsG*| ΔSacT ΔSacY       | 10                                     | 327     | 85             | 166           | 335     |
| GP389    | *ptsG*| ΔlicT             | 12                                     | 412     | 222            | 422           | 312     |
| GP385    | *ptsG*-R1| wild type       | 6                                      | 8       | 34             | 37            | 102     |
| GP386    | *ptsG*-R1| *gltC*           | 10                                     | 18      | 48             | 49            | 121     |
| GP390    | *ptsG*-R1| ΔlicT            | 16                                     | 11      | 17             | 14            | 13      |
| GP413    | *ptsG*-R2| wild type        | 3                                      | 2       | 2              | 3             | 4       |
| GP415    | *ptsG*-R3| wild type        | 14                                     | 11      | 12             | 7             | 14      |
| GP416    | *ptsG*-R4| wild type        | 11                                     | 3       | 8              | 6             | 94      |
| GP396    | *ptsG*| *gltC*            | 7                                      | 9       | 34             | 20            | 90      |
| GP417    | *ptsG*| ΔlicT             | 10                                     | 10      | 11             | 10            | 9       |
| GP404    | *ptsG*| wild type         | 34                                     | 33      | 174            | 121           | 877     |
| GP400    | *ptsG*| *gltC*            | 24                                     | 100     | 457            | 178           | 633     |
| GP402    | *ptsG*| ΔlicT             | 12                                     | 18      | 66             | 80            | 16      |
| GP455    | *ptsG*-R5| ΔlicT ΔSacY      | 7                                      | 9       | 42             | 40            | 8       |
| GP454    | *ptsG*-R5| ΔlicT ΔSacT      | 6                                      | 10      | 10             | 9             | 9       |
| GP456    | *ptsG*-R5| ΔlicT ΔSacY ΔSacT | 10                                     | 9       | 10             | 11            | 8       |
| GP434    | *ptsG*| *gltC*            | 9                                      | 8       | —              | —             | —       |
| GP436    | *ptsG*| ΔlicT             | 32                                     | 70      | —              | —             | —       |
| GP435    | *ptsG*| ΔlicT ΔSacY       | 33                                     | 73      | —              | —             | —       |
| GP408    | *ptsG*-R6| wild type        | 34                                     | 16      | 45             | 50            | 702     |
| GP399    | *ptsG*-R6| *gltC*           | 19                                     | 50      | 239            | 156           | 514     |
| GP409    | *ptsG*-R6| ΔlicT            | 14                                     | 16      | 20             | 20            | 12      |
| GP464    | *ptsG*-R7| wild type        | 37                                     | 32      | 36             | 37            | 60      |
| GP419    | *ptsG*-R8| wild type        | 147                                    | 188     | 157            | 238           | 1038    |
| GP420    | *ptsG*-R8| *gltC*           | 160                                    | 174     | 660            | 798           | 786     |
| GP421    | *ptsG*-R8| ΔlicT            | 209                                    | 150     | 300            | 287           | 217     |

*Arrows indicate construction by transformation.*
Table 3. Conversion analysis of the sacB RAT

| Strain   | RAT | Relevant genotype | β-Galactosidase activity (U/mg protein)* | CSE | CSE-Glp | CSE Suc (0.1%) | CSE Suc (2%) | CSE Sal |
|----------|-----|-------------------|----------------------------------------|-----|---------|----------------|-------------|--------|
| GP437    | sacB| wild type         | 9                                      | 6   | 54      | 78             | 5           |        |
| GP440    | sacB| ΔsacT             | 5                                      | 4   | 5       | 28             | 4           |        |
| GP438    | sacB| ΔsacY             | 6                                      | 5   | 50      | 49             | 6           |        |
| GP441    | sacB| ΔsacT ΔsacY       | 6                                      | 3   | 5       | 4              | 3           |        |
| GP461    | sacB| ΔsacT ΔlicT       | 7                                      | 8   | 132     | 104            | 35          |        |
| GP465    | sacB| ΔsacY ΔlicT       | 5                                      | 4   | 3       | 4              | 2           |        |
| GP466    | sacB| ΔsacT ΔlicT       | 5                                      | 5   | 80      | 74             | 3           |        |
| GP463    | sacB| ΔsacT ΔlicT       | 11                                     | 8   | 10      | 12             | 28          |        |
| GP462    | sacB| ΔlicT             | 7                                      | 9   | 85      | 82             | 5           |        |
| GP460    | sacB| wild type         | 7                                      | 6   | 93      | 145            | 8           |        |
| GP472    | sacB| ΔsacY ΔlicT       | 3                                      | 5   | 97      | 112            | 5           |        |
| GP471    | sacB| ΔsacT ΔlicT       | 4                                      | 3   | 4       | 7              | 3           |        |
| GP519    | sacB| wild type         | 5                                      | 2   | 96      | 88             | 4           |        |
| GP521    | sacB| ΔlicT             | 4                                      |      | 4       |                | 9           |        |
| GP540    | sacB| wild type         | 8                                      | 8   | 70      | 117            | 7           |        |
| GP541    | sacB| ΔsacT             | 7                                      | 8   | 9       | 35             | 7           |        |
| GP542    | sacB| ΔsacT ΔsacY       | 4                                      | 3   | 3       | 4              | 2           |        |
| GP520    | sacB| ΔlicT             | 4                                      | 5   | 216     | 208            | 6           |        |
| GP522    | sacB| ΔlicT             | 4                                      | 2   | 5       | 7              | 2           |        |
| GP537    | sacB| wild type         | 4                                      | 6   | 349     | 463            | 11          |        |
| GP538    | sacB| ΔsacT             | 6                                      | 5   | 6       | 22             | 9           |        |
| GP539    | sacB| ΔsacT ΔsacY       | 3                                      | 4   | 4       | 4              | 3           |        |
| GP476    | sacB| wild type         | 3                                      | 6   | 176     | 145            | 92          |        |
| GP544    | sacB| ΔsacT             | 5                                      | 4   | 7       | 8              | 136         |        |
| GP484    | sacB| ΔlicT             | 3                                      | 2   | 195     | 185            | 2           |        |
| GP536    | sacB| ΔsacT ΔlicT       | 2                                      | 2   | 3       | 5              | 2           |        |
| GP477    | sacB| wild type         | 6                                      | 7   | 248     | 294            | 218         |        |
| GP486    | sacB| ΔlicT             | 2                                      | 1   | 305     | 188            | 2           |        |
| GP487    | sacB| ΔsacT             | 5                                      | 4   | 9       | 8              | 376         |        |
| GP480    | sacB| wild type         | 7                                      | 11  | 280     | 253            | 510         |        |
| GP492    | sacB| ΔlicT             | 3                                      | 4   | 291     | 189            | 3           |        |
| GP493    | sacB| ΔlicT ΔsacT       | 4                                      | 2   | 3       | 4              | 2           |        |
| GP494    | sacB| ΔlicT ΔsacY       | 3                                      | 3   | 213     | 193            | 3           |        |
| GP444    | sacB| wild type         | 6                                      | 3   | 24      | 18             | 10          |        |
| GP543    | sacB| ΔsacT             | 5                                      | -   | 5       | 8              | 12          |        |
| GP446    | sacB| wild type         | 48                                     | 32  | 43      | 34             | 23          |        |
| GP448    | sacB| wild type         | 5                                      | 4   | 4       | 6              | 4           |        |
| GP450    | sacB| wild type         | 28                                     | 95  | 112     | 129            | 33          |        |
| GP453    | sacB| ΔlicT ΔsacT ΔsacY | 48                                     | 86  | 104     | 80             | 28          |        |
| GP451    | sacB| glcT              | 31                                     | 33  | 20      | 36             | 18          |        |

*Representative values of lacZ expression. All measurements were performed at least twice.

Sources indicated. Cells were harvested at OD_{600} 0.6–0.8. Cell extracts were obtained by treatment with lysozyme and DNAse. β-Galactosidase activities were determined as previously described using o-nitrophenyl-galactoside as a substrate (27). One unit is defined as the amount of enzyme which produces 1 nmol of o-nitrophenol per min at 28°C.

DNA manipulation

Transformation of E.coli and plasmid DNA extraction were performed using standard procedures (25). Restriction enzymes, T4 DNA ligase and DNA polymerases were used as recommended by the manufacturers. DNA fragments were purified from agarose gels using the QIAquick gel extraction kit (Qiagen®, Hilden, Germany). Pfu DNA polymerase was used for the PCR as recommended by the manufacturer. The combined chain reaction and the multiple mutation reaction were performed with Pfu DNA polymerase and thermostable DNA ligase (Ampligase®, Epicentre, Wisconsin, USA). DNA sequences were determined using the dideoxy chain termination method (25). Chromosomal DNA of B.subtilis was isolated as described (27).

Construction of sacT and sacY mutant strains by allelic replacement

To construct sacT and sacY mutant strains, the long flanking homology PCR (LFH-PCR) technique was used (28). Briefly, cassettes carrying the cat and spe resistance genes were amplified from the plasmids pGEM-cat and pDG1726, respectively (29,30). DNA fragments of ~1000 bp flanking the regions to be deleted at their 5' and 3' ends were amplified. The 3' end of the upstream fragment as well as the 5' end of the downstream fragment extended into the gene(s) to be deleted in a way that all expression signals of genes up- and downstream of the targeted genes remained intact. The joining of the two fragments to the resistance cassette was performed in a second PCR as described previously (31). In these reaction we used the primer pairs cat-fwd (5'-CCGGCAATAGTTACC-3') and cat-rev (5'-CCAGCGTGGACCAGGCGTGGCAGTTACAC-3') and spec-fwd/spec-rev (31) for the
amplification and joining of the cat and spc cassettes, respectively. The PCR products were directly used to transform B. subtilis. The integrity of the regions flanking the integrated resistance cassettes was verified by sequencing PCR products of ~1000 bp amplified from chromosomal DNA of the resulting mutants. The resulting strains were GP425 (ΔsacY::cat) and GP429 (ΔsacT::spc).

Site-directed mutagenesis

Translational fusions of variants of the ptsG and sacB regulatory regions with the lacZ gene were constructed using the vector pAC7 (32) containing the kanamycin resistance gene aphA3. The plasmid harbours a lacZ gene without a promoter located between two fragments of the B. subtilis amyE gene. To construct a translational sacB-lacZ fusion, the DNA upstream from the sacB gene [−464 to +15 nt relative to the translational start point of sacB (33)] was amplified by PCR using the primers OS49 (5′-AAAAGATTCGATCCGTTTTAACCCTACCATATAAC-3′) and OS50 (5′-TTTGGATCCCTTTTATGATTCATGTC-3′). The primers introduced BamHI and EcoRI cloning sites at the ends of the amplified fragment and created an in-frame translational fusion of the lacZ gene with the 5th codon of sacB. The PCR product was inserted into pAC7, both linearized with the same enzymes producing plasmid pGP437.

To study the effect of point mutations in the RAT sequences the following strategy was applied: a DNA fragment carrying the mutant form of the RAT was constructed by site-directed mutagenesis using either the combined chain reaction or the multiple mutation reaction (to introduce three or more mutations simultaneously) as outlined previously (34,35). Plasmids pGP66 (10) and pGP437 containing the ptsG and sacB promoter regions, respectively, served as templates. The mutagenic primers and the resulting plasmids are available upon request. The oligonucleotides JS11 (10)/IL5 (36) and OS49/OS50 (see above) were used as outer primers for ptsG and sacB, respectively. The final PCR products were purified and cut by BamHI and MfeI (for ptsG) or BamHI and EcoRI (for sacB) sites introduced by the PCR primers. To introduce the constructed lacZ fusions into the chromosome of B. subtilis, competent cells of the wild type strain 168 were transformed with the plasmids carrying the respective mutations linearized with Scal.

Construction of expression vectors for the RNA-binding domains of antiterminator proteins

A plasmid allowing the fusion of any protein to a Strep tag at the C-terminus was constructed as follows: First, the expression vector pET3C (Novagen) was digested with NdeI and BamHI. The insert containing a small multiple cloning site and the Strep tag was prepared by annealing the complementary oligonucleotides OS91 (5′-TATGAGGCTCGAGATCCTGGAACCCGGATTTTCTGCGAAATACTGATT-3′) and OS92 (5′-GATCCTACTATTTCGAACTGCGGGTTGCTCCAGGATGGATCCTCA-3′). The resulting DNA fragment carries ends compatible with NdeI and BamHI. Upon ligation, the NdeI site was conserved whereas the BamHI site was lost. The resulting plasmid, pGP574, carries an IPTG-inducible promoter, a small cloning site (NdeI–SacI–BamHI) for the insertion of the coding sequences, and the sequence encoding the Strep tag followed by two stop codons.

To fuse the RNA-binding domains of GlcT, LicT and SacT to a Strep tag at their C-termini, plasmids pGP575, pGP576 and pGP577 were constructed: DNA fragments corresponding to amino acids 1–60 of GlcT, and 1–57 of LicT and SacT were amplified by PCR using chromosomal DNA of B. subtilis QBS448 and the primer pairs OS93/OS94, OS95/OS96, and OS97/OS98, respectively (the primer sequences are available upon request). The PCR products were digested with NdeI and BamHI, and the resulting fragments were cloned into the expression vector pGP574 cut with the same enzymes.

Protein purification

E. coli BL21(DE3)/pLysS was used as host for the over-expression of recombinant proteins. Expression was induced by the addition of IPTG (final concentration 1 mM) to exponentially growing cultures (OD600 of 0.8). Cells were lysed using a French press. After lysis the crude extracts were centrifuged at 15 000 × g for 30 min and then passed over a Streptactin column (IBA, Gottingen, Germany). The recombinant protein was eluted with desthiobiotin (Sigma, final concentration 2.5 mM). After elution the fractions were tested for the desired protein using 12.5% SDS–PAGE gels. The relevant fractions were combined and dialysed overnight. Purified proteins were concentrated using Microsep™ Microconcentrators with a molecular weight cut-off of 3 kDa (Pall Filtron, Northborough, MA). The protein concentration was determined according to the method of Bradford using the Bio-rad dye-binding assay and BSA as the standard.

Assay of interaction between the RNA-binding domains and RAT-RNA

To obtain templates for the in vitro synthesis of the ptsG RAT-RNA, the primers OS25/OS26 (24) were used to amplify a 99 bp PCR product using pGP66 or the plasmid carrying the desired mutation as template. Similarly, a 99 bp DNA fragment encompassing the sacB RATS was amplified using the oligonucleotides OS86 (5′-CCAAGTTAATACGACTCAGTTAGCGAAAAGTAAATCGCGCG-3′) and OS87 (5′-GTATACACTTTGCCCTTCTAC-3′) and pGP437 or a mutant variant as template. The presence of a T7 RNA polymerase recognition site on primers OS25 and OS86 (underlined) allowed the use of the PCR product as a template for in vitro transcription with T7 RNA polymerase (Roche Diagnostics). The integrity of the RNA transcripts was analyzed by denaturing agarose gel electrophoresis (12).

Binding of the RNA-binding domains to RAT-RNA was analyzed by gel retardation experiments. The RAT-RNA (in water) was denatured by incubation at 90°C for 2 min and renatured by dilution 1:1 with ice cold water and subsequent incubation on ice. Purified protein was added to the RAT-RNA and the samples were incubated for 10 min at room temperature in TAE buffer in the presence of 300 mM NaCl. After this incubation, glycerol was added to a final concentration of 10% (v/v). The samples were then analyzed on 10% Tris–acetate PAA gels.
Northern blot analysis
RNA was prepared by the modified ‘mechanical disruption protocol’ described previously (12). Briefly, 20 ml of cells were harvested at the exponential phase. After mechanical cell disruption, the frozen powder was instantly resuspended in 3 ml lysis buffer [4 M guanidine isothiocyanate; 0.025 M sodium acetate, pH 5.3; 0.5% N-laurylsarcosine (w/v)]. Subsequently, total RNA was extracted using the RNeasy Mini kit (Quiagen, Germany). Digoxigenin RNA probes specific for the E.coli lacZ gene were obtained by in vitro transcription with T7 RNA polymerase (Roche Diagnostics) using a PCR-generated fragment as templates. The primers used for PCR were SHU55 (5’-GTGTGAACTCGTGCAGCT) and SHU56 (5’-CTAATACGACTCATATAGGAGTTGCAGTTCAACCACC). The reverse primers contained a T7 RNA polymerase recognition sequence. In vitro RNA labelling, hybridization and signal detection were carried out according to the manufacturer’s instructions according to the instructions of the manufacturer (DIG RNA labelling kit and detection chemicals; Roche Diagnostics).

Uptake of radioactive glucose in vivo
B.subtilis strains were grown in CSE medium with glucose (10 g/l). Sugar uptake assays were performed as described previously (10). Exponentially growing cells were harvested at an OD600 of 0.6–0.8 and washed once with the incorporation medium. Labelled [14C] glucose (184 mCi mmol−1) and non-labelled glucose (final concentration 0.4 mM) were added. Samples were taken and treated as described (10).

RESULTS
Analysis of the loop structures in the ptsG RAT
The ptsG RAT differs from all other RAT sequences recognized by antiterminator proteins of the BglG/SacY family in the structure of the lower loop (Figure 1B). In a previous work, we have demonstrated that the insertion of one base into the lower loop of the ptsG RAT (the ptsG-R1 mutation, see Figure 2) makes its structure similar to that recognized by the other antiterminator proteins and results in exclusive binding of LicT to this structure whereas it is not bound by GlcT (24). From this result it was concluded that structure rather than the nucleotide sequence is important for antiterminator protein-RAT recognition.

Since the antitermination proteins bind as dimers to the RAT, and LicT contacts different structures of the lower and the upper loop, we asked whether a RAT with an inversion of the lower and upper RAT structures might be recognized by any of the antitermination proteins (see Figure 2). The activity of this mutant RAT, ptsG-R2, was assayed by analyzing the expression of a translational fusion of the mutated ptsG control region to a promoterless lacZ gene (see Table 2). While the presence of glucose in the growth medium resulted in a strong GlcT-dependent induction of the ptsG promoter in the wild type, salicin induced the ptsG-R1 promoter region in a LicT-dependent manner. In contrast, the ptsG-R2 promoter region did not allow expression of the lacZ fusion irrespective of the potential inducing carbohydrate present in the medium. Thus, this RAT is not bound by any of the antitermination proteins in B.subtilis (Table 2).

The ptsG-R1 structure was obtained by inserting an A after position 3 of the RAT sequence. This did not only create a lower loop structure similar to those present in RAT structures bound by LicT, SacT, and SacY, but also generate an additional base pair between the lower and upper loops (see Figures 1 and 2). To rule out any effect of this extra base pair we constructed the ptsG-R3 RAT mutant by deleting the U at position 25. This results in a lower loop identical to that in the ptsG-R1 RAT, but separated from the upper loop by only 2 bp (Figure 2). The biological activity of this RAT mutant was determined by studying its effect on the expression of a ptsG-lacZ fusion. As shown in Table 2, the ptsG-R3 RAT did not confer induction under any of the conditions tested.
suggested that this structure is not recognized by any of the antitermination proteins (see Discussion).

The inability of GlcT to bind to the ptsG-R2 and ptsG-R3 RATs was verified by a northern blot analysis. The amounts of lacZ mRNA were compared in the wild type strain QB5448 and the two mutant strains GP413 and GP415. As can be seen in Figure 3, the lacZ mRNA was strongly induced in cells grown in the presence of glucose whereas no induction was observed in the two mutant strains. This result is in perfect agreement with those obtained by the reporter gene assays (Table 2).

Contribution of individual bases to the recognition of the RAT sequence by antitermination proteins

The cognate RATs bound by LicT, SacT, and SacY are very similar to each other both in terms of structure and sequence (Figure 1B). However, the ptsG-R1 RAT is recognized by LicT only and not by SacY or SacT. Therefore, we decided to introduce further mutations into the ptsG-R1 RAT that allow the evaluation of the contribution of individual nucleotides to protein–RNA recognition.

The RATs recognized by LicT and SacT contain an A at position 3 in the lower loop rather than a G as in the ptsG-R1 RAT. Therefore, we exchanged the G3 for an A. The effect of this mutation, present in the ptsG-R4 RAT (see Figure 4), was tested by the analysis of a ptsG-R4-lacZ fusion. This mutation resulted in a lacZ expression comparable to that observed with the ptsG-R1 RAT. As determined for ptsG-R1, the expression driven by the ptsG-R4 promoter region was completely dependent on a functional licT gene (Table 2). Thus, LicT is the only antiterminator protein binding to both the ptsG-R1 and ptsG-R4 RATs.

Another important difference between the ptsG-R1 RAT and all other RATs recognized by LicT, SacT, or SacY is the U:A base pair above the lower loop, which is A:U in ptsG-R1 (see Figures 1B and 2). Previous results suggested that inversions of base pairs in the stems of the RAT are tolerated as long as the general structure is conserved (24). However, due to the strict conservation of the U:A pair in this position in all RATs except ptsG-R1, we addressed the effect of such a base pair inversion. As can be seen in Table 2, this inversion, present in the ptsG-R5 mutation (Figure 4), resulted in an increased expression of the fusion under all conditions tested. However, the ptsG-R5 RAT conferred a strong induction in the presence of salicin, and this induction was dependent on the presence of the LicT antiterminator protein. In contrast, glucose did not induce this fusion suggesting that GlcT is unable to bind this RAT (Table 2). These observations were verified by an electrophoretic mobility shift analysis using the purified RNA-binding domains of GlcT and LicT (see Figure 5). As reported previously (24), GlcT efficiently bound the ptsG-R5 RAT-RNA. In contrast, LicT was unable to bind this RNA. In good agreement with the reporter gene analysis, LicT but not GlcT was capable of binding the ptsG-R5 RAT in vitro (Figure 5). For the ptsG-R1 RAT, induction in the presence of salicin and, to a lesser extent, sucrose, is strictly LicT-dependent (24, Table 2). We tested therefore whether the induction of the ptsG-R5 RAT by sucrose was also due to binding of LicT. As mentioned above, a deletion of the licT gene resulted in loss of ptsG-R5 induction by salicin. However, the licT mutation did not abolish the induction by sucrose at the ptsG-R5 RAT (Table 2) suggesting that either SacT or SacY (or both) bind this RNA and cause antitermination.

To test this possibility, we assayed the expression of the ptsG-R5-lacZ fusion in strains containing combinations of mutations of the three antiterminator genes. In a licT sacY double mutant, a slight reduction of sucrose induction of the ptsG-R5-lacZ fusion was observed as compared to the licT mutant strain (Table 2). In contrast, induction was completely lost in the licT sacY double and triple mutant strains carrying deletions of two and of three antiterminator protein-encoding genes, respectively. From this result we may conclude that SacT can recognize the ptsG-R5 RAT in addition to LicT. A northern blot analysis of the lacZ mRNA confirmed the strong induction by salicin and, to a lesser extent, by sucrose, conferred by the ptsG-R5 RAT (Figure 3).
demonstrate that the U:A base pair just above the lower loop is important to facilitate binding of LicT to this structure (compare the high β-galactosidase activity to that driven by the ptsG-R1 fusion, Table 2) and to allow binding of SacT.

The role of the base at position 3 in the lower loop was also analyzed in the context of the ptsG-R5 RAT. However, as observed with ptsG-R4, only minor effects of a substitution of G3 by A were observed (Table 2, see ptsG-R6, see Figure 4). Induction by salicin was slightly decreased, and the induction with sucrose was also completely dependent on LicT indicating that SacT did not bind the ptsG-R6 RAT. A substitution of G3 by U (ptsG-R7, as present in the sacB RAT recognized by SacY, Figure 4) resulted in loss of induction by sucrose and only weak induction upon the addition of salicin (15-fold reduction as compared to ptsG-R5, Table 2). Taken together, these results indicate that the G at position 3 facilitates binding of the antitermination proteins. In contrast, an U at this position strongly diminishes binding by LicT. These conclusions are validated by an analysis of the sacB RAT (see below).

Binding of the antitermination proteins to their RAT targets allows the formation of otherwise non-favoured RAT structures and prevents concomitantly the formation of the transcription terminators. The relative stability of the RAT structures may therefore be important for the level of gene expression. The RATs recognized by LicT, SacT, and SacY contain two G:C base pairs in the bottom stem whereas the ptsG RAT contains a A:U and a C:G base pair at this position. It seemed therefore possible that the replacement of the A:U base pair by a C:G base pair would result in a more stable RAT structure and thus affect transcription. To test this idea, the ptsG-R8 RAT was constructed based on ptsG-R5 and analyzed (see Figure 4, Table 2). While the ptsG-R5 RAT allowed only a weak basal expression in the absence of any inducer (CSE medium), a strongly increased basal expression was found for ptsG-R8 (34 versus 147 U of β-galactosidase). This was also reflected in a northern blot analysis of lacZ mRNA if expressed under the control of the ptsG-R5 and ptsG-R8 RAT (compare Figure 3, lanes 7 and 11). The ptsG-R8-lacZ fusion was also induced by salicin, and the induced expression was the sum of read through (∼150 U) and real induction (∼900 U, see ptsG-R5, Table 2). However, the increased read through might also result from a destabilization of the terminator even though an extra mutation was introduced in the terminator to restore base pairing.

Carbon catabolite repression interferes with the transport of glucose by BglP

The analysis of the ptsG-R5 RAT revealed that this structure is efficiently bound by LicT but not by GlcT. The disruption of the glcT gene in a strain carrying the ptsG-R5-lacZ fusion resulted in induction of β-galactosidase by salicin and sucrose (see Table 2) as expected due to the binding of LicT and SacT, respectively (see above). Surprisingly, glucose did also activate expression of this fusion in a glcT mutant strain. Since GlcT is not available in this mutant, LicT or SacT must be activated in the presence of glucose in the glcT mutant. To test this idea, we studied the activity of the ptsG-R5 control region in glcT licT or glcT sacT double mutants. As shown in Table 2, only a minor effect of the sacT deletion was observed, whereas the deletion of the licT gene resulted in complete loss of induction by glucose. Thus, glucose can activate LicT in a glcT mutant strain.

Two scenarios for the activation of LicT by glucose can be envisaged. First, there might be some non-specificity in BglP that results in the transport of glucose by this permease and the subsequent dephosphorylation and activation of the cognate antiterminator LicT. Second, there might be some cross-talk between the glucose permease PtsG and the LicT antiterminator that results in LicT activation upon glucose transport. Several lines of evidence demonstrate that the former possibility reflects the truth: (i) The glucose permease PtsG is not expressed in a glcT mutant strain and is therefore unable to activate LicT in a glcT mutant (10,15). (ii) BglP phosphorulates and thereby inactivates LicT in the absence of the substrate salicin, and this regulation would be dominant over any minor PtsG-dependent dephosphorylation of LicT (21,22). (iii) To provide direct evidence for glucose uptake by BglG we measured the glucose transport of glucose-grown cells of a wild type strain (B. subtilis 168), a ptsG mutant (QB5435) and a ptsG bglP double mutant (GP470). As shown in Figure 6, glucose was efficiently transported by the wild type strain (initial uptake rate 620 ± 110 pmol glucose per minute and OD600), whereas a significant reduction was observed in the ptsG mutant (initial uptake rate 62 ± 5 pmol glucose per minute and OD600). These results are in good agreements with previous studies of glucose transport in ptsG mutants (10,14). In the ptsG bglP double mutant GP470, the transport of glucose was further reduced (see Figure 6, initial uptake rate 24 ± 1 pmol glucose per minute and OD600), confirming that BglP has some minor glucose transport activity which may explain glucose-dependent activation of LicT in the glcT mutant background (see Discussion).

Conversion analysis of the sacB RAT sequence towards new recognition specificities

The similarity of the ptsG-R6 RAT to that of the sacPA operon (Figures 1 and 4) suggests that both RNA structures
might be recognized by the same proteins. However, as shown in Table 2, ptsG-R6 is bound exclusively by LicT whereas the sacPA RAT is the target of SacT and is not recognized by LicT (17). Similarly, the ptsG-R7 RAT which is poorly recognized by LicT but by none of the other antiterminator proteins resembles strongly the sacB RAT which is the target of SacY (see Figures 1 and 4). Thus, additional components seem to play a role in RAT-antiterminator protein recognition. To unravel these factors, we decided to perform an in-depth conversion analysis of the sacB RAT to mutate it and shift it gradually to sequences that are not longer recognized by SacB but rather by one of the three other family members. We chose the sacB RAT for this purpose since sacB lacks any additional regulation by carbon catabolite repression (37). This analysis was aimed at the identification of bases that are responsible for the specificity for one or the other antiterminator protein.

**Discrimination between SacY and SacT**

First, we determined the regulation mediated by the wild type sacB RAT. If the lacZ gene was expressed under the control of this RAT, induction was observed only in the presence of sucrose confirming that neither GlcT nor LicT bind the sacB RAT. Induction by sucrose occurred both at low and high sucrose concentrations which activate SacT and SacY, respectively. Indeed, induction at a low sucrose concentration was lost in the sacT mutant. In the sacY mutant strain, induction was still visible at both concentrations suggesting that SacT is active under both conditions. In the sacT sacy double mutant, the sacB RAT-terminator couple did not allow induction under all the conditions tested (Table 3).

The ptsG-R5 RAT which is recognized by LicT and SacT, closely resembles the sacB RAT but contains a G at position 3 rather than a U as in the sacB RAT. We constructed therefore the sacB-R1 RAT by replacing U3 by a G (see Figure 7). This single mutation resulted in a significant specificity shift. The sacB-R1 RAT was not longer a target for SacY, whereas the activation of SacT allowed a higher β-galactosidase expression as compared to the wild type sacB RAT (Table 3). Moreover, the sacB-R1 RAT allowed LicT-dependent induction by salicin. These results are in good general agreement with the observed affinity of the similar ptsG-R5 RAT for SacT and LicT, however, the preference for the two antitermination proteins was inverse. The sacPA RAT, which is the cognate target of SacT, contains also a purine base at position 3, i.e. an A (see Figure 1B). Therefore, the sacB-R2 containing an A at position 3 was constructed (Figure 7). The presence of this RAT conferred induction by sucrose but neither by salicin nor glucose (Table 3). Thus, this RAT is not bound by LicT. To distinguish whether it is recognized by SacY or SacT we analyzed the expression driven by the sacB-R2-lacZ fusion in licT sacT and licT sacY double mutant strains. As shown in Table 3, the sacB-R2 RAT is efficiently bound by SacT at both low and high sucrose concentrations whereas it is not recognized by SacY. The results obtained with the sacB-R1 and -R2 mutants suggest that the U at position 3 of the RAT is important for recognition by SacY. In contrast, SacT tolerates all three tested bases at this position.

The data presented above demonstrate that U3 is important for SacY binding in the context of the sacB RAT. However, since SacY is capable of recognizing the sacPA RAT (17), it seems to be able to accept bases different from U at position 3. To address this question we exchanged the three bases in the sacB RAT that are different from the sacPA RAT. A replacement of U8 in the middle loop by a G as in sacPA (sacB-R3, Figure 7) resulted in loss of binding by SacY whereas SacT bound this RAT as judged from loss of sucrose induction in the sacT mutant (see Table 3). Both the sacB and sacPA RATs contain a UAAA tetraloop at the top. This loop is flanked by A-G and G-G pairs in sacB and sacPA, respectively (see Figure 1B). Therefore, we constructed the sacB-R4 mutant RAT with a G-G pair at the bottom of the top-loop (A13G exchange, Figure 7). This mutation did not affect binding by SacY and SacT as compared to the wild type sacB RAT. Moreover, it did not confer induction by salicin indicating that it is no target for LicT (see Table 3). A combination of the two mutations of sacB-R2 and sacB-R3 (U3A and U8G) present in the sacB-R5 RAT (Figure 7) resulted in enhanced induction by sucrose which was exclusively dependent on SacT as determined using a sacT mutant strain (Table 2). Indeed, the sacB-R5 RAT was efficiently bound by the RNA-binding domain of SacT (see Table 3). In contrast, the wild type sacB RAT was only weakly bound by SacT. These observations are in very good agreement with the high SacT-dependent induction of gene expression mediated by sacB-R5 as compared to induction conferred by the wild type sacB RAT. An additional mutation of the base pair at the bottom of the top loop (A13G) made the resulting sacB-R6 RAT identical to that of sacPA, but in a sequence context of sacB (see Figures 1B and 7). As expected, this RAT is most efficiently recognized by SacT. In the sacT mutant, only a very weak induction by sucrose was observed which was lost in the sacT sacy double mutant strain (Table 3). Thus, the A3 and G8 do both discriminate against binding by SacY. However, as demonstrated using the sacB-R6 RAT, the G-G base pair at

---

**Figure 7.** Secondary structures of the wild type sacB, sacB-R5, and sacB-R6 mutant RATs. Bases that differ from sacB RAT are boxed in the sacB-R5 double- and in the sacB-R6 triple mutant. The sacB-R1 to R4 mutant RATs are indicated as single base exchanges into sacB RAT. The sacB-R6 triple mutant RAT is identical to the sacPA wild type RAT (see Figure 1).
the bottom of the top loop seems to weaken this discrimination and does thus allow weak binding by SacY.

**Discrimination between SacY and LicT**

The *sacB* RAT differs from the *licS* and *bglPH* RATs that are the cognate targets of LicT by two bases in the lower loop. Additionally, the *bglPH* RAT contains a C-G base pair at the bottom of the top loop and a GAAA tetraloop at the top (see Figure 1B). The first step in the conversion of the *sacB* RAT to a structure expected to be recognized by LicT was the *sacB*-R2 mutation (U3A) described above. This RAT was bound by SacT but not by LicT (see Table 3).

With the introduction of a second mutation in the lower loop (G26A) the resulting *sacB*-R7 RAT was identical to that of *licS* (see Figures 1B and 9). The determination of β-galactosidase regulation conferred by this RAT demonstrated induction not only by sucrose but also by salicin (see Table 3). Induction by salicin was completely lost in a LicT mutant whereas sucrose induction was lost in the *sacT* RAT that made its structure gradually more similar to this distinct structure rather than the details of the actual nucleotide sequence is important for recognition by GlcT. Thus, for LicT and SacT, position 3 seems to be most important to maintain the proper RAT structure, whereas the opposing A at position 26 is important for allowing efficient binding by LicT.

**Discrimination between SacY and GlcT**

The *ptsG* RAT is most different from all other RAT structures in *B. subtilis* due to the triple base pairing in the lower loop region (24, see above, Figure 1B). It has been proposed that this distinct structure rather than the details of the actual nucleotide sequence is important for recognition by GlcT. To verify this assumption we introduced mutations into the *sacB* RAT that made its structure gradually more similar to that of the *ptsG* RAT. In a first step, the U at position 4 was deleted (Figure 11). This mutation is a reversal of the base conversion from the *ptsG* to the *ptsG*-R1 RAT (insertion of one base at position 4, see Figure 2), but in the context of the *sacB* RAT. The resulting *sacB*-R10 RAT allowed a very weak SacT-dependent induction by sucrose. In contrast, this RAT was not at all recognized by GlcT as concluded from the absence of induction by glucose. Here, the U3 might form a base pair with either A24 or A26 thus forming a structure weakly recognized and sufficiently stabilized by SacT.

---

**Figure 8.** Electrophoretic mobility shift analysis of the interaction between the *sacB* and *sacB*-R5 RATs, and the RNA-binding domain of SacT. Lanes 1–4 and 5–8 contain 100 pmol of *sacB* and *sacB*-R5 RAT-RNAs, respectively. Increasing concentrations of SacT were added to the RNA in lanes 2–4 and 6–8 prior to electrophoresis. Aliquots of 75, 150 and 300 pmol SacT were used.

**Figure 9.** Secondary structures of the wild type *sacB*, *sacB*-R2, *sacB*-R7 and *sacB*-R8 RATs. Bases that differ from the *sacB* RAT are boxed. The mutations introduced into the *sacB* RAT convert it gradually to the *licS* (*sacB*-R7) and the *bglPH* RATs (*sacB*-R8).

**Figure 10.** Electrophoretic mobility shift analysis of the interaction between the wild type *sacB* and several mutant RAT-RNAs (see Figure 8) with the RNA-binding domain of LicT. In all lanes, 100 pmol of RNA were used. In the lanes labelled with ‘+’, 250 pmol of LicT were added prior to electrophoresis.
to allow antitermination. The sacB-R10 mutant RAT was then parent to two further variants. In the sacB-R11-RAT, the U3 was replaced by a G (Figure 11). In the ptsG RAT, a G at this position contacts the nucleotides at positions 24 (U) and 26 (A). However, this does not seem to be the case with the two adenines in sacB-R11-RAT. This RAT does not confer induction to the lacZ gene, however, the read through was somewhat increased even in the absence of any inducer (see Table 3). A replacement of A24 present in sacB-R10 by a U (as in the ptsG RAT at this position) resulted in complete loss of expression of the reporter gene (see Table 3, sacB-R12, Figure 11). By replacing the U3 of sacB-R12 by a G, we obtained a lower loop that is identical to that found in the ptsG RAT (sacB-R13, Figure 11). Indeed, the sacB-R13-lacZ fusion was induced by glucose and sucrose (see Table 3). Since salicin and sucrose are known to activate GlcT (36), we tested the expression of this fusion in a licT sacT sacY triple mutant as well as in a gltC mutant strain. As expected, the combined deletion of licT, sacT, and sacY did not affect the induction by any of the sugars whereas no induction was observed in the gltC mutant strain (see Table 3). Thus, this RAT is exclusively recognized by GlcT. To verify this observation we performed electrophoretic mobility shift assays in the presence of the RNA-binding domain of GlcT using the RATs of ptsG and sacB as controls as well as the sacB-R11- and sacB-R13-RATs (see Figure 12). As previously observed, GlcT is capable of binding its cognate ptsG RAT. In contrast, the sacB RAT was not recognized by GlcT. Similarly, the sacB-R11-RAT was not retarded. As expected from the transcription regulation conferred by GlcT and the sacB-R13 RAT, an RNA-fragment containing this RAT was bound by the RNA-binding domain of GlcT. These findings confirm the important role of G3, U24, and A26 for the formation of the structure in the lower loop and in GlcT binding.

**DISCUSSION**

Several distinct mechanisms contribute to the specificity of the four antitermination systems present in *B. subtilis*. These include, first, the sugar permeases and their interactions with their substrates and with the cognate antitermination proteins. Second, carbon catabolite repression limits the conditions under which certain systems are expressed and the antiterminator proteins active. Finally, the interaction between the antiterminator proteins and the RAT-RNAs makes a major contribution to regulatory specificity.

The sugar permeases of the PTS can transport phosphorylate only one substrate, or they can exhibit a relaxed specificity, i.e. they may transport more than one sugar. The glucose permease PtsG is known to transport sucrose and salicin in addition to glucose thus explaining the induction of ptsG expression by these sugars (36, see Table 2). A relaxed specificity has also been observed for the GlcB permease from *Staphylococcus carnosus* which is also capable of transporting salicin in addition to glucose (38). Similarly, the β-glucoside permease BglP is able to transport glucose, although with a low efficiency (see Figure 6). In wild type strains, the bglP gene is strongly repressed in the presence of glucose, thus, this relaxed specificity has no biological consequence. In contrast, the two sucrose permeases seem to be highly specific for sucrose, and the SacX permease is regarded as being inactive since it does not contribute to sucrose transport (39). All experiments with the different antitermination systems published so far did not provide any indication that a permease might interact with a non-cognate antiterminator protein. This might reflect the parallel evolution of the permeases and their targets, the PRD-I domains of the antitermination proteins (40). Indeed, the control of the antitermination proteins by the corresponding sugar-specific permeases works beyond the species barrier as shown for *B. subtilis* LicT in *E. coli* or *S. carnosus* GlcT in *B. subtilis* (20,41).

Bacteria use carbon sources in a hierarchical order, i.e. those that are most easily metabolized with a maximum yield of energy are preferred. In *B. subtilis*, glucose is the preferred carbon source, and the presence of glucose prevents the activity of many enzymes as well as the expression of genes and operons that are required for the utilization of alternative carbon sources. Among the genes studied here, only *ptsG* is induced by glucose (via antitermination) whereas *sacPA*, *bglPH* and *licS* are repressed. This repression is achieved by two independent mechanisms: First, the CcpA...
LicT recognizes structures that are highly similar to the targets of SacT and SacY. An inspection of the RAT structures reveals that the LicT targets are unique in having an A at position 26 (see Figure 2B). The importance of this position is underlined by our mutation analysis of the sacB RAT. The importance of A26 for recognition by LicT is also supported by the structure of the LicT–RAT complex. There are several contacts of LicT with A26 and the sugar phosphate backbone in its immediate neighbourhood (23). The differential role of guanine and adenine residues for recognition of nucleic acids by proteins is well established (43). Moreover, the data indicate that the A at position 26 is necessary but not sufficient for LicT binding. In addition, a purine base is required at the opposing position 3 of the RAT. This is in good agreement with a previous study (19). Interestingly, a G at position 3 (sacB-R9) allows much higher LicT-, but also SacT-dependent antitermination as compared to a similar RAT containing an A at this position (sacB-R7).

There are conflicting reports on the recognition of the sacB RAT by SacT (17,19). We observed that the sacB RAT is recognized by both SacT and SacY. Interestingly, the SacT-dependent induction of sacB is stronger than the induction mediated by the cognate antiterminator, SacY. Thus, SacT induces both the sacPA operon and the sacB gene encoding levansucrase. In contrast, SacY exerts only a very minor effect at the sacPA RAT (identical to sacB-R6). Since SacT is active at both high and low sucrose concentrations whereas SacY is active only in the presence of large amounts of sucrose, SacT may be regarded as the major antiterminator protein controlling sucrose utilization. The minor role of SacY is also illustrated by the weak affinity of this protein to the sacB RAT which is two orders of magnitude lower than the affinities observed for LicT and GlcT with their respective targets (44, I. Langbein and J. Stülke, unpublished data). A step-wise conversion of the sacB RAT into a sacPA RAT like structure revealed the following observations: Single base mutations (Figure 7, sacB-R2, sacB-R3, sacB-R4, Table 3) all enhanced the binding of SacT, whereas the double and triple mutations (Figure 7, sacB-R5, sacB-R6) had additive effects. Most single mutations and the double mutation prevented SacY binding. However, the triple mutation which did also affect the top loop neutralized the negative effect of the two other nucleotides and restored binding of SacY.

Taken together, our data indicate that SacT is the most promiscuous of the antitermination proteins whereas GlcT at the other end of the spectrum is strictly confined to its cognate ptsG RAT due to its specific structural demands. In the living cell, glucose plays a special role as the by far most preferred carbon source. Therefore it is advantageous for the bacteria to have a regulatory system for glucose utilization that avoids any risk of cross-talk. On the other hand, SacT induces both sucrose catabolic systems, but it does not mediate antitermination at the bgIP RAT in wild type bacteria (45). As shown here and in previous publications, SacT can bind bgIP-like RAT structures (17, Table 3). It is so far unknown why SacT does not induce the bgIPH operon in wild type bacteria. More factors such as the sequence context surrounding the RAT, the top loop, and the overall stability of the different RAT/terminator couples may provide additional levels for controlling the effective interaction with the antiterminator
proteins. A careful analysis of the data presented here indicates that this is indeed the case. More work will be required to study the contributions of these factors.

ACKNOWLEDGEMENTS

We are grateful to Matthias H. Schmalisch for helpful discussions, and Timo Hupfeld for the help with some experiments. Thorsten Mascher is acknowledged for the gift of primers for LFH-PCR. This work was supported by grants from the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie to J.S. Funding to pay the Open Access publication charges for this article was provided by the DFG.

Conflict of interest statement. None declared.

REFERENCES

1. Helmann,J.D. (1999) Anti-sigma factors. Curr. Opin. Microbiol., 2, 135–141.
2. Stock,A.M., Robinson,V.L. and Goudreau,P.N. (2000) Two-component signal transduction. Annu. Rev. Biochem., 69, 183–215.
3. Panè-Farré,J., Lewis,R.J. and Stülke,J. (2005) The RsbRST stress response in Bacillus subtilis: discovery of an antiterminator protein. J. Mol. Biol., 249, 335–336.
4. Hoffman,J.L. and Brennan,R.G. (2002) Prokaryotic transcription regulators: more than just the helix-turn-helix motif. Curr. Opin. Struct. Biol., 12, 98–106.
5. Ramos,J.L., Martinez-Bueno,M., Molina-Henares,A.J., Teran,W., Watanabe,K., Zhang,X., Gallegos,M.T., Brennan,R. and Tobe,R. (2005) The TetR family of transcriptional repressors. Microbiol. Mol. Biol. Rev., 69, 326–356.
6. Galvao,T.C. and de Lorenzo,V. (2006) Transcriptional regulators à la carte: engineering new effector specificities in bacterial regulatory proteins. Curr. Opin. Biotechnol., 17, 34–42.
7. Garmentia,J., Devos,D., Valencia,A. and de Lorenzo,V. (2001) A la carte transcriptional regulators: unlocking responses of the prokaryotic enhancer-binding protein XylIR to non-natural effectors. Mol. Microbiol., 42, 47–59.
8. Hoch,J.A. and Varughese,K.I. (2001) Keeping signals straight in phosphorylase signal transduction. J. Bacteriol., 183, 4941–4949.
9. Stülke,J. and Hillen,W. (2000) Regulation of carbon catabolism in Bacillus subtilis. Annu. Rev. Microbiol., 54, 849–880.
10. Stülke,J., Martin-Verstraete,J., Yang,Y., Strub,M.P., Kochoyan,M., Steinmetz,M. and Stülke,J. (1999) Specific interaction of the CcpA antiterminator with multiple mutations into the glpK gene of Mycoplasma pneumoniae. Mol. Microbiol., 34, 141–148.
11. Jordan,S., Junker,A., Helmann,J.D. and Mascher,T. (2006) Regulation of LysR-dependent gene expression in Bacillus subtilis: identification of inhibitor proteins, regulator binding sites, and target genes of a conserved cell envelope stress-sensing two-component system. J. Bacteriol., 177, 2403–2407.
12. Weirich,Y., Monda,T., Hupfeld,T. and Dubnau,D. (1991) Signal and properties of comQ, a new competence regulatory gene of Bacillus subtilis. J. Bacteriol., 173, 5685–5693.
13. Steinmetz,M., Le Coq,D., Aymerich,S. and Gay,P. (1985) The DNA sequence of the gene for the secreted Bacillus subtilis enzyme levanuconase and its genetic control sites. Mol. Gen. Genet., 200, 220–228.
14. Weihep,M. and Stambrook,P.J. (1998) Site-directed mutagenesis by combined chain reaction. Anal. Biochem., 256, 137–140.
15. Bacher,S. and Stülke,J. (1997) Characterization of the presumptive phosphorylation sites of the Bacillus subtilis glucose permease by site-directed mutagenesis: Implication in glucose transport and catabolite repression. FEMS Microbiol. Lett., 156, 233–238.
16. Schmalisch,M., Bacher,S. and Stülke, J. (1999) Control of the Bacillus subtilis antiterminator protein GlcT by phosphorylation: Elucidation of the phosphorylation chain leading to inactivation of GlcT. J. Biol. Chem., 278, 51108–51115.
antiterminator GlcT with its RNA target, RAT. J. Mol. Biol., 293, 795–805.

37. Steinmetz, M., Le Coq, D. and Aymerich, S. (1989) Induction of saccharolytic enzymes by sucrose in Bacillus subtilis: evidence for two partially interchangeable regulatory pathways. J. Bacteriol., 171, 1519–1523.

38. Christiansen, I. and Hengstenberg, W. (1999) Staphylococcal phosphoenolpyruvate-dependent phosphotransferase system—two highly similar glucose permeases in Staphylococcus carnosus with different glucoside specificity: protein engineering in vivo? Microbiology, 145, 2881–2889.

39. Kunst, F., Pascal, M., Lepesant, J.A., Walle, J. and Dedonder, R. (1974) Purification and some properties of an endocellular sucrase from a constitutive mutant of Bacillus subtilis 168. Eur. J. Biochem., 42, 611–620.

40. Greenberg, D.G., Stülke, J. and Saier, M.H. (2002) Domain analysis of transcriptional regulators bearing PTS regulatory domains. Res. Microbiol., 153, 519–526.

41. Knezevic, I., Bachem, S., Sickmann, A., Meyer, H.E., Stülke, J. and Hengstenberg, W. (2000) Regulation of the glucose-specific phosphotransferase system (PTS) of Staphylococcus carnosus by the antiterminator protein GlcT. Microbiology, 146, 2333–2342.

42. Krüger, S., Gertz, S. and Hecker, M. (1996) Transcriptional analysis of bglPH expression in Bacillus subtilis: evidence for two distinct pathways mediating carbon catabolite repression. J. Bacteriol., 178, 2637–2644.

43. Nobeli, I., Laskowski, R.A., Valdar, W.S.I. and Thornton, J.M. (2001) On the molecular discrimination between adenine and guanine by proteins. Nucleic Acids Res., 29, 4294–4309.

44. Declerck, N., Vincent, F., Hoh, F., Aymerich, S. and van Tilbeurgh, H. RNA recognition by transcriptional antiterminators of the BglG/SacY family: functional and structural comparison of the CAT domain from SacY and LicT. J. Mol. Biol., 294, 389–402.

45. Krüger, S. and Hecker, M. (1995) Regulation of the putative bglPH operon for aryl-β-glycoside utilization in Bacillus subtilis. J. Bacteriol., 177, 5590–5597.

46. Le Coq, D., Lindner, C., Krüger, S., Steinmetz, M. and Stülke, J. (1995) New β-glucoside (bgl) genes in Bacillus subtilis: the bglP gene product has both transport and regulatory functions similar to those of BglF, its Escherichia coli homolog. J. Bacteriol., 177, 1527–1535.