Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 10410–10414, November 1992
Genetics

Specificity determinants and structural features in the RNA target of the bacterial antiterminator proteins of the BglG/SacY family

(*Bacillus subtilis*/*sacB* gene/levansucrase/RNA-binding protein/transcription)

STÉPHANE AYMERICH* AND MICHEL STEINMETZ

Laboratoire de Génétique, Institut National de la Recherche Agronomique et Centre National de la Recherche Scientifique (Unité de Recherche Associée 537), 78850 Thiverval-Grignon, France

Communicated by Michael J. Chamberlin, July 15, 1992 (received for review April 17, 1992)

ABSTRACT **Induction of the *Bacillus subtilis sacB* gene and *sacPA* operon and *Escherichia coli bgl* operon is mediated by structurally homologous antiterminators encoded by the *sacY*, *sacT*, and *bglG* genes, respectively. When activated, these proteins prevent early transcription termination at terminators located in the leader regions of the three operons. BglG was previously shown to bind *in vitro* to an imperfectly palindromic 29-nucleotide RNA sequence located upstream of the terminator and partially overlapping with it [Houman, F., Diaz-Torres, M. R. & Wright, A. (1990) *Cell* 62, 1153–1163]. Similar motifs, here termed ribonucleic antiterminators (RATs), strongly conserved in sequence and in position, are found in the leader of both *sacB* and *sacPA*. Mutations were created in *sacB* RAT and tested in *B. subtilis*; this showed that *sacB* RAT is the target for SacY-mediated induction of *sacB* and that a stem–loop structure in the mRNA is required for regulatory function. Mutations increasing the similarity of the *sacB* RAT with those of *sacPA* or *bgl* rendered *sacB* inducible by SacT or BglG, respectively; most of these changes did not strongly affect induction by SacY, suggesting that the nucleotides at these variable positions act as negative specificity determinants.**

Induction of the *Bacillus subtilis* levansucrase (*sacB*) gene by sucrose and the *Escherichia coli* β-glucoside (*bgl*) operon by β-glucosides is controlled by similar transcriptional antitermination mechanisms (1–3). In both systems, transcription initiates constitutively from the promoter; in the absence of inducer, most transcripts terminate at a terminator located in the leader region, between the promoter and the first gene; in the presence of inducer, a trans-acting protein (antiterminator) is activated and prevents early transcription termination (1, 2). Sucrose induction of the *B. subtilis* sucrase (*sacPA*) operon appears to follow a similar mechanism (4, 5). The antiterminator BglG is encoded by the first gene of the *bgl* operon. Between *bglG* and the second gene (*bglF*) lies a second BglG-controlled terminator (6). In *B. subtilis*, the *sacY* and *sacT* gene products are required for induction of *sacB* and *sacPA*, respectively (3–5). BglG, SacY, and SacT are very similar proteins, with 48% identity between SacY and SacT and 30% between SacY and BglG (5). It was shown *in vitro* that BglG is an RNA-binding protein that recognizes a specific sequence located upstream of each of the two *bgl* conditional terminators (7). These 29-nucleotide sequences partially overlap with the terminators; thus, binding of BglG to these sequences, hereafter referred to as ribonucleic antiterminator (RAT) sequences, might prevent the formation of the termination structures. Similar RAT sequences are present in the *sacB* and *sacPA* leader regions and overlap with the terminator sequences in exactly the same way (5). This suggests that they are binding sites for SacY and SacT in the leader of *sacB* and *sacPA* mRNAs, respectively. The same model can therefore be proposed for the three systems: when activated by the inducer, the antiterminator binds to the RAT sequence in the nascent mRNA as soon as it emerges from the transcription complex; this binding prevents formation of the terminator structure, allowing transcription read-through. This general model raises several questions. How does binding of the antiterminator to the RAT element prevent terminator formation? Is it by direct sequestration of nucleotides required for terminator formation or by stabilization of a secondary structure of the RAT, itself sequestering critical nucleotides? Do the mRNA sequences upstream of RAT have any role? SacY and SacT control the expression of *sacB* and *sacPA*, respectively, although weak crosstalk can be detected under certain conditions (4). This raises the question: What determines the specificity of the RNA/antiterminator interaction? To test the model for the *sacB*/SacY system, we constructed mutants of the *sacB* RAT region and developed a system allowing *in vivo* measurement of the interactions between SacY and these mutant leader regions. This experimental system was also used to identify the nucleotides that control the specificity of the interaction between one RAT and its cognate antiterminator.

MATERIALS AND METHODS

Bacterial Strains, Transformation, and Transduction. The *E. coli* strain used, TGI (3), was transformed by the calcium-shock procedure. *B. subtilis* was transformed by using natural competence. SPβ lysate was prepared by thermoinduction, and *B. subtilis* strains were transduced as described (8). *B. subtilis* GM856 (*sacB*Δ*23*, SPβSA1) was derived from strain PY480 (8); *sacB*Δ*23* is a deletion mutant of *sacB* (9); the SPβSA1 prophage, whose genotype is SPβ*c2*Δ*2*::Tn*917*::*lacZ* (see Fig. 1), was constructed by replacing the transposon present in PY480 prophage, with the Tn917-*lac* transposon present in pTV32 (10). All *B. subtilis* strains used are isogenic derivatives of GM856; GM882 contains the *sacXY*Δ*3* allele (3); GM904 contains the *sacT*Δ*4* allele in which a *Bcl* I–*Bgl* II deletion has removed all the *sacT* coding sequence except the first 16 and last 4 codons. GM905 is GM856 with both the *sacXY*Δ*3* and *sacT*Δ*4* deletions. The *sacXY*Δ*3* and *sacT*Δ*4* alleles were introduced into GM856 chromosome by using the *in vivo* allele-exchange methodology previously described (3). GM906 was obtained by transformation of GM905 with the replicative plasmid pBG4 expressing the *E. coli bglG4* gene and conferring phleomycin resistance. GM982, a *sacY*-overexpressing GM904 derivative, was obtained by replacing the 1-kilobase-pair *Sal* I–*Sst* I segment containing the 5′ end of *sacX* (3) in the GM904 chromosome with the *aphA′* cassette, a kanamycin-resistance gene devoid of transcription terminator (11); this substitution created an

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: RAT, ribonucleic antiterminator.
*To whom reprint requests should be addressed.

aphA'::*sacY* transcriptional fusion that expressed high levels of SacY and that was constitutively active because of the *sacX* deletion.

Plasmids. pBG4 is replicative in *B. subtilis* and contains the *bglG4* gene, coding for a β-glucoside-independent BglG variant, under control of the inducible *spac* promoter. pBG4 was constructed in two steps: first, the *bglG4*-containing *Eco*109–*Pvu* II fragment from pT7OAC-G4 (12) was inserted at the *Xba* I site of pDG148 (13); second, 5 base pairs were deleted at the *Eco*109 site regenerated at the *bglG4* 5' end to optimize the distance between the ribosome binding site and the initiation codon of the *bglG4* gene. Plasmid pIC38 was obtained by ligation of the large *Nar* I–*Pvu* II fragment of pUC18 (14) with a *Nar* I–*Hpa* I fragment of the plasmid pTSBG6.3 (15), which contains a tripartite fusion of the *B. subtilis trpE* promoter (−48 to +16, constitutive), the leader region of *sacB* (−20 to +199, plus the five first codons of *sacB*), and the 5' end of *lacZ* (Fig. 1).

***B. subtilis* Liquid Cultures and β-Galactosidase Assays.** MMHG is MM minimal medium (9) containing 0.25% (wt/vol) casein hydrolysate and 1% (wt/vol) glucose; CHg medium is minimal C medium (4) containing 0.25% (wt/vol) casein hydrolysate and 1% (wt/vol) glucitol. Strains GM856, GM904, and GM982 were grown in MMHG medium; strains GM856 and GM904 were induced (i.e., SacY-activated) by addition of 2% (wt/vol) sucrose. Strain GM882 was grown in CHg medium and induced (i.e., SacT-activated) by addition of 0.1% (wt/vol) sucrose. Strain GM906 was grown in CHg medium in the presence of phleomycin (0.2 μg/ml); the *spac*::*bglG4* fusion was induced by addition of 0.5 mM isopropyl β-D-thiogalactoside. Samples of cultures (4) were assayed for β-galactosidase activity by the method of Miller (16).

***In Vitro* Mutagenesis.** The small *Bam*HI–*Hin*dIII fragment of pIC38 (Fig. 1) was inserted into pBluescript (Stratagene). Double-stranded plasmid DNA was used for site-directed mutagenesis (17). After mutagenesis, single-stranded DNA was prepared and the sequence of the *Bam*HI–*Acc* I or *Acc* I–*Hin*dIII fragment was checked (14). The mutated fragments were then substituted for the wild-type fragments in pIC38.

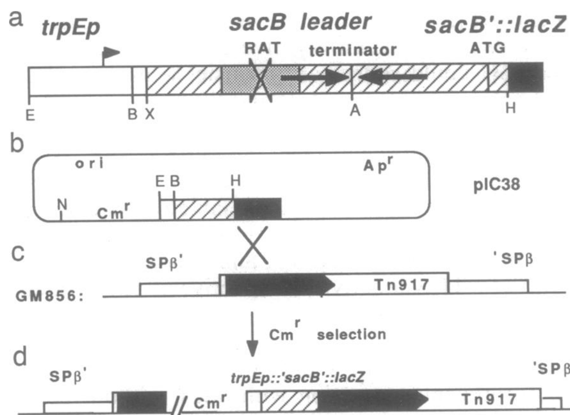

FIG. 1. Genetic system for phenotypic characterization of mutations in the *sacB* leader region. (*a*) Structure of the tripartite fusion of a constitutive promoter (*trpEp*), a *Bam*HI–*Xba* I linker, the promoter-free *sacB* leader region (hatched box), and the 5' end of the *lacZ* gene (black box); the *sacB* leader is the −20 to +199 *sacB* regulatory region plus the first 5 codons of the *sacB* coding sequence (*sacB'*); *sacB'* is fused in frame to the eighth codon of the *lacZ* coding sequence by a *Hin*dIII linker. Restriction sites: E, *Eco*RI; B, *Bam*HI; X, *Xba* I; A, *Acc* I; H, *Hin*dIII. (*b*) pIC38 contains the tripartite fusion, a chloramphenicol-resistance gene (Cm^r) functional in *B. subtilis*, and the ampicillin-resistance gene (Ap^r) and replication origin (ori) from pUC18; some unique restriction sites are indicated. N, *Nar* I. (*c*) GM856 contains the Tn*917*–*lac* transposon inserted into a thermoinducible SPβ prophage (not to scale). (*d*) GM856 transformation to Cm^r with pIC38 DNA created a *trpEp*::'*sacB'*::*lacZ* fusion within the SPβ prophage. SPβ phage lysate prepared from the transformant can be used to transduce the fusion into recipient strains.

RESULTS

The *sacB* RAT may form several different stem–loop structures. Assuming that *sacB, sacPA*, and *bgl* RAT elements (Fig. 2*a*), which are the putative targets of homologous regulatory proteins, form similar structures *in vivo*, we have constructed a model for the *sacB* RAT that is a compromise between stability (18) and applicability to the four RAT sequences (Fig. 2*b*). The differences between the four RAT sequences lie only in bulges or in the loop, except for *bglF* RAT, where a pair of compensatory differences are present in the stem (Fig. 2). The predicted stabilities (18) of the proposed RAT structures (ΔG = −4 to −5 kcal/mol) are much lower than those of the associated terminator structures (ΔG = −18 to −30 kcal/mol). This model is similar to that previously proposed for *bglG* RAT (7).

The system to test *in vivo* mutations created in the *sacB* leader region is schematized in Fig. 1. An integrative plasmid (pIC38 or a mutant derivative) can be inserted by homologous recombination into the *lacZ* gene present within the SPβ prophage in GM856. This results in a tripartite fusion of a constitutive promoter, the *sacB* leader region, and the *lacZ* reporter gene. SPβ phage lysate prepared from the transformant can be used to transduce the fusion into mutant strains; depending on the recipient, the β-galactosidase activity is a marker for the *in vivo* interaction of SacY, SacT, or BglG with the *sacB* wild-type or mutant leader sequence.

Substitution Analysis of the *sacB* Leader Sequence Upstream of RAT. There is no similarity between the sequences upstream of the two *bgl* RATs (19). The model proposed for the regulation of this operon did not assign any role to them (7). Those present upstream of *sacB* and *sacPA* RATs are dissimilar, except for the three nucleotides preceding the 5' end of the RAT, and dissimilar to the corresponding *bgl* sequences (Fig. 2*a*). To test whether the *sacB* RAT sequence is sufficient, upstream of the terminator, for efficient induc-

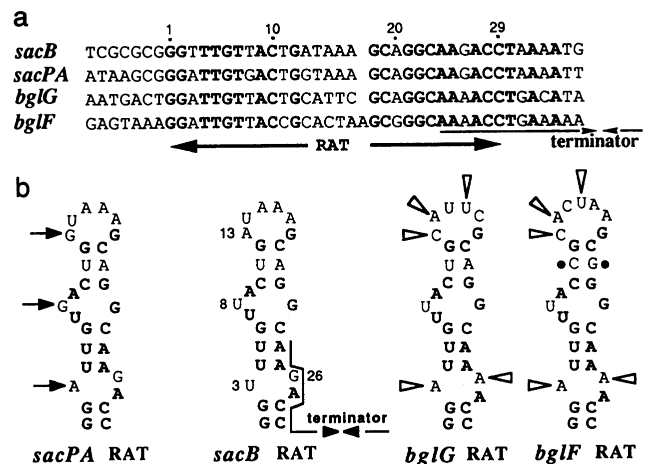

FIG. 2. (*a*) Sequences upstream of the conditional terminators of the *B. subtilis sacB* gene and *sacPA* operon and the *E. coli bgl* operon. The terminator sequences overlap with the four RAT elements in exactly the same way. (*b*) Predicted secondary structures of *sacPA, sacB, bglG*, and *bglF* RAT elements. *sacPA* and *sacB* RAT sequences differ at three positions (indicated by arrows in *sacPA* RAT); five positions are identical in the two *bgl* RATs but differ in the *sacB* RAT sequence (indicated by triangles in *bglG* and *bglF* RATs); a pair of compensatory differences in a stem region of *bglF* RAT is marked with black circles.

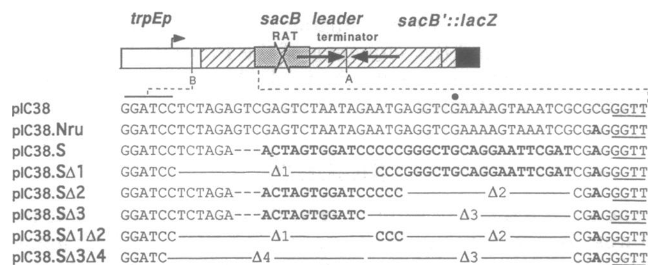

FIG. 3. Substitutions upstream of *sacB* RAT. The heterologous sequences substituted upstream of *sacB* RAT are in boldface letters. The mutations were introduced into the *trpEp*::'*sacB*'::*lacZ* fusion of pIC38 (schematized at the top) by replacement of the *Bam*HI–*Acc* I (B–A) fragment: a C → A point mutation created a *Nru* I site (pIC38Nru); the replacement of the sequence located between this site and the *Bam*HI site led to pIC38.S; and deletions internal to the polylinker led to the pIC38.SΔ set of plasmids. Hyphens indicate gaps introduced for sequence alignment. The 5′ end of RAT is underlined. A black circle indicates the first nucleotide of *sacB* mRNA in *B. subtilis* wild type; in the *trpEp*::'*sacB*'::*lacZ* test fusion the transcription initiates 23 nucleotides upstream of the *Bam*HI site.

tion of *sacB*, the sequence located upstream of RAT was replaced by heterologous sequences (Fig. 3). First, we created an *Nru* I site centered 4 bases upstream of the 5′ end of the *sacB* RAT sequence (mutation Nru). Second, the sequence present upstream of this site was replaced with a polylinker (substitution S). In a third step, deletions were created in the polylinker (Fig. 3). Plasmids carrying these substitutions were inserted into SPβSA1 prophage as described in Fig. 1. The *lacZ* reporter gene was inducible by sucrose in mutant Nru but the β-galactosidase activity was only 80% that of wild type (Table 1). Inducibility was preserved by substitutions S and SΔ1, but expression of *lacZ* reached only 55% that of the wild type. Substitutions SΔ2, SΔ3, SΔ1Δ2, and SΔ3Δ4 had more drastic effects: inducibility was completely lost in the SΔ3 mutant, and *lacZ* expression was largely constitutive in SΔ1Δ2. Two other substitutions by longer and entirely dissimilar and independent sequences were also created: one led to a phenotype similar to that of substitution S, the other to constitutive (and non-overinducible) expression of *lacZ* (data not shown). Overall, these results show that the upstream context of RAT is important for optimal functioning of the system, but also that some sequences dissimilar to that of the wild type can supply an appropriate context.

Table 1. *lacZ* expression in GM856 derivative strains containing a *trpEp*::'*sacB*'::*lacZ* fusion with mutations in the *sacB* leader region

Mutation	β-Galactosidase activity, units	
	− Inducer	+ Inducer
Wild type (pIC38)	1	22
Nru	1	18
S	–	12
SΔ1	–	12
SΔ2	2	10
SΔ3	–	–
SΔ1Δ2	4	10
SΔ3Δ4	1	5
2A	–	–
19U	1	1
20Δ	–	–
6A	–	–
6A/23U	–	3
6C/23U	–	–
6C/23G	–	8

Mutations are described in Figs. 3 and 4. Cultures were grown and induced and extracts were prepared and assayed as described in *Materials and Methods*. Assays were repeated three to five times with samples from independent clones. Standard deviations were 15% or less of the mean when activities of 4 units or more were measured, and 40% or less of the mean for smaller activities. One representative result is given for each strain/mutation combination. –, 0.5 unit or less.

Genetic Evidence for Stem–Loop Pairing of *sacB* RAT. To test the secondary structure model, mutations were created in *sacB* RAT and tested in *B. subtilis*. The 2A, 6A, 19U, and 20Δ mutations (Fig. 4), expected to alter the proposed structure, abolished the inducibility of the *lacZ* reporter gene (Table 1). The 6A/23U double mutation, which consists of 6A plus the presumed compensatory mutation, restored significant inducibility to *lacZ*. Furthermore, in the 6C/23G double mutant (G·C → C·G), *lacZ* was inducible. This demonstrates that pairing between the nucleotides at positions 6 and 23 is essential for *sacB* RAT function. The probable lower stability of the RAT structure in the 6U/23A mutant and a possible modification of RAT three-dimensional geometry in the 6U/23A and 6C/23G mutants could be the reason for the weaker level of *lacZ* induction. These substitutions could also affect direct interactions between the antiterminator and the nucleotides 6 and/or 23.

Mutagenesis of the Nucleotides in *sacB* RAT That Differ from Those in *sacPA* RAT. The *sacPA* and *sacB* RAT sequences differ at only three positions. According to the model, these nucleotides are unpaired. We replaced each of these three nucleotides in *sacB* RAT with those present in *sacPA* RAT (mutations 3A, 8G, and 13G; Fig. 4). The mutated fusions were introduced into strain GM882 [Δ(*sacY*), *sacT*+], where β-galactosidase activity reflects the SacT antitermination activity. Mutation 3A had a significant positive effect on β-galactosidase activity, 13G had a weaker positive effect, and 8G appeared to be neutral (Table 2). The phenotypes of the 3A/8G double mutant and 3A/8G/13G triple mutant showed that the effects are cumulative (Table 2). In GM904 [Δ(*sacT*), *sacY*+], 3A and 13G appeared to be neutral whereas 8G had a negative effect. The neutrality of 3A and 13G could reflect an intrinsic neutrality. Alternatively, SacY could be abundant and therefore saturate *sacB* RAT; this is unlikely, since SacY appears limiting in the wild type (3, 15). This was confirmed: the SPβSA1::pIC38 mutant and wild-type phages were transduced into the SacY-overproducing strain GM982. β-Galactosidase activities were about 20-fold higher than those obtained in induced GM904; the presence of 3A, 8G, and 13G led to 85%, 50%, and 110% of wild-type activity, respectively. Thus, mutations 3A and 13G have little effect on induction of the fusion by SacY.

The U at position 3 and A at position 13 of *sacB* RAT thus appear to prevent interaction with SacT without being required for the interaction with SacY. The mutation 8G seems

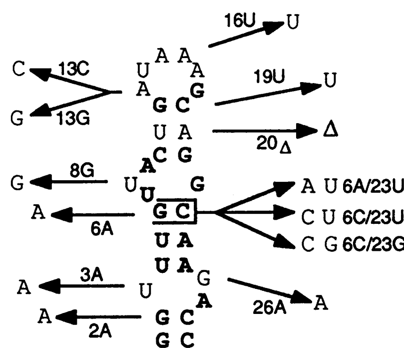

FIG. 4. Mutations in *sacB* RAT. Conserved nucleotides in *sacB*, *sacPA*, and *bgl* RAT sequences are in boldface letters.

Table 2. Phenotype of *sacB* RAT mutations in different genetic backgrounds

Mutation	β-Galactosidase activity, units					
	GM904 (*sacY*+)		GM882 (*sacT*+)		GM906 (*bglG4*)	
	−i	+i	−i	+i	−i	+i
Wild type (pIC38)	−	13	−	4	−	4
3A	−	13	−	34	−	11
8G	−	6	−	5	ND	ND
13G	1	19	1	11	ND	ND
3A/8G	−	11	−	28	ND	ND
3A/8G/13G	1	15	1	45	ND	ND
13C	1	16	ND	ND	1	5
26A/t	1	14	ND	ND	1	5
26A	3	17	ND	ND	ND	ND
t	1	19	ND	ND	ND	ND
3A/26A/t	1	11	ND	ND	1	25
3A/13C/26A/t	1	14	ND	ND	1	40
16U	−	11	−	3	−	3

Mutations are described in Fig. 4, except the t allele, which is a compensatory mutation of 26A (see text). Cultures, extracts, assays, and standard deviations are described in Table 1 legend. −i/+i, Culture in the absence/presence of inducer; −, 0.5 unit or less; ND, not done.

to lower the affinity of the RAT sequence for SacY without increasing that for SacT, so the G at position 8 might play a similar, symmetrical, role in *sacPA* RAT. It is clear, however, that changing the nucleotides at position 3 or 13 of the *sacB* RAT to those present in the *sacPA* RAT increases *sacB* induction by SacT.

Mutations in *sacB* RAT That Allow Recognition by BglG. Role of the Loop. The *sacB*, *bglG*, and *bglF* RAT sequences differ at several positions, most of them being in the loop. There are only five positions where the nucleotides present in both *bgl* RATs are the same but different from that present in the *sacB* RAT (Fig. 2*b*). We replaced the nucleotides present at three of these positions in *sacB* RAT with those present in the two *bgl* RATs (mutations 3A, 13C, and 26A; Fig. 4). The mutation 26A, a G → A substitution, affects both *sacB* RAT and the 5′ end of the terminator; it was therefore associated with the mutation t (a C → T substitution 61 nucleotides downstream) expected to compensate the destabilizing effect of 26A in the terminator structure. These mutations were tested in strain GM906 [Δ(*sacY*), Δ(*sacT*), *bglG4*]. Mutation 3A had a positive effect on the interaction with BglG, whereas 13C and 26A appeared to be neutral (Table 2). There was synergy between 3A and 13C or 26A: when 26A was associated with 3A, *lacZ* expression was double that with 3A alone. When 13C was introduced into the 3A/26A double mutant, *lacZ* expression again doubled (Table 2). This demonstrates the involvement of residues at positions 3, 13, and 26 of *bglG* RAT in the efficient interaction with BglG. In GM904 [Δ(*sacT*), *sacY*+], the presence of the 3A, 13C, and 26A mutations, either independently or in combination, had no significant effect on *lacZ* expression (Table 2).

The central parts of the two *bgl* RATs (i.e., the upper part of the loop) are different (Fig. 2). BglG can also efficiently interact with the 3A/13C/26A mutant *sacB* RAT, which is identical to *bglG* RAT except for the loop. A mutation (16U; Fig. 4) in the loop of *sacB* RAT was neutral for induction by SacY, SacT, or BglG (Table 2). These observations suggest that recognition by the antiterminator is largely independent of the nucleotides of the top of the loop in RAT sequences.

DISCUSSION

The *bgl*, *sacB*, and *sacPA* RAT sequences have the potential to form similar stem–loop structures (Fig. 2); since their 3′ end overlaps with the 5′ end of the terminator sequences, the proposed structure for RAT and that of the terminator cannot both form simultaneously. However the predicted stability of the RAT structure is lower than that of the terminator. The antiterminator might stabilize the RAT structure, thus preventing the formation of the terminator. Mutations expected to destabilize the *sacB* RAT structure (2A, 6A, 19U, 6C/23U, and 20Δ) were shown to abolish induction of the *sacB′*::*lacZ* gene fusion, although mutations affecting nucleotides supposed to be unpaired (3A, 8G, 13G, 13C, 16U, and 26A) had little effect on induction of the fusion by SacY, with the exception of 8G, which had a weak negative effect. The mutations that abolished induction may affect the interaction between RAT and the antiterminator, base pairings crucial for the RAT structure, or both. But the phenotype conferred by two pairs of compensatory mutations (6A/23U and 6C/23G) gave strong evidence for the pairing of nucleotides at positions 6 and 23. We conclude that this pairing is necessary for efficient function of *sacB* RAT *in vivo* and underline that the phenotypes of all mutations were consistent with the proposed model (Fig. 5).

The sequences upstream of the four RAT sequences differ. However, the sequence upstream of the *sacB* RAT has a role: substitutions with heterologous sequences significantly lowered or abolished induction of the reporter gene. Assuming that these substitutions did not modify the stability of the mRNA the results suggested that the wild-type sequence upstream from *sacB* RAT gives the context required for the folding of RAT in the nascent mRNA before completion of transcription of the terminator. This could be important for such a target because RNA molecules begin folding as they are synthesized, and sequences could become kinetically trapped in a structure that is not the one required for function. A similar (non-neutral) role has been attributed to the sequence context of the RNA target of the human immunodeficiency virus Rev protein (20). In *E. coli*, sequences linked to promoters can affect the efficiency of downstream Rho-independent termination sites, probably by provoking the change of the RNA polymerase conformation from a resistant to a competent state for termination (21). Such a phenome-

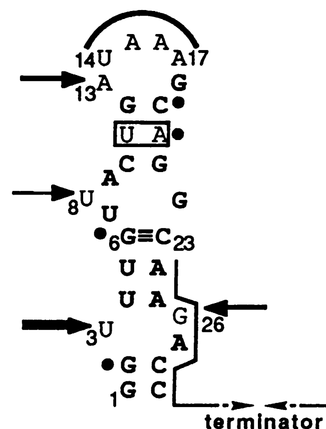

FIG. 5. Structural model of *sacB* RAT and specificity determinants. Conserved nucleotides in the four RAT sequences are in boldface letters; the 29 residues of *sacB* RAT are numbered as in Fig. 2; the 6 bases overlapping the terminator are indicated. Base pairing between nucleotides 6 and 23 (indicated by a triple hyphen) was demonstrated with pairs of compensatory mutations; point mutations at positions indicated by black circles abolished RAT function; the positions framed (U·A base pairs) are C·G in *bglF* RAT. Nucleotides at positions indicated by arrows are involved in the control of the specificity of the RAT/antiterminator interaction: the thickness of the arrow is proportional to the magnitude of the phenotype of the mutation; recognition by the antiterminator appears largely independent of the nucleotides of the upper part of the loop (positions 14–17).

non, if it exists in *B. subtilis*, could explain the phenotype of some of the mutants.

BglG binding to *bglG* RAT has been demonstrated *in vitro* (7). Mutant *sacB* RAT with nucleotide 3 or 13 replaced by that present in *sacPA* RAT was an efficient target for SacT; similarly, *sacB* could be efficiently induced by BglG when substitutions made *sacB* RAT more similar to *bgl* RAT. This shows that the RAT sequences are the *in vivo* targets of the antiterminator proteins.

An unexpected observation was the synergic effect observed between a mutation (3A) that increased *sacB* induction by BglG and mutations (13C and 26A) that appeared neutral. The *in vivo* experimental procedure used in this study results in data reflecting a sum of interactions and events. RAT mutations might differentially affect the association and dissociation parameters of the RAT/antiterminator complex; 3A may be particularly efficient with BglG by increasing the turnover of the complex and, therefore, the concentration of free BglG; 13C and 26A may stimulate the association without stimulating induction if they do not increase the turnover. The synergy could also reflect a two-step process in the RAT/antiterminator binding, associated with a conformational change of the target: a U at position 3 could prevent the transition from a poor to an efficient interaction with BglG; nucleotides 13 and 26 might affect the interaction only after this transition. These detailed mechanisms are, however, speculations.

Mutations of *sacB* RAT that result in *sacB* induction by SacT or BglG did not strongly affect induction by SacY; the nucleotides at positions 3, 13, and 26 in wild-type *sacB* RAT appear to prevent binding with noncognate antiterminators and therefore play a role of negative specificity determinants. This type of specificity determinant has been described for a few RNA–protein interactions (22). In the case of yeast $tRNA^{Asp}$, three posttranscriptional modifications are essential to prevent mis-aminoacylation by arginyl-tRNA synthetase (23).

In conclusion (Fig. 5), some bases in the stem regions of the model appear to be paired and thus determine the formation of the *sacB* RAT structure required for function. Some of the unpaired bases control the specificity of the RAT/antiterminator interaction: the residues 3U, 13A, and 26A of *sacB* RAT appear to be negative specificity determinants. The results do not allow us to identify the bases that directly interact with the antiterminators; this function may be fulfilled by some of the paired nucleotides and/or the unpaired nucleotides conserved in the four RAT sequences and located in the immediate proximity of the nucleotides involved in negative specificity discrimination.

Sequence similarities among the three RAT/antiterminator systems suggest that they evolved from one common ancestor. The original regulatory system might have been duplicated to control expression of different metabolic pathways in the same cell. Our results suggest that the divergence of the duplicated regulatory systems could have occurred by selection of substitutions that reduced inappropriate interactions rather than increasing the appropriate ones.

We gratefully acknowledge Orna Amster-Choder, whose enthusiasm made this work possible. We thank Dennis Henner, Patrick Stragier, and Philip Youngman for giving strains and plasmids. We thank Dick D'Ari for critical reading of the manuscript, Alex Edelman for help with the English language, and Claude Gaillardin. We thank Dominique Le Coq and Roselyne Richter for cloning the *sacT* gene.

1. Mahadevan, S. & Wright, A. (1987) *Cell* **50**, 485–494.
2. Shimotsu, H. & Henner, D. (1986) *J. Bacteriol.* **168**, 380–388.
3. Crutz, A. M., Steinmetz, M., Aymerich, S., Richter, R. & Le Coq, D. (1990) *J. Bacteriol.* **172**, 1043–1050.
4. Steinmetz, M., Le Coq, D. & Aymerich, S. (1989) *J. Bacteriol.* **171**, 1519–1523.
5. Débarbouillé, M., Fouet, A., Arnaud, M., Klier, A. & Rapoport, G. (1990) *J. Bacteriol.* **172**, 3966–3973.
6. Schnetz, K. & Rak, B. (1988) *EMBO J.* **7**, 3271–3278.
7. Houman, F., Diaz-Torres, M. R. & Wright, A. (1990) *Cell* **62**, 1153–1163.
8. Poth, H. & Youngman, P. (1988) *Gene* **73**, 215–226.
9. Steinmetz, M., Le Coq, D., Aymerich, S., Gonzy-Tréboul, G. & Gay, P. (1985) *Mol. Gen. Genet.* **200**, 220–228.
10. Perkins, J. B. & Youngman, P. J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 140–144.
11. Zagorec, M. & Steinmetz, M. (1991) *J. Gen. Microbiol.* **137**, 107–112.
12. Amster-Choder, O., Houman, F. & Wright, A. (1989) *Cell* **58**, 847–855.
13. Stragier, P., Bonamy, C. & Karmazyn-Campelli, C. (1988) *Cell* **52**, 697–704.
14. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed.
15. Henner, D. J., Yang, M., Band, L., Shimotsu, H., Ruppen, M. & Ferrari, E. (1987) in *Genetics of Industrial Microorganisms*, Proceedings of the Fifth International Symposium on the Genetics of Industrial Microorganisms, eds. Alacevic, M., Hranueli, D. & Toman, Z. (Ognjen Prica Printing Works, Karlovac, Yugoslavia), pp. 81–90.
16. Miller, J. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
17. Hemsley, A., Arnheim, N., Toney, M. D., Cortopassi, G. & Galas, D. J. (1989) *Nucleic Acids Res.* **17**, 6545–6551.
18. Zuker, M. (1989) *Methods Enzymol.* **180**, 262–288.
19. Schnetz, K., Toloczyki, C. & Rak, B. (1987) *J. Bacteriol.* **169**, 2579–2590.
20. Bartel, D. P., Zapp, M. L., Green, M. R. & Szostak, J. W. (1991) *Cell* **67**, 529–536.
21. Telesnitsky, A. P. W. & Chamberlin, M. J. (1989) *J. Mol. Biol.* **205**, 315–330.
22. Schulman, L. H. (1991) in *Progress in Nucleic Acid Research and Molecular Biology*, eds. Cohn, W. E. & Moldave, K. (Academic, San Diego), Vol. 41, pp. 23–88.
23. Putz, J., Puglisi, J. D., Florentz, C. & Giegé, R. (1991) *Science* **252**, 1696–1699.