Published online March 5, 2004

1610–1614 *Nucleic Acids Research, 2004, Vol. 32, No. 4*
DOI: 10.1093/nar/gkh321

A theophylline responsive riboswitch based on helix slipping controls gene expression *in vivo*

Beatrix Suess*, Barbara Fink, Christian Berens, Régis Stentz and Wolfgang Hillen

Lehrstuhl für Mikrobiologie, Friedrich-Alexander-Universität Erlangen-Nürnberg, Staudtstrasse 5, 91058 Erlangen, Germany

Received January 13, 2004; Revised and Accepted February 11, 2004

ABSTRACT

Riboswitches are newly discovered regulatory elements which control a wide set of basic metabolic pathways. They consist solely of RNA, sense their ligand in a preformed binding pocket and perform a conformational switch in response to ligand binding resulting in altered gene expression. We have utilized the enormous potential of RNA for molecular sensing and conformational changes to develop novel molecular switches with predetermined structural transitions in response to the binding of a small molecule. To validate these *in vivo*, we exploit the distance-dependent inhibitory potential of secondary structure elements placed close to the bacterial ribosome binding site. We created a translational control element by combining the theophylline aptamer with a helical communication module for which a ligand-dependent one-nucleotide slipping mechanism had been proposed. This structural element was inserted at a position just interfering with translation in the non ligand-bound form. Addition of the ligand then shifts the inhibitory element to a distance which permits efficient translation. We present here a novel regulatory mechanism in the first rationally designed, *in vivo* active RNA switch. Its use of a slippage mechanism to control gene expression makes it different from natural riboswitches which are based on sequestration or antitermination.

INTRODUCTION

RNA adopts a wide range of sophisticated structures that can function as precise genetic switches by direct interaction between small metabolites and allosteric mRNAs. Naturally occurring RNA switches are structural domains that are embedded within the non-coding sequences of certain mRNAs where they directly serve as metabolite-responsive genetic control elements. Thereby, the RNA element can adopt distinct structures with one of these serving as a precise binding pocket for its target metabolite. The metabolite-induced allosteric interconversion between structural states causes a change in the level of gene expression. In this way, naturally occurring riboswitches serve as master control elements for a wide set of basic metabolic pathways, sensing metabolites like coenzyme B_1 and B_{12}, FMN, SAM, lysine, guanine and adenine (1–11). They utilize a variety of different mechanisms such as interference with ribosomal accessibility by masking binding sites (1) or attenuation (4). In addition, examples of splicing regulation and influences on RNA stability are discussed for eukaryotes (12).

This enormous potential of RNA for molecular sensing and conformational changes predestines the use of such switches to develop novel RNA-based conditional gene expression systems. Current reports about RNA–small molecule-based regulatory systems in eukaryotes make use of the conformational control of RNA aptamers. Aptamers recognize their ligands with high specificity and affinity, and complex formation is often accompanied by structural changes and/or increased rigidity (13,14). Such complexes become an obstacle for ribosome binding or movement when the aptamer is placed in non-coding regions of an mRNA. As a result, gene expression is switched off in the presence of the ligand, but proceeds in its absence as has been shown for aptamers binding dyes (15), biotin (16) or tetracycline (17,18).

In prokaryotes, the spacing between ribosome binding site (RBS), called Shine-Dalgarno sequence, and start codon ranges from 5 to 13 nt (19). An aptamer-based regulatory system as described above for eukaryotes is therefore not realizable in bacterial systems.

Initiation of translation of many bacterial mRNAs depends simply on the RBS and the nearby start codon being accessible to ribosomes (20). The formation of secondary structure elements which sequester the RBS and/or start codon in a region of double-stranded RNA serves as the main determinant in translational regulation. Mutations that disrupt such structures increase the translational efficiency (21,22). We exploited the regulatory potential of inhibitory secondary structure on translational initiation by creating a new kind of riboswitch with the capability to shift its location dependent on ligand binding. Thus, at the one location, the riboswitch is close to the RBS and interferes with ribosome accessibility. Binding of the ligand then increases the distance to the RBS which, in turn, facilitates the ribosome accessibility to the translational initiation region.

Such a ligand-dependent distance switch has been proposed for the communication modules of allosterically regulable ribozymes (23). These helical elements were originally selected *in vitro* as a structural bridge for coupling a catalytic

*To whom correspondence should be addressed. Tel: +49 9131 852 88 18; Fax: +49 9131 852 80 82; Email: bsuess@biologie.uni-erlangen.de

with a receptor domain. Ligand binding to the receptor domain triggers a conformational change in the communication domain, dictating the activity of the adjoining ribozyme. A slipping mechanism of the communication module resulting in a reorganization of the bridge was discussed as possible mechanism for the structural transition (23). These allosteric ribozymes have a highly modular nature (23,24). The FMN-binding aptamer which was used as the receptor domain during the *in vitro* selection process was replaceable by aptamers responding to ATP or theophylline (23).

We created an RNA switch composed of two independent domains: the theophylline aptamer as a receptor domain for specific ligand binding (25) and the communication module I-2 (23) proposed to perform helix slipping. This element was inserted close to the RBS so that its non-bound conformation interfered with ribosome accessibility. Binding of the ligand theophylline then induced a structural transition in the bridge helix leading to a 1 nt shift which moved the element away by exactly the critical distance to allow ribosome binding. Thus, we have rationally designed an *in vivo* active molecular riboswitch which employs helix slippage as a novel mechanism to regulate gene expression.

MATERIALS AND METHODS

Construction of the *Bacillus subtilis* strains

A 615 bp internal fragment of the *xylA* gene was amplified by genomic PCR using the primer pairs xabs1 (5′-GATCT-CTAGAAGATTCGACTAATCCTTTAGC) and xabs2 (5′-GATCGAGCTCGCATAATCCACTGCCATATGC). The fragment was cloned into pBluescript SK$^-$ (Stratagene, La Jolla, USA) using restriction sites for XbaI and SacI. The resulting plasmid was named pWH1179. The BamHI/EcoRI fragment of the plasmid pSpec (P.Trieu-Cuot, Institut Pasteur, France) containing the spectinomycin resistance gene *aad(9)* was cloned into the BamHI and EcoRI sites of pWH1179, resulting in pWH1180. An 814 bp DNA fragment containing the wild type *xylR* gene and the regulatory region of the divergently oriented genes *xylA* and *xylR* was amplified by genomic PCR using the primer pairs xrbs1 (5′-TGC-AGAATTCGTTGATTGAACTGGAATGAG) and xrbs4 (5′-TGCACTCGAGGTACTTCTCTTGTATCGAAG). The fragment was cloned into pWH1180 using restriction sites for EcoRI and XhoI, resulting in pWH1181. pWH1181 was integrated into the chromosome of *B.subtilis* WH335 [*amyE::Pxyl*$_{396bp}$, EcoRI–HindIII-*spoVG-lacZ*, *Cm*^{R}, *trpC2*, (26)] which contains a *xylA::lacZ* fusion integrated at the *amyE* locus. The resulting *B.subtilis* strain was WH441.

In order to introduce sequences upstream of the RBS of the *xylR* gene, an overlap extension PCR was carried out using internal primers to introduce the additional sequence motif and the outside primer pairs xrbs1 and xrbs4. The resulting PCR products were cloned into pWH1181 using EcoRI and XhoI. The resulting plasmids were integrated into the chromosome of *B.subtilis* WH335. Sequences of the internal overlapping primers are available upon request.

LacZ measurements

Bacillus strains were grown in minimal medium supplemented with 0.5% succinate at 37°C (27) and harvested at an A_{600} = 0.4 and 0.8, respectively. β-Galactosidase was determined by the method of Miller (28) and normalized to cell density. Measurements were repeated at least twice for two independent cultures.

RESULTS

Distance-dependent translation inhibition by hairpins

We wanted to devise an inhibitory RNA element which interferes with ribosome binding and with the capability to reversibly shift its location by 1 nt. It was, therefore, necessary to identify the critical distances from the RBS for an inhibitory stem–loop at which a difference of 1 nt allows regulation of translation. We inserted a 23 nt fragment predicted to form a stem–loop structure at different positions upstream of the RBS of the *xylR* gene. Expression of the encoded Xyl repressor was indirectly determined by repression of a *xylA::lacZ* fusion. This allows us to quantify the inhibitory effect on *xylR* expression by measuring β-galactosidase activity in the absence of xylose. A schematic representation of the genomic arrangement is shown in Figure 1A. The secondary structure prediction of a 103-nt long fragment containing the 5′ leader region and the first 30 nt of the *xylR* gene was performed using mfold version 3.1 [http://www.bioinfo.rpi.edu/applications/mfold (29)] and is shown in Figure 1B. We placed the 23 nt stem–loop directly upstream of the RBS in construct K1 (Fig. 1B). In constructs K2–K4, we inserted one and two additional nucleotides between the stem–loop and the RBS sequence in order to introduce different distances (K2, +A; K4, +AA; K3, +A combined with an additional disruption of the first base pair of the stem–loop). The secondary structure prediction of all constructs was done in the context of the 103-nt long fragment. Sections containing the RBS and the introduced stem–loop of all constructs are displayed in Figure 1B and were integrated in single copy into the chromosome of *B.subtilis* WH335 (26). β-Galactosidase activity of the *xylA::lacZ* fusion was measured in the absence and presence of 0.2% xylose. The data are displayed in Figure 1C.

All constructs show full expression of the *xylA::lacZ* fusion in the presence of xylose indicating that XylR does not bind the operator sequences in its ligand-bound form. The β-galactosidase activity of the wild type strain (wt) is completely repressed in the absence of xylose. K1 shows complete loss of repression indicating that the inserted stem–loop abolishes expression of *xylR*. Moving away the stem–loop structure by only 1 nt restores repression of *xylA::lacZ* expression almost completely. This effect is further increased by moving the stem–loop away by 2 nt.

Taken together, a stem–loop inserted directly besides the RBS interferes with gene expression. Shifting the element away by only 1 nt is sufficient to abrogate this effect.

A theophylline riboswitch controls gene expression

We then replaced the stem–loop by a rationally designed RNA switch which is composed of the theophylline aptamer as receptor domain [boxed in Fig. 2B, (25)] and a communication domain [class II-1 inhibition element (23), highlighted in grey in Fig. 2B]. A schematic view of the genomic organization is shown in Figure 2A and a speculative secondary structure of

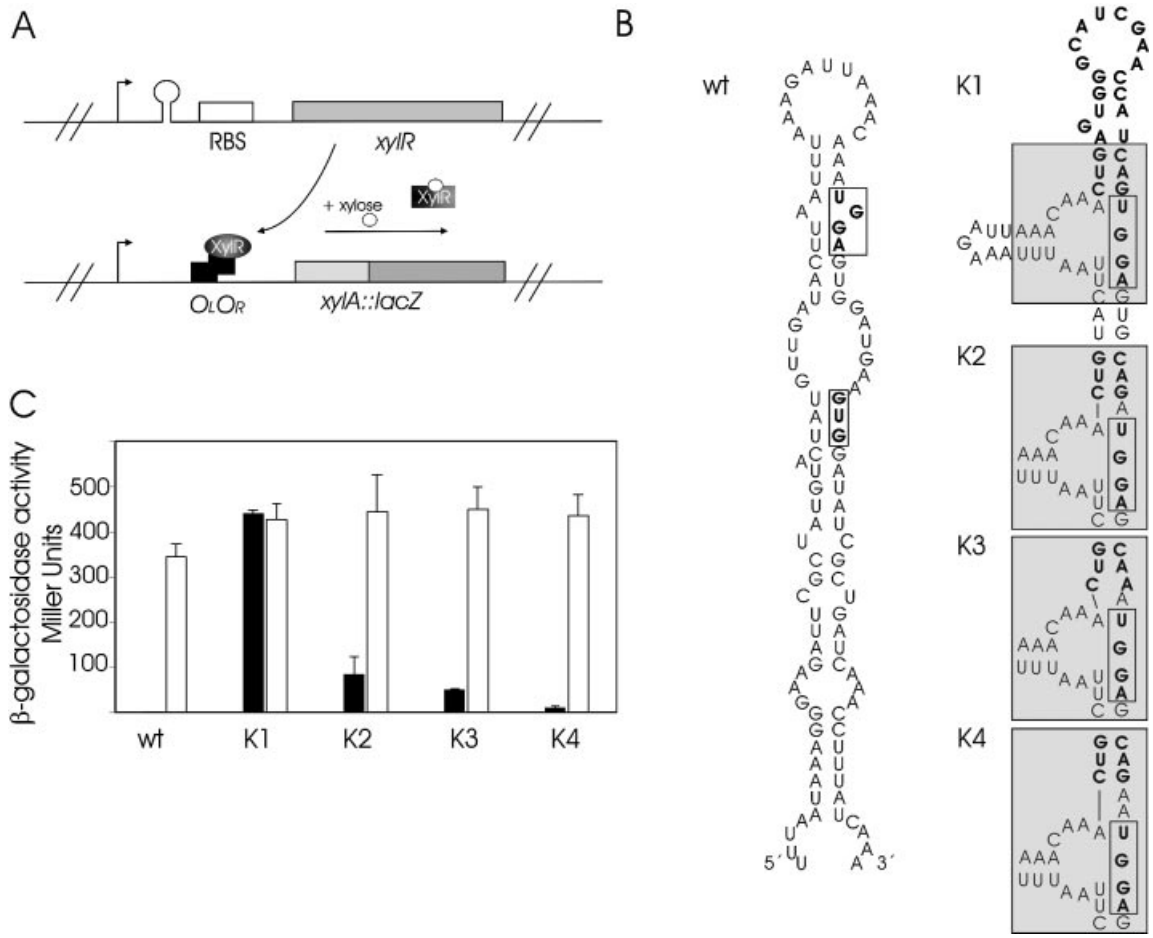

Figure 1. Inhibitory potential of a stem–loop inserted at varying distances from RBS. (**A**) Schematic view of the genomic situation of the regulatory system. The *xylR* gene (grey box) is constitutively expressed. The transcriptional start site (arrow), the RBS (open box) and the stem–loop element are indicated. The Xyl repressor (XylR, oval) binds the two *xyl* operators (closed boxes), thereby repressing the *xylA::lacZ* fusion gene. Xylose-induced XylR (rectangle) does not bind the operator sequences allowing gene expression. (**B**) Secondary structure prediction [mfold web server version 3.1; http://www.bioinfo.rpi.edu/applications/mfold (29)] of a 103-nt long fragment containing the translation initiation region and 30 nt of the coding region of *xylR* without (wt) and with the introduced stem–loop elements K1–K4. For K1–K4 only the section containing the introduced stem–loop and the RBS is depicted. The ribosomal binding site (UGGA) and the start codon GUG are boxed, the nucleotides introduced are shown in bold letters. (**C**) The plot displays the β-galactosidase activity of the *xylA::lacZ* fusion measured in the absence (closed) and presence (open bars) of 0.2% xylose.

the RNA switch derived from the proposed structure of the communication module (23) is pictured in Figure 2B together with the proposed slipping mechanism. The construct was integrated in the chromosome of *B.subtilis* WH335 resulting in the strain T1. β-Galactosidase activity expressed from the *xylA::lacZ* fusion was measured in the absence and presence of 0.2% xylose and 10 mM theophylline. In addition, we analysed K1, which lacks a theophylline responsive element, to exclude side effects of theophylline on either *xylR* or *xylA* expression. The data are displayed in Figure 2C. Furthermore, we monitored the growth of both strains in the absence and presence of 10 mM theophylline. No influence on growth rate was detectable (data not shown).

The presence of theophylline has no influence on the β-galactosidase expression of the control constructs K1 and the xylose-induced form of T1. However, in the absence of xylose the reporter activity is only repressed in the presence of theophylline. We then asked whether the observed effect is dose dependent and specific for theophylline. *Bacillus* strains K1 and T1 were grown with varying concentrations of theophylline and the closely related compound caffeine, which binds to the aptamer with 1000-fold lower affinity (25). The data are displayed in Figure 2D. We observed a clear dose-dependent decrease in reporter activity for T1 in the presence of theophylline. Maximum 8-fold repression is obtained with 6 mM theophylline. No inhibition by caffeine was detectable. This indicates the specificity of the riboswitch for theophylline.

Taken together, we have used rational design to create a theophylline responsive RNA switch which is able to control gene expression.

DISCUSSION

Binding of a translational initiation site by the eubacterial ribosome occurs only if the sequences surrounding the RBS are not involved in stable base pairing (19,21). We inserted different stem–loops upstream of the RBS of *xylR* which

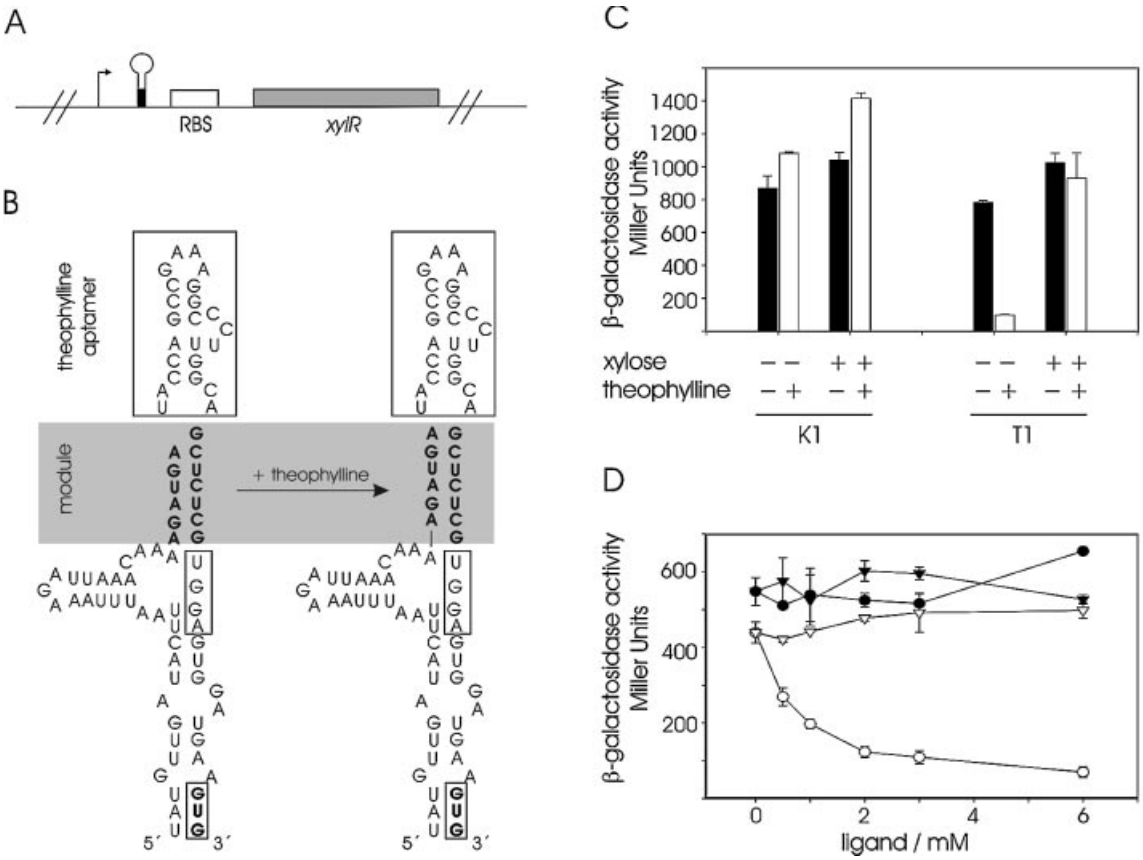

Figure 2. Theophylline responsive riboswitch controls gene regulation. (**A**) Schematic view of the genomic situation of the regulatory system. The *xylR* gene (grey box) is constitutively expressed. The transcriptional start site (arrow) and the RBS (open box) are indicated. The receptor domain of the riboswitch is shown as an open stem–loop and the communicator module as a closed stem. (**B**) Predicted secondary structure of the translation initiation region of *xylR* with the introduced regulatory element consisting of the theophylline aptamer (boxed) fused to a bridge domain (highlighted in grey). The proposed theophylline-mediated slipping mechanism of the bridge domain is shown schematically. (**C**) The plot displays the β-galactosidase activity measured in the absence and presence of 0.2% xylose and 10 mM theophylline, respectively. The presence (+) and absence (–) of xylose and/or theophylline is labelled beneath the plot. (**D**) β-Galactosidase activity of T1 (open) and K1 (closed symbols) measured in the absence of xylose and with increasing concentrations of theophylline (circles) and caffeine (triangles).

introduce stable base pairing into the translation initiation region to determine the minimal necessary accessible area at which translation can occur. In the strain K1, the structural element is placed 13 nt away from the start codon. Repression of the *xylA::lacZ* fusion is completely abolished in this construct indicating that the stable base pairing prevents translation of XylR. In the constructs K2 and K3, the distances between the start codon and the stem–loops are elongated by only 1 nt, to 14 (see Fig. 1B). In contrast to K1, these constructs show β-galactosidase activities comparable with wt indicating that 14 nt free of stable base-pairing seems to be sufficient to allow efficient binding of the ribosome to the initiation site in our system. The result fits well with experiments in which ~15 nt on each side of the start codon are protected by the 70S ribosome against digestion by ribonucleases (30).

The stability of the stem–loop additionally contributes to its regulatory properties. Destabilization of the stem–loop K2 by exchanging a Watson–Crick CG base pair into a non-canonical CA base pair in the strain K3 leads to better repression. This is in agreement with mutagenesis data of the translational initiation region of the MS2 coat protein. These studies revealed a strict correlation between translational efficiency and the stability of the secondary structure formed by the initiation region (21,22). For the implementation of the riboswitch, we used the theophylline aptamer as receptor domain since theophylline is not a natural metabolite in *B.subtilis* and does not interfere with the cellular metabolism. It is therefore a convenient inducer for conditional gene expression. In addition, it has already been successfully coupled with several communication modules [VI, VII (23)]. We attached the aptamer to the communication module II-1 since the two conformations proposed for II-1 correspond well with the base arrangements of K1 and K3. The ligand-free form of the communication module correlates with K1. In both cases, the U of the AU pair at the end of the stem is part of the RBS sequence. Ligand binding then shifts the communication module to a base arrangement similar to the situation in K3. This is achieved by disrupting a base pair in the Shine–Dalgarno site of the stem, while creating one on the opposing receptor site. The result is an unpaired G directly adjacent to the RBS that is followed by a non-canonical AC base pair, the same situation as in K3 with an unpaired A and a mispaired CA. Thus, the shift in the communication module mimics the base arrangements of K1 and K3. The regulatory factor of 8.0 obtained nearly reaches the maximum regulatory window of

8.8 spanned by K1 (441 Miller Units) and K3 (50 Miller Units).

Modular rational design has proven to be an effective means for conferring additional chemical and kinetic complexity on existing proteins and RNA enzymes. But, until now, for none of the RNA elements created by rational design has an *in vivo* function been demonstrated. For the first time, we present a rationally designed RNA element which is able to efficiently control translation dependent on a small molecule ligand *in vivo*. Thereby, we were able to verify the proposed slipping mechanism of the communication modules by using an *in vivo* reconstruction in which we could demonstrate its *in vivo* functionality and general applicability.

ACKNOWLEDGEMENTS

We thank Dr Trieu-Cuot, Institut Pasteur, for providing the plasmid pSpec. We are grateful to the Volkswagenstiftung and the Fond der Chemischen Industrie. B.S. was a recipient of a personal grant from the 'Bayerischer Habilitationsförderpreis'.

REFERENCES

1. Winkler,W., Nahvi,A. and Breaker,R.R. (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature*, **419**, 952–956.
2. Mironov,A.S., Gusarov,I., Rafikov,R., Lopez,L.E., Shatalin,K., Kreneva,R.A., Perumov,D.A. and Nudler,E. (2002) Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell*, **111**, 747–756.
3. Nahvi,A., Sudarsan,N., Ebert,M.S., Zou,X., Brown,K.L. and Breaker,R.R. (2002) Genetic control by a metabolite binding mRNA. *Chem. Biol.*, **9**, 1043–1049.
4. Winkler,W.C., Cohen-Chalamish,S. and Breaker,R.R. (2002) An mRNA structure that controls gene expression by binding FMN. *Proc. Natl Acad. Sci. USA*, **99**, 15908–15913.
5. Winkler,W.C., Nahvi,A., Sudarsan,N., Barrick,J.E. and Breaker,R.R. (2003) An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nat. Struct. Biol.*, **10**, 701–707.
6. Epshtein,V., Mironov,A.S. and Nudler,E. (2003) The riboswitch-mediated control of sulfur metabolism in bacteria. *Proc. Natl Acad. Sci. USA*, **100**, 5052–5056.
7. Sudarsan,N., Wickiser,J.K., Nakamura,S., Ebert,M.S. and Breaker,R.R. (2003) An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes Dev.*, **17**, 2688–2697.
8. Mandal,M., Boese,B., Barrick,J.E., Winkler,W.C. and Breaker,R.R. (2003) Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. *Cell*, **113**, 577–586.
9. Winkler,W.C. and Breaker,R.R.(2003) Genetic control by metabolite-binding riboswitches. *Chembiochem.*, **4**, 1024–1032.
10. Nahvi,A., Barrick,J.E. and Breaker,R.R. (2004) Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes. *Nucleic Acids Res.*, **32**, 143–150.
11. Nudler,E. and Mironov,A.S. (2004) The riboswitch control of bacterial metabolism. *Trends Biochem. Sci.*, **29**, 11–17.
12. Sudarsan,N., Barrick,J.E. and Breaker,R.R. (2003) Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA*, **9**, 644–647.
13. Patel,D.J., Suri,A.K., Jiang,F., Jiang,L., Fan,P., Kumar,R.A. and Nonin,S. (1997) Structure, recognition and adaptive binding in RNA aptamer complexes. *J. Mol. Biol.*, **272**, 645–664.
14. Hermann,T. and Patel,D.J. (2000) Adaptive recognition by nucleic acid aptamers. *Science*, **287**, 820–825.
15. Werstuck,G. and Green,M.R. (1998) Controlling gene expression in living cells through small molecule–RNA interactions. *Science*, **282**, 296–298.
16. Harvey,I., Garneau,P. and Pelletier,J. (2002) Inhibition of translation by RNA–small molecule interactions. *RNA*, **8**, 452–463.
17. Suess,B., Hanson,S., Berens,C., Fink,B., Schroeder,R. and Hillen,W. (2003) Conditional gene expression by controlling translation with tetracycline-binding aptamers. *Nucleic Acids Res.*, **31**, 1853–1858.
18. Hanson,S., Berthelot,K., Fink,B., McCarthy,J.E.G. and Suess,B. (2003) Tetracycline-aptamer-mediated translational regulation in yeast. *Mol. Microbiol.*, **49**, 1627–1637.
19. Kozak,M. (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene*, **234**, 187–208.
20. deSmit,M.H. and van Duin,J. (1990) Control of prokaryotic translational initiation by mRNA secondary structure. *Prog. Nucleic Acid Res. Mol. Biol.*, **38**, 1–35.
21. deSmit,M.H. and van Duin,J. (1990) Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proc. Natl Acad. Sci. USA*, **87**, 7668–7672.
22. deSmit,M.H. and van Duin,J. (1994) Control of translation by mRNA secondary structure in *Escherichia coli*. A quantitative analysis of literature data. *J. Mol. Biol.*, **244**, 144–150.
23. Soukup,G.A. and Breaker,R.R. (1999) Engineering precision RNA molecular switches. *Proc. Natl Acad. Sci. USA*, **96**, 3584–3589.
24. Silverman,S.K. (2003) Rube Goldberg goes (ribo)nuclear? Molecular switches and sensors made from RNA. *RNA*, **9**, 377–383.
25. Jenison,R.D., Gill,S.C., Pardi,A. and Poliski,B. (1994) High-resolution molecular discrimination by RNA. *Science*, **263**, 1425–1429.
26. Kraus,A., Hueck,C., Gärtner,D. and Hillen,W. (1994) Catabolite repression of the *Bacillus subtilis* xyl operon involves a cis element functional in the context of an unrelated sequence and glucose exerts additional xylR-dependent repression. *J. Bacteriol.*, **176**, 1738–1745.
27. Gärtner,D., Geissendörfer,M. and Hillen,W. (1988) Expression of the *Bacillus subtilis xyl* operon is repressed at the level of transcription and is induced by xylose. *J. Bacteriol.*, **170**, 3102–3109.
28. Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
29. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
30. Steitz,J.A. and Jakes,K. (1975) How ribosomes select initiator regions in mRNA: base pair formation between the 3′ terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **72**, 4734–4738.