Another Face of RNA: Metabolite-Induced "Riboswitching" for Regulation of Gene Expression

Sabine Müller*[a]

KEYWORDS:

gene expression · regulation · riboswitches · RNA · thiamine

Control of gene expression is a crucial feature of all organisms. At any given time only a fraction of the total number of genes is active. Regulation of cell metabolism primarily takes place at the level of transcription but is supplemented by translational control and several other mechanisms. The entire spectrum of controls provides the cell with a sensitive tool for regulating the kinds and numbers of cellular molecules produced. In 1965 Francois Jacob, André Lwoff, and Jacques Monod received the Nobel Prize for their discoveries concerning genetic control of enzyme and virus synthesis. The *lac* operon described by Jacob and Monod[1] is found in all genetic textbooks nowadays as the paradigm of transcriptional control. The model describes genetic material organized in units of activity, the operons (Figure 1). A regulator gene *I* produces a repressor protein that upon interaction with its receptor, the operator *O*, prevents transcription of genes *Z*, *Y*, and *A*, which encode proteins that are involved in lactose metabolism. In the presence of lactose (inducer), binding of the repressor to the operator is prevented by interaction of the repressor protein with allolactose (metabolized from lactose by transglycosylation). Consequently, RNA polymerase can bind to the promoter *P* next to the operator region to start transcription of the following genes. In agreement with this model, it has been widely accepted that cells are controlled predominantly at the DNA level by regulatory proteins with the capacity to recognize individual genes and to adjust-

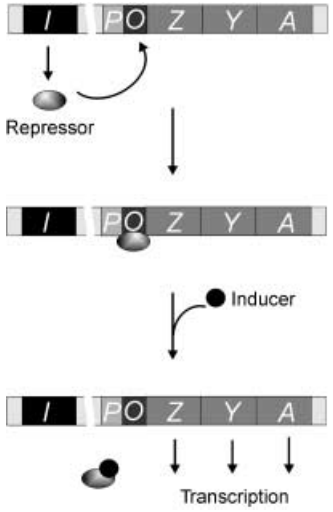
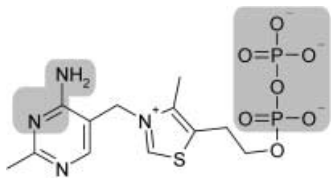


Figure 1. *Expression of the* lac *operon. In the absence of inducer, the repressor produced by the* I *gene binds to the operator (O) and thus prevents transcription of the structural genes* Z, Y, *and* A *encoding lactose-metabolizing enzymes. Binding of the inducer leads to dissociation of the repressor from the operator. The promotor (P) becomes accessible to RNA polymerase so that transcription of* Z, Y, *and* A *can occur, followed by synthesis of the corresponding proteins.*

their activity. Nevertheless, in 1965 Francois Jacob already expected that there might be room for other control mechanisms: "The simplest hypothesis compatible with the results of genetic analysis ··· is that the promoter represents the punctuation of transcription, providing the signal for RNA polymerase to start the synthesis of the messenger for this operon on one of the two DNA strands. If this is correct, the operator is not transcribed into messenger and repression can be exerted only at the level of DNA. This is the interpretation that now seems the most plausible to the geneticists; but it is clear that, as usual, the last word will belong to the chemist."[2] Indeed, important new experiments concerning this matter have recently been reported. Ronald Breaker and co-workers have now demonstrated that bacterial gene expression can be directly regulated at the level of mRNA without the participation of proteins.[3] That genetic control proceeds not only at the transcriptional level as described above, but also at the translational level has been known for some years. Each mRNA contains protein-coding stretches as well as noncoding regions used to control translation of these RNAs into proteins. However, in all examples known to date, metabolite-sensing proteins bind to the control regions to alter the conformation of the mRNA and thus to regulate the access of the ribosome for initiation of translation. The group led by Ronald Breaker has now provided the first examples of translational control by direct RNA – metabolite interaction.[3] In a recent work, Winkler et al.[3a] showed that mRNAs encoding enzymes involved in thiamine (vitamin B_1) biosynthesis in *Escherichia coli* are able to directly bind thiamine or its phosphate (TP) and pyrophosphate (TPP) derivatives (Scheme 1) and that the resulting mRNA – metabolite complex



Scheme 1. *Chemical structure of thiamine pyrophosphate. Sites that make contact with RNA are marked.*

[a] *Prof. Dr. S. Müller*
Ruhr-Universität Bochum, Fakultät für Chemie
Universitätsstrasse 150
44780 Bochum (Germany)
Fax: (+49) 234-3214355
E-mail: sabine.mueller@orch.ruhr-uni-bochum.de

 DOI: 10.1002/cbic.200300640

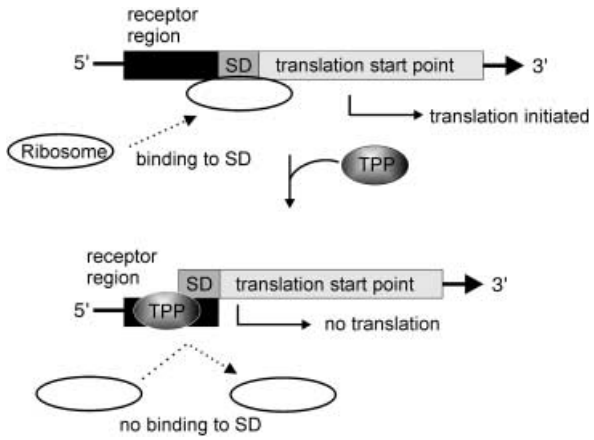


Figure 2. *Thiamine-pyrophosphate-sensing mRNA. In the absence of TPP, translation of the RNA downstream of the receptor region takes place. When TPP is present, it binds to the receptor region and sequesters the Shine – Dalgarno (SD) element, which leads to inhibition of translation.*

adopts a distinctly different conformation relative to the free mRNA. This conformational change leads to sequestering of the ribosome binding site (Shine – Dalgarno sequence) and in consequence to reduction of gene expression (Figure 2).

It has been proposed previously that vitamin biosynthesis might be regulated by direct interaction of the respective mRNA with its cognate metabolite.[4] In searching for more direct evidence of mRNA – metabolite interaction, Winkler et al. synthesized the untranslated leader sequences of two *E. coli* mRNAs important for vitamin B_1 synthesis, *thiM* and *thiC*. Structure probing with hydroxyl radicals in the presence and absence of thiamine, TP, and TPP revealed that both mRNAs undergo substantial structural alteration upon addition of the metabolite. Comparing the binding properties of thiamine, thiamine phosphate, and thiamine pyrophosphate revealed that TPP is bound more readily than thiamine or TP, with up to more than 1000-fold discrimination. By examination of several thiamine analogues, important contact sites of TPP for specific binding to the RNA could be identified (Scheme 1). Furthermore, the pattern obtained after hydroxyl-radical-induced cleavage helped to establish the secondary structure of the investigated mRNAs and showed a great degree of congruence with stem and bulge structures identified by secondary structure folding algorithms. Thus, the authors identified a region of the mRNAs that becomes protected against cleavage upon addition of TPP. The obtained results also showed that both mRNAs, *thiC* and *thiM*, are folded into similar secondary structures containing several common sequence elements. Furthermore, a different region corresponding to the Shine – Dalgarno sequence, which is required for ribosome binding and translation initiation, was accessible in the metabolite-free samples but became protected in the presence of TPP. These results led the authors to suggest that there exists a specific TPP binding motif and that gene expression is controlled by binding of TPP to this motif and the consequent switching of mRNA conformations such that the ribosome cannot gain access to the Shine – Dalgarno sequence (Figure 3). These two predicted conformations with the Shine – Dalgarno sequence freely accessible ("on" state) or, upon addition of the metabolite, buried in a double-stranded mRNA structure ("off" state) were further evaluated by generation of a series of mutants favoring the one or the other conformation. From the obtained data it became evident that TPP binding restricts the structural freedom of the Shine – Dalgarno element in the appropriate RNA variants and that this effect correlates with genetic control.

The described system is one of the first representatives of a genetic control strategy in which mRNA serves as a metabolite – sensing switch. Although several previous papers have proposed that certain messenger RNAs might use allosteric mechanisms to mediate regulatory responses depending on specific metabolites,[4] and thus have anticipated the findings of Breaker and colleagues,[3] the thiamine-pyrophosphate-sensing mRNA is the first true demonstration of a natural "riboswitch". So far, three more examples of gene regulation by "riboswitching"

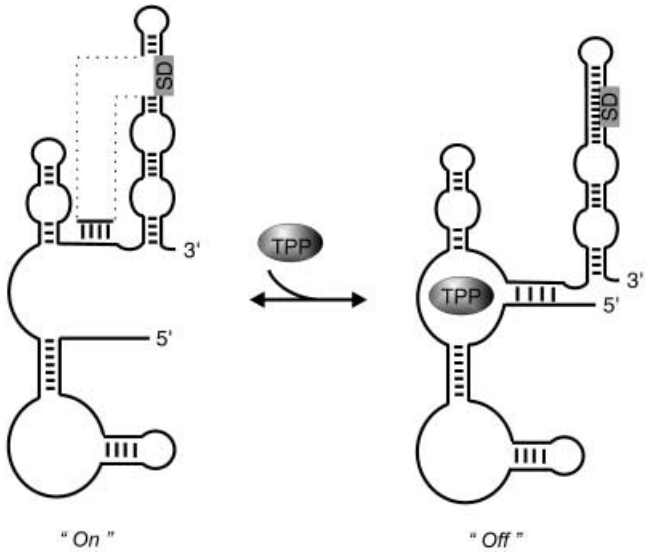


Figure 3. *Schematic representation of the riboswitch in the "on" and "off" states, as suggested by Winkler et al.[3a] The SD element is marked as a grey bar. In the absence of TPP, the SD sequence is single stranded and can interact with the ribosome. Binding of TPP leads to a conformational change upon which the SD element becomes occluded in a double-stranded structure and thus is not accessible to the ribosome. As a result, translation is inhibited.*

have been found: the synthesis of coenzyme B_{12} (5′-deoxy-5′-adenosylcobalamin),[3b, 5] the synthesis of vitamin B_2 (flavin mononucleotide, FMN),[3c, 5] and very recently the control of sulphur metabolism in bacteria.[6] It is fair to assume that the evolutionary origin of this strategy predates the emergence of proteins. The RNAs identified as receptors for metabolites have conserved sequences that are found in certain bacterial mRNAs as well as in several species of archae,[7] which supports the idea that metabolite-binding RNA structures are widespread in nature. All three metabolites, TPP, coenzyme B_{12}, and FMN, have been suggested to have emerged as biological cofactors during the "RNA world."[8] If these cofactors were biosynthesized and used before the advent of proteins, than indeed riboswitches might represent the most ancient form of genetic control.

Beyond this hypothesis, the work of Winkler et al.[3a] once more demonstrates the ability of RNA molecules to combine receptor and functional elements. Over the past years it has become evident that RNA can bind small molecules. Based on this feature, RNAs have been created that serve as highly specific receptors for a variety of organic compounds.[9] Furthermore, ribozymes have been engineered that contain allosteric binding sites and thus are functionally controlled by effector molecules.[10] The TPP riboswitch described by Winkler et al.[3a] represents a naturally occurring allosteric system that works in the classical sense. TPP, the allosteric factor, binds to an aptamer domain of the RNA and induces a signal that is sensed in another domain, the Shine–Dalgarno element. The Shine–Dalgarno blocking mechanism is clearly used to control translation in the *thiM* mRNA. Interestingly, the second investigated mRNA, *thiC*, uses a more complex regulation system that controls gene expression both at the transcriptional and translational level. In the suggested model, TPP binding is converted into a transcription-terminating signal and additionally leads to inhibition of translation of already transcribed mRNAs.

Altogether, the work of Winkler et al.[3a] demonstrates that the many faces of RNA known today are still to be supplemented. We know that RNAs, as aptamers, can bind a wide variety of small molecules; we know that RNA can act as a catalyst and that binding and functional elements can be combined to engineer allosteric ribozymes (reviewed in ref. [11]). After having created all these RNA systems in vitro, we should not be particularly surprised to see that nature has evolved a mechanism for gene regulation that is based on exactly this simple principle: 1) use of metabolite-driven switching between two energetically balanced RNA conformations to sense changes in metabolite concentration and 2) transformation of the induced conformational transitions into an "on"—or "off"—signal for translation of metabolite-synthesizing proteins. Nevertheless, since the days of Monod and Jacob, in our minds genetic regulation has been dominated by proteins. Having been used to this fact for years, we experience the findings of the Breaker laboratory as evidence for an exciting new function of RNA molecules in nature. It remains to see which of its many faces RNA will show to the world next. We await this revelation with pleasant anticipation.

[1] F. Jacob, J. Monod, *J. Mol. Biol*. **1961**, *3*, 318–356.

[2] F. Jacob, Nobel lecture, 11. December 1965.

[3] a) W. Winkler, A. Nahvi, R. R. Breaker, *Nature* **2002**, *419*, 952–956; b) A. Nahvi, N. Sudarsan, M. S. Ebert, X. Zou, K. L. Brown, R. R. Breaker, *Chem. Biol*. **2002**, *9*, 1043–1049; c) W. C. Winkler, S. Cohen-Chalamish, R. R. Breaker, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15908–15913; d) As this manuscript was going to press, two new reports of the Breaker laboratory describing naturally occurring riboswitches appeared: (1) N. Sudarsan, J. F. Barrick, R. R. Breaker, *RNA* **2003**, *9*, 644–647; (2) M. Mandal, B. Boese, J. E. Barrick, W. C. Winkler, R. R. Breaker, *Cell* **2003**, *113*, 577–586.

[4] a) L. Gold, D. Brown, Y. He, T. Shtatland, B. S. Singer, Y. Wu, *Proc. Natl. Acad. Sci. USA* **1997**, *94,* 59–64; b) M. S. Gelfand, A. A. Mironov, J. Jomantas, Y. I. Kozlov, D. A. Perumov, *Trends Genet*. **1999**, *15*, 439–442; c) X. Nou, R. J. Kadner, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7190–7195; d) J. Miranda-Rios, M. Navarro, M. Soberón, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9736–9741; e) G. D. Stormo, Y. Yi, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9465–9467; f) A. G. Vitreschak, D. A. Rodinov, A. A. Mironov, M. S. Gelfand, *Nucleic Acids Res*. **2002**, *30*, 3141–3151.

[5] A. S. Mironov, I. Gusarov, R. Rafikov, L. E. Lopez, K. Shatalin, R. A. Kreneva, D. A. Perumov, E. Nudler, *Cell* **2002**, *111*, 747–756.

[6] V. Epstein, A. S. Mironov, E. Nudler, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5052–5056.

[7] a) M. S. Gelfand, A. A. Mironov, J. Jomantas, Y. I. Kozlov, D. A. Perumov, *Trends Genet*. **1999**, *15*, 439–442; b) J. Miranda-Rios, M. Navarro, M. A. Soberón, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9736–9741.

[8] F. Joyce, *Nature* **2002**, *418*, 214–221.

[9] a) M. Famulok, *Curr. Opin. Struct. Biol*. **1999**, *9*, 324–329; b) T. Hermann, D. J. Patel, *Science* **2000**, *287*, 820–825.

[10] G. A. Soukup, R. R. Breaker, *Curr. Opin. Struct. Biol.* **2000**, *10*, 318–325; b) S. Seetharaman, M. Zivarts, N. Sudersan, R. R. Breaker, *Nat. Biotechnol*. **2001**, *19*, 336–341.

[11] S. K. Silverman, *RNA* **2003**, *9*, 377–383.

Received: April 29, 2003 [H640]

