The Plant Cell, Vol. 19: 3437–3450, November 2007, www.plantcell.org © 2007 American Society of Plant Biologists

Riboswitch Control of Gene Expression in Plants by Splicing and Alternative 3′ End Processing of mRNAs [W][OA]

Andreas Wachter,[a,1] Meral Tunc-Ozdemir,[b] Beth C. Grove,[c] Pamela J. Green,[d] David K. Shintani,[b] and Ronald R. Breaker[a,c,e,2]

[a] Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, Connecticut 06520
[b] Department of Biochemistry and Molecular Biology, University of Nevada, Reno, Nevada 89557
[c] Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520
[d] Department of Plant and Soil Sciences and Delaware Biotechnology Institute, University of Delaware, Newark, Delaware 19711
[e] Howard Hughes Medical Institute, Yale University, New Haven, Connecticut 06520

The most widespread riboswitch class, found in organisms from all three domains of life, is responsive to the vitamin B_1 derivative thiamin pyrophosphate (TPP). We have established that a TPP-sensing riboswitch is present in the 3′ untranslated region (UTR) of the thiamin biosynthetic gene *THIC* of all plant species examined. The *THIC* TPP riboswitch controls the formation of transcripts with alternative 3′ UTR lengths, which affect mRNA accumulation and protein production. We demonstrate that riboswitch-mediated regulation of alternative 3′ end processing is critical for TPP-dependent feedback control of *THIC* expression. Our data reveal a mechanism whereby metabolite-dependent alteration of RNA folding controls splicing and alternative 3′ end processing of mRNAs. These findings highlight the importance of metabolite sensing by riboswitches in plants and further reveal the significance of alternative 3′ end processing as a mechanism of gene control in eukaryotes.

INTRODUCTION

Riboswitches are metabolite-sensing gene control elements typically located in the noncoding portions of mRNAs. Twelve structural classes of riboswitches in bacteria have been characterized to date that sense small organic compounds, including coenzymes, amino acids, and nucleotide bases (Mandal and Breaker, 2004; Soukup and Soukup, 2004; Winkler and Breaker, 2005; Fuchs et al., 2006; Roth et al., 2007) or magnesium ions (Cromie et al., 2006). In most instances, riboswitches can be divided into aptamer and expression platform regions that represent two functionally distinct but usually physically overlapping domains responsible for ligand binding and gene control, respectively.

The complexity of the structures formed by aptamers and their mechanisms of ligand recognition are evident upon examination of the atomic resolution models elucidated by x-ray crystallography for several riboswitch classes, including those that bind guanine and adenine (Batey et al., 2004; Serganov et al., 2004), *S*-adenosylmethionine (Montange and Batey, 2006), thiamin pyrophosphate (TPP) (Edwards and Ferre-D'Amare, 2006; Serganov et al., 2006; Thore et al., 2006), and glucosamine-6-phosphate (Kline and Ferré-D'Amaré, 2006; Cochrane et al., 2007). The nucleotide sequences of the ligand binding core and supporting architectures of each aptamer class are highly conserved among different species as a result of their need to form a precise receptor for a specific ligand. By contrast, the expression platforms for riboswitches can vary considerably among species or even among multiple representatives of a riboswitch class in a single organism.

The high level of aptamer conservation allows researchers to employ bioinformatics methods to identify new riboswitch candidates (e.g., Grundy and Henkin, 1998; Gelfand et al., 1999; Barrick et al., 2004; Corbino et al., 2005; Weinberg et al., 2007) and to determine the distribution of known riboswitch classes in various organisms (e.g., Rodionov et al., 2002; Vitreschak et al., 2003; Nahvi et al., 2004; Abreu-Goodger and Merino, 2005). To date, these searches have revealed that only members of the TPP-sensing riboswitch class are present in all three domains of life (Sudarsan et al., 2003). In eukaryotes, TPP aptamers were found in thiamin metabolic genes from plants and filamentous fungi, but the proposals for the mechanism of riboswitch function remained speculative (Kubodera et al., 2003; Sudarsan et al., 2003). In the fungus *Neurospora crassa*, a TPP aptamer resides in an intron within the 5′ region of *NMT1* mRNA, and recently it has been shown that TPP binding by the aptamer regulates *NMT1* gene expression by controlling alternative splicing (Cheah et al., 2007). Specifically, TPP binding by the riboswitch prevents removal of intron sequences carrying upstream open reading frames (ORFs) that preclude expression of the main ORF.

Herein, we report that TPP riboswitches are present in a variety of plant species where they reside in the 3′ untranslated regions

[1] Current address: Heidelberg Institute for Plant Sciences, University of Heidelberg, 69120 Heidelberg, Germany.
[2] Address correspondence to ronald.breaker@yale.edu.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Ronald R. Breaker (ronald.breaker@yale.edu).
[W]Online version contains Web-only data.
[OA]Open Access articles can be viewed online without a subscription.
www.plantcell.org/cgi/doi/10.1105/tpc.107.053645

(UTRs) of *THIC* genes. Formation of *THIC* transcripts with alternative 3′ UTR lengths is dependent on riboswitch function and mediates feedback regulation of *THIC* expression in response to changes in cellular TPP levels. Our data indicate that 3′ UTR length correlates with transcript accumulation, thereby establishing a basis for gene control by alternative 3′ end processing. We propose a detailed mechanism for TPP riboswitch function in plants, which includes aptamer-mediated control of splicing and differential 3′ end processing of *THIC* mRNAs. This study further reveals the versatility of riboswitch control in organisms from different domains of life and expands our knowledge of previously unknown aspects of eukaryotic gene regulation.

RESULTS AND DISCUSSION

TPP Aptamers Are Widely Distributed in Plant Species

The presence of highly conserved TPP binding aptamers in the 3′ UTRs of the *THIC* genes from the plant species *Arabidopsis thaliana*, *Oryza sativa*, and *Poa secunda* had been reported previously (Sudarsan et al., 2003). The collection of plant TPP aptamer representatives was expanded by sequencing *THIC* genes from additional plant species and by conducting database searches for nucleotide sequences that conform to the TPP aptamer consensus. After cDNA sequences were obtained, the corresponding regions from genomic DNAs of each species were cloned and sequenced (see Methods for details), thus providing the sequences of both the precursor and the processed mRNA molecules.

An alignment of all available TPP aptamer sequences from plants reveals a high level of conservation of nucleotide sequence and a secondary structure consisting of stems P1 through P5 (Figure 1A). The major differences of eukaryotic TPP riboswitch aptamers from plants (Figure 1B) and filamentous fungi (Cheah et al., 2007) compared with their bacterial and archaeal counterparts (Figure 1C) (Rodionov et al., 2002; Winkler et al., 2002) are the consistent absence of a P3a stem frequently present in bacterial representatives and the variable length of the P3 stem in eukaryotes. Neither region is involved directly in TPP binding (Edwards and Ferre-D'Amare, 2006; Serganov et al., 2006; Thore et al., 2006; Cheah et al., 2007); therefore, these differences should not affect ligand binding specificity.

The TPP aptamer is found in the 3′ UTR of all known *THIC* examples from monocots, dicots, and the conifer *Pinus taeda*. Interestingly, in the moss *Physcomitrella patens*, the TPP aptamer is present in the 3′ UTR of *THIC* (Ppa1) and also resides in the 3′ region of two genes that are homologous to the thiamin biosynthetic gene *THI4* (Ppa2 and Ppa3). This latter observation, and the observation that fungi also have TPP aptamers associated with multiple different genes (Cheah et al., 2007), indicate that eukaryotes likely use representatives of the same riboswitch class to control multiple genes in response to changing concentrations of a key metabolite.

A striking characteristic of TPP aptamers from plants is the high level of nucleotide sequence conservation. Approximately 80% of the nucleotides (excluding the P3 stem) are conserved in all plant examples. By contrast, <40% are conserved in filamentous fungi. Most differences among plant TPP aptamers are found in the P3 stem, which varies both in length and sequence. Also, the length of the P3 stem varies among TPP aptamer representatives in the same species, as is observed in *P. patens* (Figure 1A). The presence of both an extended P3 stem in the aptamer present in *THIC* and very short P3 stems in those present in *THI4* suggests that there is no species-specific requirement for this component of the aptamer.

THIC 3′ UTRs Vary in Length and Sequence

We analyzed the nucleotide sequences corresponding to the 3′ regions of *THIC* mRNAs cloned from six plant species or obtained from GenBank (*O. sativa*) (see Supplemental Figure 1 online; see also Methods for details). Interestingly, the exon-intron organization of the 3′ regions of *THIC* genes is conserved among these seven species, and the formation of three major types of RNA transcripts with varying 3′ UTR lengths is always observed (Figure 2A). The stop codon for the *THIC* ORF is commonly followed by an intron that resides near a TPP aptamer. Type I (*THIC*-I) RNAs represent precursors that carry the complete aptamer and can extend to variable lengths at their 3′ ends. *THIC*-I RNAs can be differentially processed either upstream of the aptamer to yield type II (*THIC*-II) RNAs with a shorter 3′ UTR or can be spliced to yield type III (*THIC*-III) RNAs, which are missing the 5′ portion of the TPP aptamer. *THIC*-II and *THIC*-III RNAs are mutually exclusive fates of *THIC*-I processing.

Quantitation of the lengths of various regions (designated 1 through 6) within the *THIC* 3′ UTRs of these species reveals that some regions (2 through 5) exhibit considerable conservation of the numbers of nucleotides bridging key features within the UTR (see Supplemental Figure 2 online). By contrast, the length of the first intron (region 1) and the length of the 3′-most portion of *THIC*-I and *THIC*-III (region 6) are highly variable. For example, *THIC*-I and *THIC*-III can extend by more than 1 kb at their 3′ ends. We speculated that conservation of the distances between certain 3′ UTR features might be important for TPP-mediated gene regulation.

RT-PCR was used to analyze *THIC* transcript types. RT-PCR using a polyT primer (primer b) and a primer specific for the *THIC* ORF (primer a; amplifies all *THIC* transcript types) results predominantly in amplification of *THIC*-II (Figure 2B). This indicates that the short transcript form is most abundant in all species examined. RNA gel blot analysis of RNA from *Arabidopsis* with a probe that binds to the coding region of the *THIC* mRNA also results in one major signal corresponding to the size of *THIC*-II, which further supports our conclusion that *THIC*-II is the dominant transcript type (see further discussion below).

THIC-I and *THIC*-III were detected by RT-PCR using a reverse primer c that is specific for the extended 3′ region in each species examined but that do not recognize *THIC*-II RNAs (Figure 2C). The smallest PCR product band for each species corresponds to *THIC*-III, whereas additional bands represent products derived from *THIC*-I that still retain one or both introns of the 3′ UTR or represent minor splicing variants. RNA gel blot analysis using a probe specific for the 3′ UTR of *THIC*-I and *THIC*-III from *Arabidopsis* confirmed that these transcript types are present in low copy number (see further discussion below) and also revealed heterogeneity of transcript length.

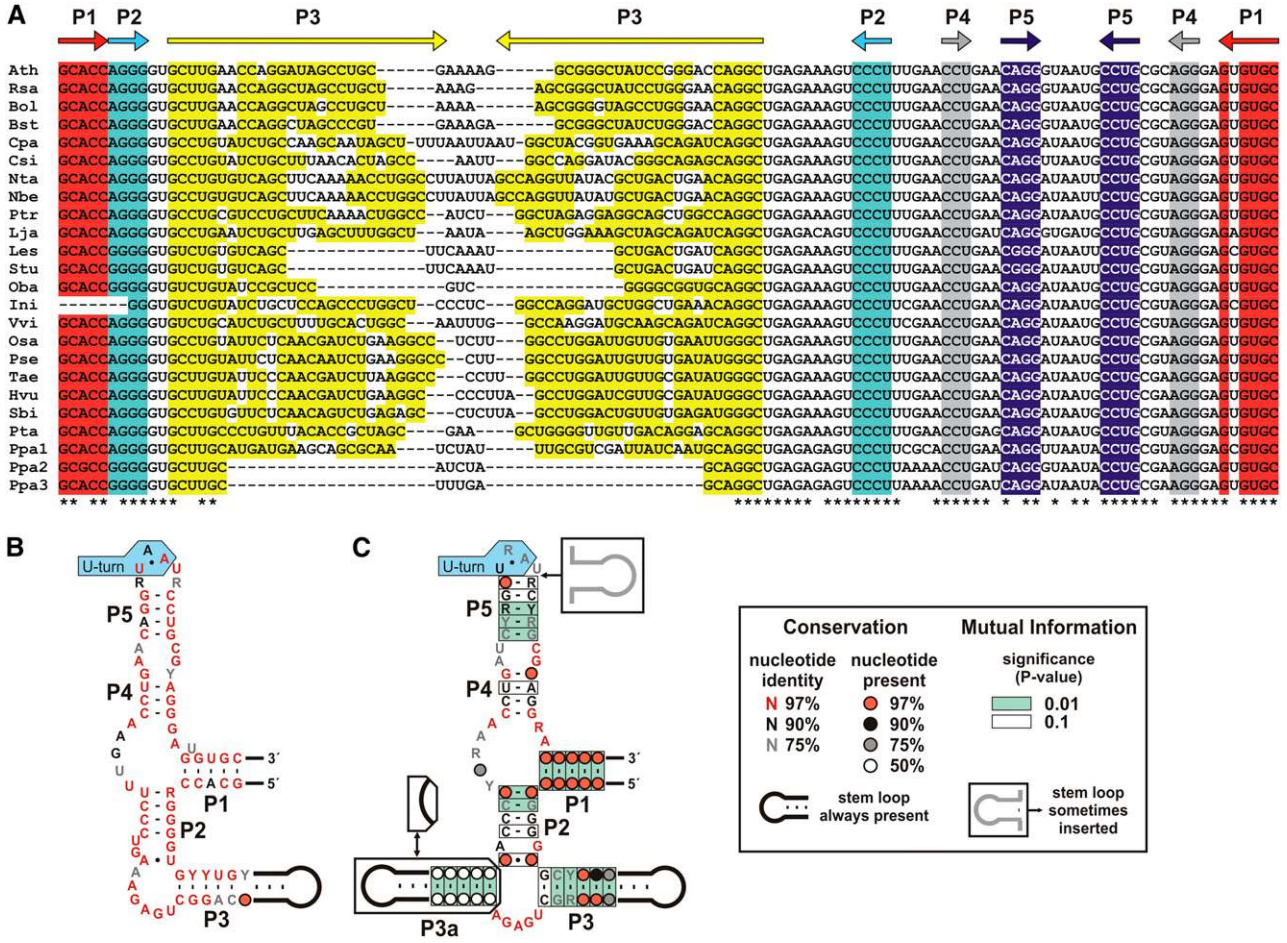

Figure 1. TPP Aptamers Are Conserved and Widespread in Plant Species.

(A) Alignment of TPP aptamer sequences from various plant species reveals high conservation of sequence and structure. Nucleotides forming stems P1 through P5 are highlighted in color, and asterisks identify nucleotides that are conserved between all examples. Sequences are derived from *Arabidopsis* (Ath, NC003071), *Raphanus sativus* (Rsa, EF588038), *Brassica oleracea* (Bol, BH250462), *Boechera stricta* (Bst, DU681973), *Carica papaya* (Cpa, DX471004), *Citrus sinensis* (Csi, DY305604), *Nicotiana tabacum* (Nta, EF588039), *Nicotiana benthamiana* (Nbe, EF588040), *Populus trichocarpa* (Ptr, Joint Genome Initiative, Populus genome, LG_IX: 7897690-7897807), *Lotus japonicus* (Lja, AG247551), *Lycopersicon esculentum* (Les, EF588041), *Solanum tuberosum* (Stu, DN941010), *Ocimum basilicum* (Oba, EF588042), *Ipomoea nil* (Ini, BJ566897), *Vitis vinifera* (Vvi, AM442795), *Oryza sativa* (Osa, NC008396), *Poa secunda* (Pse, AF264021), *Triticum aestivum* (Tae, CD879967), *Hordeum vulgare* (Hvu, BM374959), *Sorghum bicolor* (Sbi, CW250951), *Pinus taeda* (Pta, CCGB, Contig116729 RTDS2_8_E12.g1_A021: 551-686), and *Physcomitrella patens* (Ppa, gnl|ti|856901678, gnl|ti|893553357, gnl|ti|876297717; Lang et al., 2005). The sequence for *I. nil* represents a splice variant derived from cDNA and is therefore lacking the 5′ end of the aptamer.

(B) and **(C)** Consensus sequences and secondary structure models of TPP riboswitch aptamers based on all representatives from plants **(B)** or bacterial and archaeal species **(C)**. The mutual information reflects the probability for the occurrence of the boxed base pairs (Barrick and Breaker, 2007).

To assess whether 3′ end processing differs for the various transcript types in *Arabidopsis*, RT-PCR was conducted using primers that permit separate amplification of the various *THIC* RNAs. Interestingly, the ratio of amplification products corresponding to *THIC*-I and *THIC*-III RNAs differs depending on the primer used for RT (Figure 2D; see also Figures 4E and 5E). Amplification products from *THIC*-III dominate when RT is conducted with polyT primers, whereas amplification products from *THIC*-I are increased relative to *THIC*-III product amounts when either random hexamers or gene-specific primers are used for RT. This indicates that a major fraction of *THIC*-I RNAs is not polyadenylated at the sites used for 3′ end processing of the other *THIC* RNA types. Furthermore, amplification product variants of *THIC*-I with the upstream intron spliced or unspliced were detected. These findings suggest that *THIC*-I represents the unprocessed *THIC* precursor transcript. By contrast, both *THIC*-II and *THIC*-III RNAs are further processed and polyadenylated at the indicated sites, and these RNAs represent the two alternative processing fates of the precursor transcript.

Also, cDNAs generated with primers binding far downstream of the aptamer sequence yielded PCR amplification products (Figure 2D), indicating that *THIC*-I and *THIC*-III can extend >1 kb

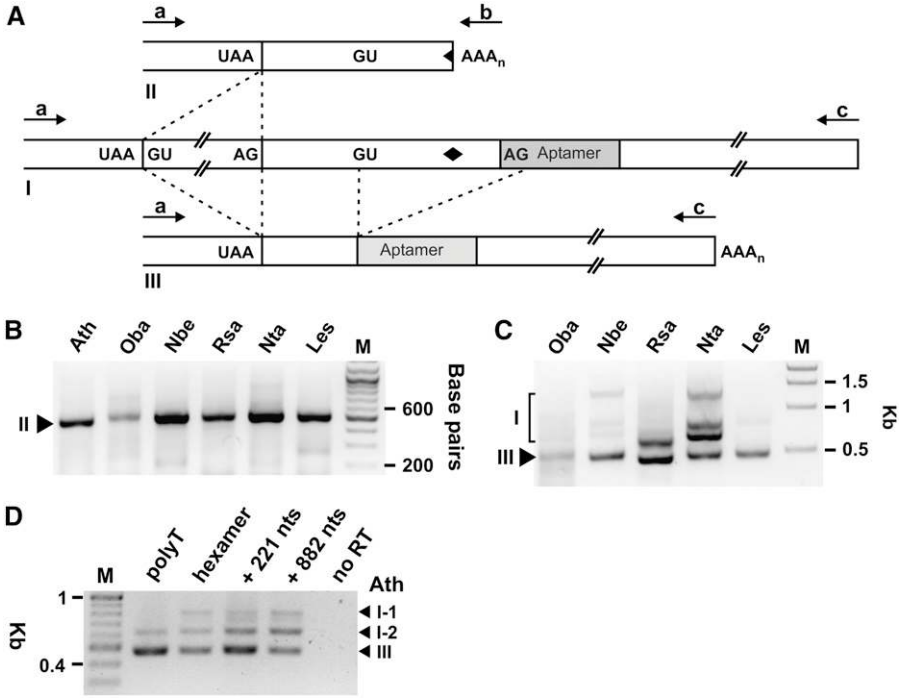

Figure 2. The Exon-Intron Organization of *THIC* 3′ UTRs Is Conserved.

(A) Organization of the 3′ region of *THIC* genes and derived transcript types are similar. The first box represents the last exon of the coding region with the stop codon UAA depicted. The stop codon is followed by an intron (except in *L. esculentum*, where the intron is located immediately in front of the stop codon), which is spliced in transcript types II and III (see **[B]** and **[D]**). GU and AG notations identify 5′ and 3′ splice sites, respectively. Arrows indicate binding sites of primers that were used for RT-PCR amplification of *THIC* 3′ UTRs shown in **(B)** to **(D)**. Dashed lines indicate splicing events, and the diamond represents the transcript processing site.

(B) PCR amplification of *THIC* 3′ UTRs with DNA primers a and b from cDNAs generated with polyT primer yields only type II RNAs in all species examined. RT-PCR products were separated using 1.5% agarose gel electrophoresis and visualized by ethidium bromide staining and UV illumination. M designates the marker lane containing DNAs of 100-bp increments.

(C) RT-PCR analysis was conducted using the same cDNAs as used in **(B)** with primers a and c. Primer c is specific for 3′ UTRs of types I and III RNAs. Kbp designates kilobase pairs.

(D) RT-PCR products of 3′ UTRs from type I and III RNAs from *Arabidopsis* cDNAs generated with different RT primers. Primers used for RT were polyT, random hexamers, or sequence-specific primers that bind near the annotated end of *THIC* (221 nucleotides downstream of the end of the aptamer) or further downstream (882 nucleotides downstream of the end of the aptamer) as indicated. PCR amplification was performed with primers a and c for all cDNAs. I-1 and I-2 represent type I RNAs with the upstream intron following the stop codon unspliced or spliced, respectively. No RT indicates a control reaction using the RNA without reverse transcription as a template source.

downstream of the annotated end of *THIC* in *Arabidopsis*. Full-length cDNA annotations in GenBank (AK068703, AK065235, and AK120238) indicate that comparable *THIC*-III mRNAs with long 3′ UTRs also exist in *O. sativa*.

Thiamin Affects *THIC* Transcript Levels

Quantitative RT-PCR (qRT-PCR) was used to determine whether *THIC* transcript levels respond to increased thiamin concentrations. *Arabidopsis* seedlings were supplemented with various amounts of thiamin, and the different *THIC* transcript types were detected separately using selective primer combinations (see Supplemental Figure 3 online). The primer combination amplifying *THIC*-II also can bind to a subset of *THIC*-I RNAs that have undergone splicing of the first 3′ UTR intron. However, the *THIC*-I–derived amplification products are minor because *THIC*-I transcripts are far less abundant and are almost undetectable when cDNAs are generated with polyT primers (Figure 2D).

After growing seedlings on medium containing 1 mM thiamin, the total amount of *THIC* transcripts decreases to ~20% of that measured when seedlings are grown without thiamin supplementation (Figure 3A). *THIC*-II transcripts exhibit an equivalent reduction, but both *THIC*-I and *THIC*-III transcripts show little or no change in copy number. Again, the simultaneous decrease of total and *THIC*-II transcripts and concomitantly unchanged levels of *THIC*-I and *THIC*-III demonstrate that *THIC*-II is the most abundant transcript form. RNA gel blot analysis of the same samples was used to confirm that *THIC*-II levels decrease and that *THIC*-I and *THIC*-III RNA levels remain relatively unchanged (Figure 3B).

The time interval in which thiamin-dependent changes in transcript levels occur was assessed by performing qRT-PCR

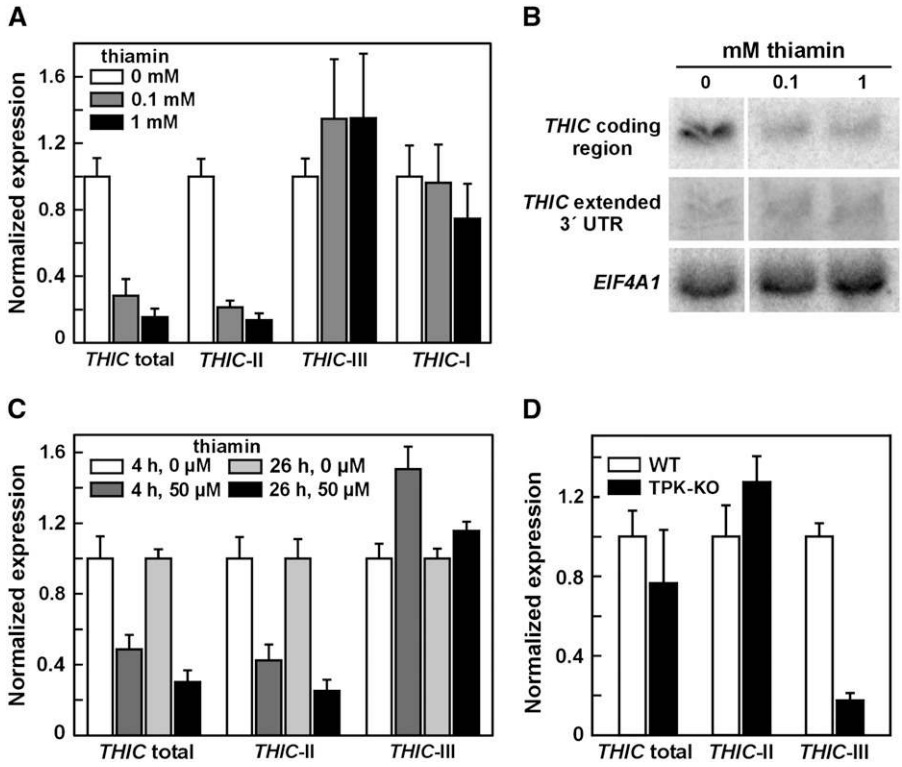

Figure 3. *THIC* Transcript Types Respond Differently to Changes in Thiamin Levels in *Arabidopsis*.

(A) qRT-PCR analysis was conducted on *THIC* transcripts from *Arabidopsis* seedlings grown for 14 d on medium supplemented with 0, 0.1, and 1 mM thiamin. Total *THIC* transcripts and types I, II, and III RNAs were separately detected using different primer combinations. cDNAs were generated using a polyT primer or random hexamers for detection of type I RNAs. Expression was normalized for each primer combination to the value measured using medium without thiamin (open bars). Values are averages from three independent experiments, and error bars represent SD.

(B) RNA gel blot analysis of *THIC* transcripts from the same samples described in **(A)**. Twenty micrograms of total RNA was loaded per lane and analyzed using probes binding to the coding region of *THIC*, the extended 3′ UTR of types I and III RNAs, or the control transcript *EIF4A1*. The signal of *THIC* probes are shown in the size range between 2 and 3 kb. The 3′ UTR probe resulted in weak signals, and exposure time was extended to 3 d compared with 1 d of exposure for the other probes.

(C) qRT-PCR analysis of the time-dependent effects of thiamin treatment on *THIC* transcripts from *Arabidopsis*. Seedlings were grown for 14 d on thiamin-free medium and subsequently sprayed with 50 μM thiamin and 0.25 mg mL^{-1} Tween 80. Control seedlings were treated with a solution containing only Tween 80. Samples were collected after 4 and 26 h and subjected to qRT-PCR analysis. Amounts of *THIC* transcripts were analyzed from cDNAs generated with polyT primer and normalized to the values of the control samples without application of thiamin (open bars). Values are averages from three independent experiments, and error bars represent SD.

(D) Relative changes of the levels of *THIC* transcript types in wild-type and TPK-KO *Arabidopsis* plants. Seedlings were grown for 12 d on thiamin-free medium, and amounts of *THIC* transcript types were analyzed by qRT-PCR. Data were normalized to the values for the wild-type samples and reflect averages from three replicates, with error bars representing SD.

of *THIC* transcripts from RNA isolated at several time points after spraying *Arabidopsis* seedlings with a thiamin solution (Figure 3C). Four hours after thiamin application, total *THIC* RNA and *THIC*-II amounts were reduced to 50% of that measured in the absence of added thiamin. After 26 h, these levels were decreased even further. Interestingly, the modest increase in *THIC*-III observed when thiamin is added to the medium (Figure 3A) is more pronounced in the early phase of the response. Because the different transcript types show opposite responses to thiamin treatment, the control mechanism most likely involves RNA 3′ end processing, and it is unlikely that the feedback mechanism acts at the level of promoter regulation. Indeed, expression of a reporter gene driven by the *THIC* promoter from *Arabidopsis* in transgenic lines was not altered after thiamin supplementation (see Supplemental Figure 4 online).

Most of the thiamin taken up by cells is expected to be converted to TPP by successive phosphorylation reactions to yield concentrations of this coenzyme that are much higher than the concentration of the unphosphorylated vitamin (I. Ajjawi, M. Rodriguez Milla, J. Cushman, and D. Shintani, unpublished data). Therefore, the observed reduction in total *THIC* RNA levels most probably reflects a riboswitch-mediated response to increased TPP concentration, given that TPP binding to plant aptamers is known to occur (Sudarsan et al., 2003; Thore et al., 2006). Furthermore, we speculated that decreased TPP concentrations should result in an opposite response in the levels of *THIC* transcripts.

To examine this possibility, we compared *THIC* expression in wild-type *Arabidopsis* plants versus those carrying a double knockout of thiamin pyrophosphokinase (TPK). These mutants are deficient in both TPK isoforms present in *Arabidopsis* and therefore cannot convert thiamin to TPP (I. Ajjawi, M. Rodriguez Milla, J. Cushman, and D. Shintani, unpublished data). It has been shown that TPK double knockout (TPK-KO) plants largely deplete the TPP stored in seeds within 2 weeks of germination and that the plants depend on TPP supplementation to complete their life cycle (I. Ajjawi, M. Rodriguez Milla, J. Cushman, and D. Shintani, unpublished data). Interestingly, qRT-PCR analysis of *THIC* RNAs from 12-d-old TPK-KO seedlings revealed a pronounced reduction of *THIC*-III compared with the wild type (Figure 3D), indicating that the formation of this RNA is triggered by the presence of TPP. However, total *THIC* and *THIC*-II RNAs do not change significantly, which might be explained by the fact that *THIC* expression is already close to its maximum under the conditions tested.

Riboswitch Function in Thiamin Feedback Response

The presence of different *THIC* RNA types and their changes in abundance in response to varying thiamin levels suggest that the TPP aptamer might affect RNA 3′ end processing and thereby control gene expression. This hypothesis was explored by analyzing the expression of reporter constructs containing enhanced green fluorescent protein (*EGFP*) fused with the complete genomic 3′ region of *THIC* (~2.2 kb downstream of the stop codon) in stably transformed *Arabidopsis* plants. Thiamin application resulted in decreased EGFP fluorescence in leaves from the rosette stage (Figures 4A and 4B). Using qRT-PCR analysis, we found that total amounts of both *EGFP* and endogenous *THIC* transcripts were reduced to ~20% of control levels after thiamin feeding (Figure 4C), which is similar to the extent of repression observed for *Arabidopsis* seedlings (Figure 3). Total *THIC* and *EGFP* transcripts were analyzed using primers binding to the coding regions of the respective cDNAs.

The 3′ UTR sequences of *EGFP* fusion and *THIC* transcripts from the transformants were amplified by RT-PCR (Figures 4D and 4E), cloned, and sequenced. Sequence analyses confirmed the formation of equivalent transcript processing types for *EGFP* and *THIC* (see also Figure 2). The difference in total transcript amount of *THIC* and *EGFP* can be explained by the action of the strong promoter used for control of the transgene. Because the reporter gene construct and the endogenous *THIC* locus yielded similar responses to thiamin and similar processed RNAs, we conclude that no additional sequences upstream of the region fused to *EGFP* are involved in the gene control mechanism.

To determine whether the effects of thiamin regulation are mediated through a TPP riboswitch, we introduced mutations M1, M2, and M3 into the aptamer (Figure 5A) that reduce TPP binding affinity. M1 and M3 mutations interfere with formation of stems P5 and P2 of the TPP aptamer, respectively. With M2, three nucleotides that are known to be involved in direct interactions with the pyrimidine moiety of TPP (Edwards and Ferre-D'Amare, 2006; Serganov et al., 2006; Thore et al., 2006) are mutated. The 3′ regions of *THIC* carrying these variants were fused to *EGFP* and stably transformed into *Arabidopsis* plants.

As expected, plants containing reporter gene constructs carrying the mutant aptamers exhibit either reduced (M1) or a complete loss (M2 and M3) of responsiveness to thiamin application compared with plants containing the wild-type construct (Figure 5B). These findings were confirmed by measuring the relative levels of transcripts using qRT-PCR (Figure 5C). In addition, a reporter construct variant of M3 containing compensatory mutations that restore formation of P2 (M4) exhibits activity similar to the wild type (Figures 5B and 5C). These results indicate that TPP binding by the aptamer is essential for mediating the response to changing TPP levels in the cell. However, the modest thiamin responsiveness exhibited by the M1 construct suggests that this mutant might affect riboswitch function other than just by diminishing the affinity of the aptamer for TPP (see further discussion below).

RT-PCR analyses of 3′ ends of the mRNAs generated from the *EGFP*-riboswitch fusions reveal that the mutant constructs maintain a high level of expression of type II RNAs (Figure 5D), as is typical of wild-type constructs. However, two major differences in type I and III RNAs between mutant and wild-type riboswitches are evident. First, the amount of type III RNA is substantially reduced in the M1 construct and was not detected from the M2 construct (Figure 5E). Second, a considerable decrease of transcripts extending far downstream of the aptamer was observed for both mutants (Figure 5E, 882 nts lane; see also the wild type in Figure 4E). These results reveal that proper riboswitch function is required for the accumulation of type III RNAs with extended 3′ UTR length.

3′ UTR Length Defines Gene Expression Levels

After demonstrating that the riboswitch controls the formation of different 3′ UTRs of RNAs, we sought to determine how the 3′ UTRs affect gene expression. The major differences between *THIC*-II and *THIC*-III 3′ UTRs are that the *THIC*-II RNA is shorter and ends in front of the TPP aptamer. By contrast, *THIC*-III includes most of the consensus aptamer sequence, but the first seven nucleotides at the 5′ end are removed due to splicing of the second intron in the 3′ UTR and are replaced with different nucleotides (Figure 6A, shaded sequence). In-line probing (Soukup and Breaker, 1999) was used to determine whether this altered aptamer retains TPP binding activity. This assay reveals structural changes of RNAs by monitoring altered patterns of spontaneous RNA degradation upon metabolite binding. Briefly, RNA linkages in unstructured portions of the molecule are generally more susceptible to spontaneous cleavage by internal phosphoester transfer than are linkages in structured portions of an RNA. Therefore, structural changes brought about by ligand binding to RNA can be quantitated, mapped, and used to establish the binding characteristics of aptamers. In-line probing has been used previously to analyze TPP aptamers (Winkler et al., 2002; Sudarsan et al., 2003).

The apparent K_d of the altered aptamer for TPP is ~60 μM (Figures 6B and 6C) compared with ~50 nM for the full-length TPP aptamer from *Arabidopsis* (Sudarsan et al., 2003), which is a loss of more than three orders of magnitude in ligand binding affinity. Furthermore, no detectable binding of thiamin occurs with the altered aptamer (data not shown), and it is unlikely that

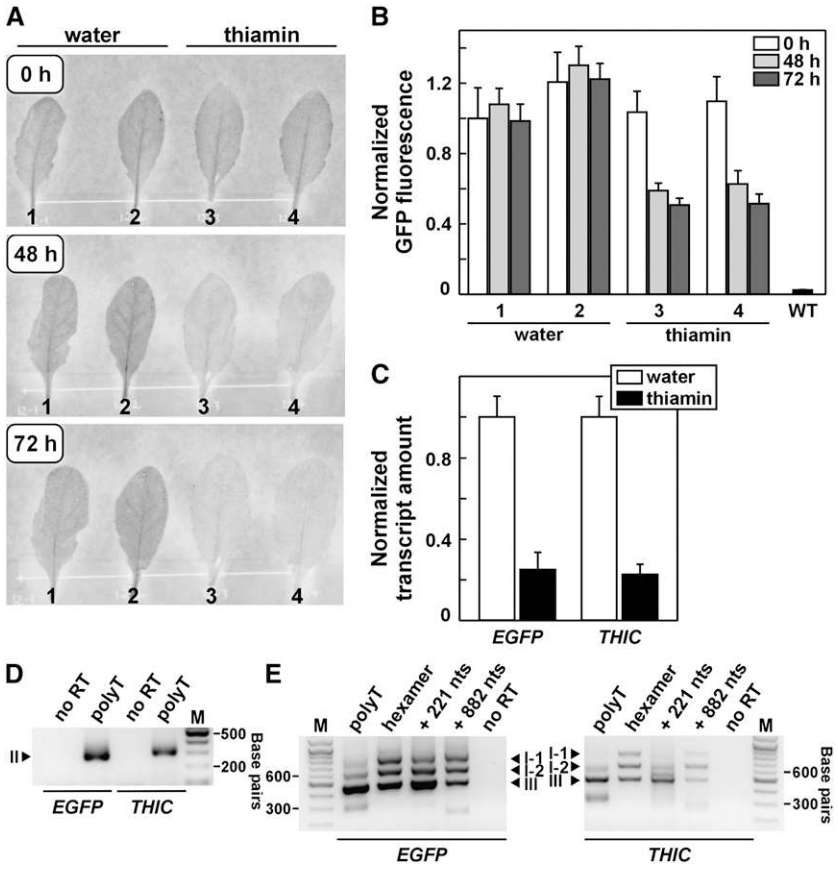

Figure 4. In Vivo Analysis of Riboswitch Function.

(A) Leaves from stably transformed *Arabidopsis* lines expressing a reporter fusion of the complete 3′ region of *AtTHIC* fused to the 3′ end of *EGFP* were abscised and incubated with the petioles in water or in water supplemented with 0.02% thiamin. EGFP fluorescence was assessed at 0, 48, and 72 h after onset of treatment. One representative set of data from three repeats is shown, and the numbers identify different leaves from one transgenic line.
(B) Quantitation of EGFP fluorescence of leaves depicted in **(A)** at three time points. The data represent average fluorescence intensity and SD for each leaf. The plot also depicts average background fluorescence of wild-type leaves.
(C) qRT-PCR analysis of total *EGFP* and *THIC* transcripts from leaves incubated for 72 h in water or 0.02% thiamin. Transcript amounts were standardized to an internal reference transcript and normalized to transcript abundance in water-treated samples. Values are averages from four independent experiments using different transgenic lines, and error bars represent SD.
(D) and **(E)** RT-PCR analysis of different 3′ UTRs of *EGFP* and *THIC* transcript types from *Arabidopsis* reporter transformants grown in the absence of exogenous thiamin. For cDNA generation, a polyT primer, random hexamers, or two different gene-specific primers (binding either 221 or 882 nucleotides downstream of the end of the aptamer) were used as indicated. The forward primers (equivalent to primer a in Figure 2A) were specific for the end of the last exon of the coding region of *EGFP* (left) or *THIC* (right), whereas the reverse primer was either a polyT primer (**[D]**; equivalent to primer b in Figure 2A) or homologous to a region 221 nucleotides downstream of the end of the aptamer (**[E]**; equivalent to primer c in Figure 2A). RT-PCR products were separated and visualized as described in the legend to Figure 2. M designates the marker lanes containing DNAs of 100-bp increments. No RT indicates a control reaction using the RNA without reverse transcription as a template source. I-1 and I-2 represent type I RNAs with the upstream intron following the stop codon unspliced or spliced, respectively. The lowest band in the polyT reaction in **(E)** results from amplification of *THIC* type II RNAs with polyT primer remaining from the RT reaction. Additional unmarked bands correspond to nonspecific amplification as confirmed by cloning and sequencing of all RT-PCR products.

other thiamin derivatives can bind because the region of the aptamer that is exchanged upon splicing is not directly involved in ligand recognition (Edwards and Ferre-D'Amare, 2006; Serganov et al., 2006; Thore et al., 2006). These findings indicate that, once splicing of the second intron of the 3′ UTR occurs, the altered version of the TPP aptamer in *THIC*-III is no longer functional.

To assess possible other effects of the two major *THIC* 3′ UTR forms on gene expression, the 3′ UTR sequences from *THIC*-II (188 nucleotides) and *THIC*-III (408 nucleotides) from *Arabidopsis* were fused to the coding region of luciferase (LUC), and these constructs were expressed in plants under control of the constitutive cauliflower mosaic virus 35S promoter and octopine synthase (OCS) terminator elements (see Supplemental Figure 5A online). *THIC*-III can extend to a variable length at the 3′ end, but the most abundant, shortest version (corresponding to GenBank accession number NM179804) was used for our

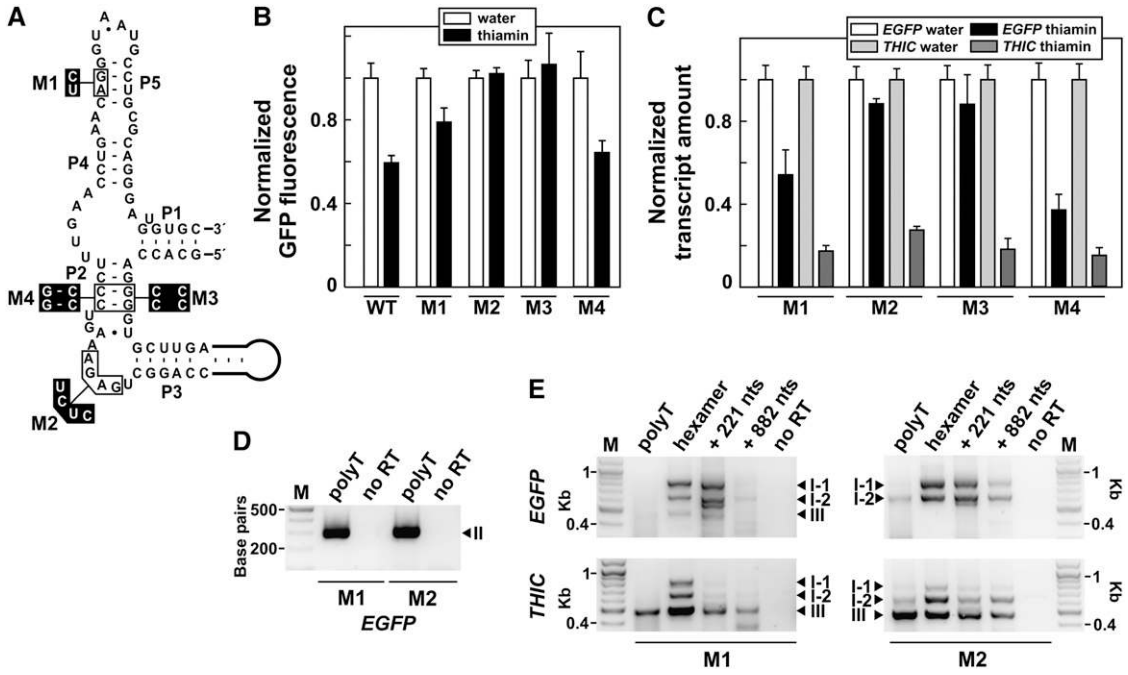

Figure 5. Effects of Aptamer Mutations on Riboswitch Function.

(A) Secondary structure model and sequence of the wild-type TPP aptamer from *Arabidopsis* located in the 3′ region of *THIC* that was fused to *EGFP*. Black boxed nucleotides were altered as indicated to generate mutants M1, M2, and M3 with impaired TPP binding and the compensatory mutant M4.
(B) Quantitation of EGFP fluorescence in leaves from *Arabidopsis* transformants expressing reporter constructs containing the wild-type aptamer sequence or mutated versions M1, M2, M3, and M4. Leaves were excised and incubated with their petioles in water or 0.02% thiamin for 72 h before fluorescence analysis. Values are averages from at least three independent experiments using different transgenic lines. Error bars represent SD.
(C) qRT-PCR analyses of *EGFP* and *THIC* transcript amounts in *Arabidopsis* transformants. Thiamin treatment was performed as described in **(B)**. Transcript amounts (standardized using a reference transcript) were normalized to transcript abundance in water-treated samples. Values are averages from two to four independent experiments using different transgenic lines. Error bars represent SD.
(D) and **(E)** RT-PCR analyses of 3′ UTRs of *EGFP* and *THIC* transcripts from *Arabidopsis* transformants with mutations M1 or M2. RT-PCR analyses were performed as described in the legends to Figures 4D and 4E. Forward primers were homologous to the end of the last exon of the coding region of *EGFP* or *THIC*, and the reverse primer was a polyT primer **(D)** or complementary to a region 221 nucleotides downstream of the end of the aptamer **(E)**. Kbp designates kilobase pairs.

expression analyses. Using RT-PCR, we confirmed that reporter transcripts can extend to and presumably end within the OCS terminator region (see Supplemental Figures 5B and 5C online).

A fusion construct containing the 3′ UTR from *THIC*-III resulted in only ~10% of the LUC activity compared with a construct carrying the 3′ UTR from *THIC*-II (Figure 6D). The possible involvement of the altered TPP aptamer in the type III construct was ruled out by introducing mutations M1 and M5 that completely abolish TPP binding but do not derepress LUC expression. Also, using the reverse complement sequence of the *THIC*-III 3′ UTR sequence did not change LUC activity significantly. These data indicate that the extended length, and not the altered TPP aptamer, plays a role in the repression of gene expression from constructs containing the 3′ UTR from type III RNAs. Equivalent results were obtained with constructs containing the reporter gene *EGFP* in place of *LUC*, and coexpression of the silencing suppressor P19 excluded the possibility that the observed differences are due to silencing effects in the reporter system (see Supplemental Figure 6 online).

We assessed whether differences in reporter activity are also reflected in transcript amounts. Using qRT-PCR, the relative amounts of reporter transcripts containing the 3′ UTRs from *THIC*-II or *THIC*-III from either *Arabidopsis* or *Nicotiana benthamiana* were determined (Figure 6E). Constructs carrying the long 3′ UTR of type III RNAs from both species were present in lower abundance compared with those that carried the short type II 3′ UTR. Since all reporter constructs were expressed under control of identical, constitutive promoters and terminators, the extents of transcription initiation and termination should be the same for all constructs.

Our findings that reporter transcripts containing the long *THIC*-III 3′ UTR are present at lower steady state levels than the respective fusions with the short *THIC*-II 3′ UTR and also that *THIC*-III RNAs do not accumulate significantly (Figures 2 and 3) suggest that long 3′ UTRs might cause increased transcript turnover. Thus, riboswitch-mediated redirection of 3′ end processing to favor the production of mRNAs with extended 3′ UTRs should reduce *THIC* expression. One possible mechanism for

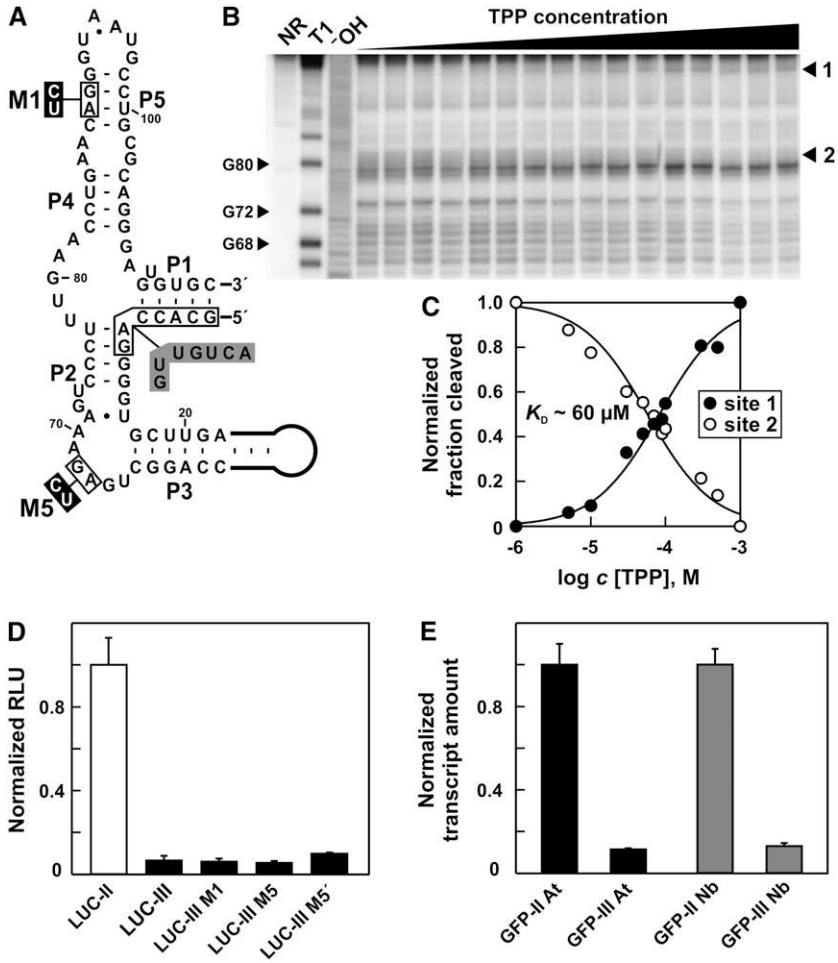

Figure 6. The Long 3′ UTR of *THIC* Causes Reduced Gene Expression Independent of Aptamer Function.

(A) Secondary structure model of the TPP aptamer generated after splicing in *THIC* type III RNA from *Arabidopsis*. Gray shaded nucleotides in stems P1 and P2 identify nucleobase changes compared with the original unspliced aptamer. Black boxed nucleotides were altered as shown to generate mutants M1 and M5 that do not bind TPP.

(B) In-line probing analysis of TPP binding by the spliced aptamer depicted in **(A)**. Lanes include RNAs loaded after no reaction (NR), after partial digestion with RNase T1 (T1), or after partial digestion with alkali ($^{-}$OH). Sites 1 and 2 were quantified to establish the K_d as shown in **(C)**.

(C) Plot indicating the normalized fraction of RNA spontaneously cleaved versus the concentration of TPP for sites 1 and 2 in **(B)**.

(D) In vivo expression analysis of reporter constructs containing the 3′ UTR of *Arabidopsis* type II or III RNAs fused to the 3′ end of the coding region of firefly luciferase (*LUC*). Constructs M1 and M5 are based on the 3′ UTR of type III RNAs but contain the mutations shown in **(A)**. LUC-III M5′ contains the inverted 3′ UTR sequence of construct LUC-III M5. Reporter constructs were analyzed in a transient *N. benthamiana* expression assay and values standardized to a coexpressed luciferase gene from *Renilla*. Expression was normalized to the fusion construct containing the 3′ UTR of type II RNA. Data shown are mean values of three independent experiments, and the error bars represent SD.

(E) qRT-PCR analysis of *EGFP* reporter fusions that contain the 3′ UTRs of *THIC* type II or III RNAs from either *Arabidopsis* (At) or *N. benthamiana* (Nb) after expression in a transient expression assay. Expression was standardized to a coexpressed *DsRED* reporter gene and normalized to the constructs containing a type II 3′ UTR. Data shown are mean values of two representative experiments, and the error bars reflect SD.

the increased transcript turnover of RNAs with long 3′ UTRs might be nonsense-mediated decay (NMD) as previous studies have shown that long 3′ UTRs induce NMD in yeast (Muhlrad and Parker, 1999) and plants (Kertesz et al., 2006). In the latter study, a reduction in the abundance of mRNAs with 3′ UTR lengths >200 nucleotides was observed, as was a correlation between 3′ UTR length and NMD efficiency. However, the extended 3′ UTR could also affect efficiency of 3′ end processing at downstream sites, and future work will be necessary to unravel the underlying mechanism of the decreased transcript accumulation that we observed.

Mechanism of Riboswitch Function

The findings described above demonstrate that the riboswitch affects gene expression by controlling 3′ UTR processing to yield

either type II or type III RNA. We investigated the mechanism by which the aptamer regulates these processing events in a TPP-dependent manner using in-line probing to identify structural changes in the *THIC* 3′ UTR from *Arabidopsis*. An aptamer construct that included 14 nucleotides upstream of the 5′ splice site for the second 3′ UTR intron exhibited TPP-dependent structural modulation of 6 nucleotides immediately upstream of the splice site and the two nucleotides at the splice site junction (Figure 7A). Specifically, TPP addition causes an increase in structural flexibility of the nucleotides near the 5′ splice site. Thus, ligand binding could increase accessibility of the splice site to the spliceosome, thereby permitting the removal of this intron.

We searched for base-pairing potential between the sequences of the modulating 5′ splice site nucleotides and the aptamer nucleotides of *THIC* genes from several plant species. In all species examined, sequences on the 5′ side of the P4-P5 stems are complementary to the sequences immediately upstream (and sometimes inclusive) of the 5′ splice site (Figure 7B). This conservation of base-pairing potential suggests that the riboswitch controls splicing by the mutually exclusive formation of structures that either mask the 5′ splice site under low TPP concentrations or expose the splice site under high TPP concentrations (Figure 7C).

This model is consistent with the in vitro and in vivo data generated in this study, including the partial thiamin responsiveness observed with the M1 variant. M1 carries two mutations that disrupt the P5 stem of the aptamer (Figure 5A), which should weaken its interaction with TPP and disrupt thiamin responsiveness. However, these mutations also weaken base pairing with the 5′ splice site region, which might allow TPP binding to compete effectively with this alternative pairing, despite the expected reduction in TPP affinity.

Our findings indicate that TPP binding to the riboswitch makes accessible a 5′ splice site in the 3′ UTR; subsequently, splicing results in the formation of type III RNAs. To determine if additional effects of TPP binding might contribute to regulation via the 3′ splice site, we mutated the 3′ splice site in the P2 stem of the aptamer using nucleotide changes that retained the structure and TPP binding activity of the RNA (see Supplemental Figure 7A online). Interestingly, plants containing a reporter gene construct

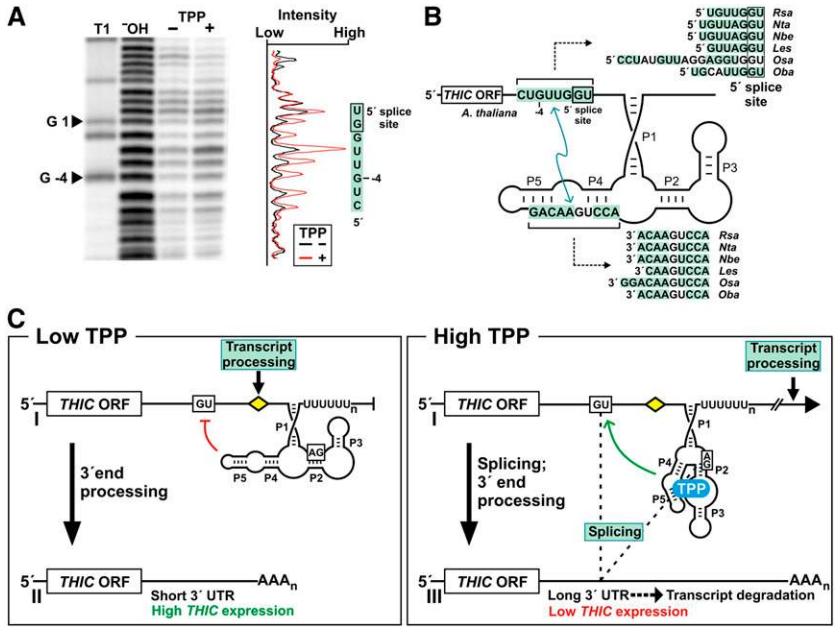

Figure 7. Mechanism of Riboswitch Function in Plants.

(A) TPP causes changes in RNA structure near to the 5′ splice site, which is important for the formation of *THIC* type III RNA. For in-line probing, a 5′ ^{32}P-labeled RNA starting 14 nucleotides upstream of the 5′ splice site (+1) and extending to the 3′ end of the TPP aptamer (nucleotides −14 to 261) from *Arabidopsis* was incubated in the absence (–) or presence (+) of 10 μM TPP, and the resulting spontaneous cleavage products were separated by PAGE. Markers are RNAs partially digested with RNase T1 (T1) or alkali (⁻OH). The graph depicts the relative band intensities in the lanes indicated.
(B) Base-pairing potential between the 5′ splice site region and the P4-P5 stems of the TPP aptamer from *Arabidopsis* (complementary nucleotides are shaded). Stretches of complementary nucleotides are also present in all other plant *THIC* mRNA sequences available.
(C) A model for *THIC* TPP riboswitch function in plants includes control of splicing and alternative 3′ end processing of transcripts. When TPP concentrations are low (left), portions of stems P4 and P5 interact with the 5′ splice site and thereby prevent splicing. The transcript processing site located between the 5′ splice site and the TPP aptamer is retained, and its use results in formation of transcripts with short 3′ UTRs that permit high expression. In the presence of elevated TPP concentrations (right), TPP binds to the aptamer cotranscriptionally, which leads to a structural change that prevents interaction with the 5′ splice site. Splicing occurs and removes the transcript processing site. Transcription continues and alternative processing sites in the extended 3′ UTR give rise to *THIC* type III RNAs. The long 3′ UTRs lead to increased RNA turnover, causing reduced expression of *THIC*.

carrying this M6 mutant aptamer still exhibit thiamin responsiveness of reporter expression (see Supplemental Figures 7B and 7C online). Analysis of 3′ UTRs of M6 reporter transcripts by sequencing RT-PCR products revealed that RNAs similar to type III can still be formed, but the M6 RNAs are now spliced from a 3′ splice site located four nucleotides upstream relative to the position in the wild-type aptamer (see Supplemental Figure 7D online). This finding, and the fact that the relative position of the 3′ splice site for formation of *THIC*-III in the TPP aptamer is variable for different species (see Supplemental Figure 1), indicates that the 3′ splice site is most probably not controlled in a TPP-dependent manner.

One remarkable feature of plant TPP riboswitches is that the 5′ splice sites under riboswitch control are located >200 nucleotides upstream of the complementary regions in the TPP aptamers (see Supplemental Figure 2 online). The complex structural organization of the sequences between the complementary regions (see Supplemental Figure 8 online) might be important to facilitate their interaction, which might also be one reason for the conservation of lengths between features of *THIC* 3′ UTRs from various plants.

Interestingly, TPP riboswitches also control alternative splicing of the *NMT1* genes of fungi, in part by forming ligand-modulated base pairing between nucleotides near a 5′ splice site and the P4-P5 region of an unoccupied TPP aptamer (Cheah et al., 2007). In contrast with these eukaryotic examples, bacteria typically use nucleotides in P1 stems to interface with expression platforms located downstream of the aptamer (Winkler et al., 2002; Sudarsan et al., 2005). Given the substantial changes in the structure of TPP aptamers upon ligand binding, it is surprising that only a portion of the P1 and P4-P5 stems are used to control expression platform function in the TPP riboswitches studied to date. One reason for this might be the need for preorganization of certain aptamer substructures to facilitate rapid ligand sensing.

Model for TPP Riboswitch Function in Plants

Based on our findings, we propose a model for TPP riboswitch regulation in plants involving the metabolite-mediated control of splicing and alternative 3′ end processing of mRNA transcripts (Figure 7C). When TPP concentration is low, the aptamer interacts with the 5′ splice site and prevents splicing. This intron carries a major processing site that permits transcript cleavage and polyadenylation. Processing from this site produces *THIC*-II transcripts that carry short 3′ UTRs, which yield high expression of the *THIC* gene.

When TPP concentrations are high, TPP binding to the aptamer prevents pairing to the 5′ splice site. As a result, the 5′ splice site becomes accessible and is used in a splicing event that removes the major processing site. Transcription subsequently can extend up to 1 kb, and the use of cryptic 3′ end processing sites located downstream of the major site gives rise to *THIC*-III RNAs that carry much longer 3′ UTRs. These findings are in agreement with previous functional studies of polyadenylation in plants (e.g., Mogen et al., 1990; Hunt, 1994) showing that mutation of sequence elements involved in 3′ end processing at a certain site result in the usage of weak and cryptic sites that can be widely spaced. The 3′ end processing at these alternative sites might be less efficient, or the long 3′ UTRs themselves might cause increased transcript turnover and subsequently result in reduced *THIC* expression. Interestingly, a polyuridine tract immediately follows the aptamer in all known TPP riboswitch examples in plants (see Supplemental Figure 1 online), and its conservation indicates a possible role in riboswitch control.

Our findings reveal a mechanism for how TPP-sensing riboswitches can control gene expression in plants and how feedback control maintains TPP levels. In addition, this study further expands the known diversity of mechanisms that riboswitches use to regulate gene expression. The TPP riboswitch in *Arabidopsis* harnesses metabolite binding to control RNA splicing, which determines alternative 3′ end processing fate and ultimately defines the steady state levels of mRNAs. The extensive conservation of sequences, structural elements, and spacing between key 3′ UTR features within the *THIC* genes of various plants indicates that this riboswitch mechanism is maintained in diverse plant species. Although it was previously established that alternative polyadenylation is employed by eukaryotic cells to produce variant products from the same gene, for example, either calciton or calcitonin gene-related peptides (Lou et al., 1998) or variants of immunoglobulin M heavy chain (Takagaki et al., 1996), our study extends this mechanism to include quantitative gene control. Independent of riboswitch-mediated regulation, the potential for the control of genes by regulating splicing and subsequently alternative 3′ end processing appears to be large; therefore, this general mechanism might be far more widespread in eukaryotes.

The unique location of TPP riboswitches in the 3′ regions of plant genes compared with their locations in fungi and bacteria might reflect adaptations to specific regulatory needs of different organisms. Nearly all known riboswitches reside in the 5′ UTRs of bacterial mRNAs (Mandal and Breaker, 2004; Soukup and Soukup, 2004; Winkler and Breaker, 2005) or in introns of 5′ UTRs or coding regions of fungi (Cheah et al., 2007) and often can suppress gene expression almost completely. However, a more subtle level of riboswitch regulation is observed in plants. Although plants can take up thiamin efficiently, most of the demand must be supplied by endogenous synthesis. In contrast with the autotrophic lifestyle of plants, fungi and bacteria sometimes grow under rich conditions that allow them to satisfy their entire requirements for compounds like thiamin by import, thus providing some rationale for different extents of regulation found in various groups of organisms.

METHODS

Plants and Plant Tissues

Arabidopsis thaliana ecotype Columbia-0 plants were grown on soil at 23°C in a growth chamber under a 16-h-light/8-h-dark photoperiod with 60% humidity unless otherwise stated. For seedling experiments, plants were grown on basal MS medium (Murashige and Skoog, 1962) supplemented with 2% sucrose and varying concentrations of thiamin and under continuous light unless otherwise specified. *Nicotiana benthamiana* plants for leaf infiltration assays were grown on soil for 3 to 5 weeks under continuous light. Plant material from other species was derived from seedlings grown from commercially available seeds.

Oligonucleotides and DNA Constructs

All synthetic DNAs and the details of cloning of DNA constructs are described in Supplemental Table 1 online.

RNA Isolation and RT-PCR Analyses

Total RNA was extracted from frozen plant tissues using the RNeasy plant mini kit (Qiagen) following the manufacturer's instructions. Two to five micrograms of total RNA were subjected to DNase treatment and subsequently reverse transcribed using SuperScript II reverse transcriptase (Invitrogen) according to the manufacturer's instructions. For cDNA generation, gene-specific primers or (if not otherwise specified) a polyT primer (DNA1) were used. cDNAs were used as templates for PCR amplification of *THIC* and *EGFP* reporter transcripts. All products obtained were cloned into TOPO-TA cloning vector (Invitrogen) and analyzed by sequencing (Howard Hughes Medical Institute Keck Foundation Biotechnology Resource Center at Yale University).

qRT-PCR was performed using the Applied Biosystems 7500 real-time PCR system and Power SYBR Green Master Mix (Applied Biosystems). Serial dilutions of the templates were conducted to determine primer efficiencies for all primer combinations. Each reaction was performed in triplicate, and the amplification products were examined by agarose gel electrophoresis and melting curve analysis. Data were analyzed using the relative standard curve method, and the abundance of target transcripts was normalized to reference transcripts reported previously (Czechowski et al., 2005) from genes *At1g13320* (PP2A catalytic subunit), *At5g60390* (EF-1α), and *At1g13440* (GAPDH).

Amplification of *THIC* Transcripts and Genomic Sequences from Plants

The 3′ UTRs from *THIC*-II RNAs were cloned using RT-PCR with a polyT primer and a degenerate primer that targets a conserved portion of the coding sequence near the stop codon. For *THIC*-III transcripts, 3′ UTRs were amplified in two fragments from polyT-generated cDNA using specific primer combinations. The 5′ portion of each 3′ UTR was PCR amplified using a degenerate primer targeting the coding region and a primer that targets the TPP aptamer. The 3′ portion of each 3′ UTR was obtained using a primer targeting the aptamer and a polyT primer. PCR products were cloned (TOPO-TA), and several independent clones were sequenced. The combined sequence information was used to design primer pairs for amplification of the corresponding genomic sequences. Genomic DNA was isolated using Plant DNAzol reagent (Gibco BRL) according to the manufacturer's instructions, and the resulting PCR products were cloned and sequenced.

RNA Gel Blot Analysis

Transcripts from *Arabidopsis* seedlings were analyzed by RNA gel blot analysis as described previously (Newman et al., 1993). Probes were specific against regions in the coding region of *THIC*, the extended 3′ UTR of *THIC* types I and III RNAs, or the control transcript *EIF4A1*.

In-Line Probing of RNA

In-line probing assays were conducted essentially as described previously (Winkler et al., 2002; Sudarsan et al., 2003). Details are described in Supplemental Methods online.

Agrobacterium-Mediated Leaf Infiltration Assay

For transient gene expression analysis, *N. benthamiana* leaves were transformed by a leaf infiltration assay as described by Cazzonelli and Velten (2006). *Agrobacterium tumefaciens* lines harboring the various reporter constructs were grown overnight in Luria-Bertani medium and centrifuged, and the pelleted cells were resuspended in water. OD_{600} was adjusted to the same value ($\sim$0.8) for cells harboring the different constructs, and *Agrobacteria* were mixed in equal amounts for cotransformation of constructs. Either luciferases from firefly (*Photinus pyralis*) or sea pansy (*Renilla reniformis*), or the fluorescent proteins EGFP and DsRed2, were used as reporter proteins. Additional details regarding reporter gene quantitation are described in Supplemental Methods online.

Stable Transformation of *Arabidopsis* by the Floral Dip Method

Arabidopsis was transformed by a floral dip method described previously (Clough and Bent, 1998). After transformation, seeds were grown under sterile conditions on medium containing 50 μg mL^{-1} kanamycin to select for transformants and 200 μg mL^{-1} cefotaxime to prevent bacterial growth. Surviving plants were transferred after 2 to 3 weeks to soil and expression of the transgene was determined after further growth.

Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession numbers EF588038 to EF588042. The Arabidopsis Genome Initiative locus identifier for *THIC* is At2g29630, and accession numbers for *THIC* sequences from additional species are listed in the legend to Figure 1.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure 1. Genomic DNA Sequence Contexts of TPP Riboswitches in *THIC* Genes from Different Plant Species.

Supplemental Figure 2. The Exon-Intron Organization of *THIC* 3′ UTRs Is Conserved, and Relative Distances of Key Features Are Similar among Different Species.

Supplemental Figure 3. Binding Sites and Sequences of DNA Primers Used for qRT-PCR Analysis of the Various *THIC* RNAs from *Arabidopsis*.

Supplemental Figure 4. The *THIC* Promoter from *Arabidopsis* Does Not Appear to be Responsible for Downregulation of *THIC* Expression after Thiamin Supplementation.

Supplemental Figure 5. Schematic Representation of Reporter Constructs and RT-PCR Detection of RNA Products Containing Two Different 3′ UTRs of At *THIC*.

Supplemental Figure 6. Effect of 3′ UTRs from Different Types of *THIC* Transcripts on Reporter Gene Expression.

Supplemental Figure 7. The Relative Position of the 3′ Splice Site for Formation of Type III RNAs Can Be Variable.

Supplemental Figure 8. TPP-Induced Modulation in the 5′ Flanking Sequence of the Aptamer.

Supplemental Table 1. Sequences of DNA Primers.

Supplemental Methods.

ACKNOWLEDGMENTS

We thank Monica Accerbi for technical assistance with the RNA gel blot analyses and Fred Souret and Justin Faletek for access to unpublished work. We thank Imad Ajjawi for providing us seeds of the TPK double

knockout *Arabidopsis* plants, Jeff Velten for the gift of a plasmid containing *Renilla* luciferase, and S.P. Dinesh-Kumar for providing the P19 vector and access to plant growth facilities. We also thank Elena Puerta-Fernandez and other members of the Breaker lab for helpful comments. A.W. was supported by a postdoctoral fellowship from the German Research Foundation, M.T. was funded by National Science Foundation Grant MCB-0236210, and B.C.G. received support from National Institutes of Health training grant GM07223. R.R.B. is supported by the Howard Hughes Medical Institute and received support for this project from National Institutes of Health Grants GM 068819 and DK070270.

Received June 14, 2007; revised October 16, 2007; accepted October 18, 2007; published November 9, 2007.

REFERENCES

Abreu-Goodger, C., and Merino, E. (2005). RibEx: A web server for locating riboswitches and other conserved bacterial regulatory elements. Nucleic Acids Res. **33:** W690–692.

Barrick, J.E., and Breaker, R.R. (2007). The distributions, mechanisms, and structures of metabolite-binding riboswitches. Genome Biol., in press.

Barrick, J.E., Corbino, K.A., Winkler, W.C., Nahvi, A., Mandal, M., Collins, J., Lee, M., Roth, A., Sudarsan, N., Jona, I., Wickiser, J.K., and Breaker, R.R. (2004). New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. Proc. Natl. Acad. Sci. USA **101:** 6421–6426.

Batey, R.T., Gilbert, S.D., and Montange, R.K. (2004). Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. Nature **432:** 411–415.

Cazzonelli, C.I., and Velten, J. (2006). An in vivo, luciferase-based, Agrobacterium-infiltration assay system: Implications for post-transcriptional gene silencing. Planta **224:** 582–597.

Cheah, M.T., Wachter, A., Sudarsan, N., and Breaker, R.R. (2007). Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. Nature **447:** 497–500.

Clough, S.J., and Bent, A.F. (1998). Floral dip: A simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J. **16:** 735–743.

Cochrane, J.C., Lipchock, S.V., and Strobel, S.A. (2007). Structural investigation of the glmS ribozyme bound to its catalytic cofactor. Chem. Biol. **14:** 97–105.

Corbino, K.A., Barrick, J.E., Lim, J., Welz, R., Tucker, B.J., Puskarz, I., Mandal, M., Rudnick, N.D., and Breaker, R.R. (2005). Evidence for a second class of S-adenosylmethionine riboswitches and other regulatory RNA motifs in alpha-proteobacteria. Genome Biol. **6:** R70.

Cromie, M.J., Shi, Y., Latifi, T., and Groisman, E.A. (2006). An RNA sensor for intracellular Mg(2+). Cell **125:** 71–84.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol. **139:** 5–17.

Edwards, T.E., and Ferre-D'Amare, A.R. (2006). Crystal structures of the thi-box riboswitch bound to thiamine pyrophosphate analogs reveal adaptive RNA-small molecule recognition. Structure **14:** 1459–1468.

Fuchs, R.T., Grundy, F.J., and Henkin, T.M. (2006). The S(MK) box is a new SAM-binding RNA for translational regulation of SAM synthetase. Nat. Struct. Mol. Biol. **13:** 226–233.

Gelfand, M.S., Mironov, A.A., Jomantas, J., Kozlov, Y.I., and Perumov, D.A. (1999). A conserved RNA structure element involved in the regulation of bacterial riboflavin synthesis genes. Trends Genet. **15:** 439–442.

Grundy, F.J., and Henkin, T.M. (1998). The S-box regulon: A new global transcription termination control system for methionine and cysteine biosynthesis genes in gram-positive bacteria. Mol. Microbiol. **30:** 737–749.

Hunt, A.G. (1994). Messenger 3′ end formation in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. **45:** 47–60.

Kertesz, S., Kerenyi, Z., Merai, Z., Bartos, I., Palfy, T., Barta, E., and Silhavy, D. (2006). Both introns and long 3′-UTRs operate as cis-acting elements to trigger nonsense-mediated decay in plants. Nucleic Acids Res. **34:** 6147–6157.

Kline, D.J., and Ferré-D'Amaré, A.R. (2006). Structural basis of glmS ribozyme activation by glucosamine-6-phosphate. Science **313:** 1752–1756.

Kubodera, T., Watanabe, M., Yoshiuchi, K., Yamashita, N., Nishimura, A., Nakai, S., Gomi, K., and Hanamoto, H. (2003). Thiamine-regulated gene expression of *Aspergillus oryzae* thiA requires splicing of the intron containing a riboswitch-like domain in the 5′-UTR. FEBS Lett. **555:** 516–520.

Lang, D., Eisinger, J., Reski, R., and Rensing, S.A. (2005). Representation and high-quality annotation of the *Physcomitrella patens* transcriptome demonstrates a high proportion of proteins involved in metabolism in mosses. Plant Biol. (Stuttg) **7:** 238–250.

Lou, H., Neugebauer, K.M., Gagel, R.F., and Berget, S.M. (1998). Regulation of alternative polyadenylation by U1 snRNPs and SRp20. Mol. Cell. Biol. **18:** 4977–4985.

Mandal, M., and Breaker, R.R. (2004). Gene regulation by riboswitches. Nat. Rev. Mol. Cell Biol. **5:** 451–463.

Mogen, B.D., MacDonald, M.H., Graybosch, R., and Hunt, A.G. (1990). Upstream sequences other than AAUAAA are required for efficient messenger RNA 3′-end formation in plants. Plant Cell **2:** 1261–1272.

Montange, R.K., and Batey, R.T. (2006). Structure of the *S*-adenosylmethionine riboswitch regulatory mRNA element. Nature **441:** 1172–1175.

Muhlrad, D., and Parker, R. (1999). Aberrant mRNAs with extended 3′ UTRs are substrates for rapid degradation by mRNA surveillance. RNA **5:** 1299–1307.

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plant. **15:** 473–497.

Nahvi, A., Barrick, J.E., and Breaker, R.R. (2004). Coenzyme B_{12} riboswitches are widespread genetic control elements in prokaryotes. Nucleic Acids Res. **32:** 143–150.

Newman, T.C., Ohme-Takagi, M., Taylor, C.B., and Green, P.J. (1993). DST sequences, highly conserved among plant SAUR genes, target reporter transcripts for rapid decay in tobacco. Plant Cell **5:** 701–714.

Rodionov, D.A., Vitreschak, A.G., Mironov, A.A., and Gelfand, M.S. (2002). Comparative genomics of thiamin biosynthesis in prokaryotes. New genes and regulatory mechanisms. J. Biol. Chem. **13:** 48949–48959.

Roth, A., Winkler, W.C., Regulski, E.E., Lee, B.W., Lim, J., Jona, I., Barrick, J.E., Ritwik, A., Kim, J.N., Welz, R., Iwata-Reuyl, D., and Breaker, R.R. (2007). A riboswitch selective for the queuosine precursor $preQ_1$ contains an unusually small aptamer domain. Nat. Struct. Mol. Biol. **14:** 308–317.

Serganov, A., Polonskaia, A., Phan, A.T., Breaker, R.R., and Patel, D.J. (2006). Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. Nature **441:** 1167–1171.

Serganov, A., Yuan, Y.R., Pikovskaya, O., Polonskaia, A., Malinina, L., Phan, A.T., Hobartner, C., Micura, R., Breaker, R.R., and Patel, D.J. (2004). Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. Chem. Biol. **11:** 1729–1741.

Soukup, G.A., and Breaker, R.R. (1999). Relationship between internucleotide linkage geometry and the stability of RNA. RNA **5:** 1308–1325.

Soukup, J.K., and Soukup, G.A. (2004). Riboswitches exert genetic control through metabolite-induced conformational change. Curr. Opin. Struct. Biol. **14:** 344–349.

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.

Sudarsan, N., Cohen-Chalamish, S., Nakamura, S., Emilsson, G.M., and Breaker, R.R. (2005). Thiamine pyrophosphate riboswitches are targets for the antimicrobial compound pyrithiamine. Chem. Biol. **12:** 1325–1335.

Takagaki, Y., Seipelt, R.L., Peterson, M.L., and Manley, J.L. (1996). The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. Cell **87:** 941–952.

Thore, S., Leibundgut, M., and Ban, N. (2006). Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. Science **312:** 1208–1211.

Vitreschak, A.G., Rodionov, D.A., Mironov, A.A., and Gelfand, M.S. (2003). Regulation of the vitamin B_{12} metabolism and transport in bacteria by a conserved RNA structural element. RNA **9:** 1084–1097.

Weinberg, Z., et al. (2007). Identification of 22 candidate structured RNAs in bacteria using Cmfinder comparative genomics pipline. Nucleic Acids Res. **35:** 4809–4819.

Winkler, W., Nahvi, A., and Breaker, R.R. (2002). Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. Nature **419:** 952–956.

Winkler, W.C., and Breaker, R.R. (2005). Regulation of bacterial gene expression by riboswitches. Annu. Rev. Microbiol. **59:** 487–517.