Genetic Control by Metabolite-Binding Riboswitches

Wade C. Winkler[a] and Ronald R. Breaker*[a]

1. Introduction

Modern organisms must coordinate the expression of many hundreds of genes in response to metabolic demands and environmental changes. Each gene product must be regulated temporally, quantitatively, and often spatially. Additionally, genetic control processes must be dynamic, rapid, and selectively responsive to the specific conditions undergoing change. Therefore, organisms require sentries of genetic regulatory factors that continuously quantify a multitude of signals. Upon measurement of a particular signal, which may be one of many possible biochemical or physical cues, these regulatory factors must modulate expression of a specific subset of the organism’s genes.

It has generally been assumed that proteins are the obligate sensors of these signals because protein is a proven medium for forming highly responsive sensors. However, recent findings demonstrate that mRNAs also are capable of acting as direct sensors of chemical [1–7] and physical conditions [8] for the purpose of genetic control. Classes of mRNA domains, collectively referred to as ‘riboswitches’, serve as RNA genetic control elements that sense the concentrations of specific metabolites by directly binding the target compound. Known riboswitches are responsible for sensing metabolites that are critical for fundamental biochemical processes including coenzyme B12 [9], thiamine pyrophosphate (TPP) [10], flavin mononucleotide (FMN) [11, 12, 13], S-adenosylmethionine (SAM) [14–16], lysine [17], guanine [18], and adenine [19]. Upon interaction with the appropriate small-molecule ligand, riboswitch mRNAs undergo a structural reorganization that results in the modulation of genes that they encode. To date, all riboswitches that have been examined in detail cause genetic repression upon binding their target ligand, although riboswitches that activate gene expression upon ligand binding are certainly possible.

In each instance, riboswitch domains have been subjected to a battery of biochemical and genetic analyses in order to convincingly demonstrate that direct interaction of small organic metabolites with mRNA receptors leads to a corresponding alteration in genetic expression. This review provides a brief summary of these efforts and of some of the general characteristics that are exhibited by riboswitches. Although these findings represent the initial steps in elucidating the principles that underlie RNA-based detection of small molecules, the prospects for riboswitches as genetic tools and as possible targets for development of antimicrobials are already beginning to emerge.

2. General Organization of Riboswitch RNAs

Bacterial riboswitch RNAs are genetic control elements that are located primarily within the 5′-untranslated region (5′-UTR) of the main coding region of a particular mRNA (Figure 1). Structural-probing studies (discussed further below) reveal that riboswitch elements are generally composed of two domains: a natural aptamer [12, 13] that serves as the ligand-binding domain, and an ‘expression platform’ that interfaces with RNA elements that are involved in gene expression (for example, Shine–Dalgarno (SD) elements, transcription terminator stems). These conclusions are drawn from the observation that aptamer domains synthesized in vitro bind the appropriate ligand in the absence of the expression platform [12–14]. Moreover, structural-probing investigations suggest that, when it is examined independently, the aptamer domain of most riboswitches adopts a particular secondary- and tertiary-structure fold that is essentially identical to the aptamer structure when it is examined in the context of the entire 5′-leader RNA. This implies that, in many cases, the aptamer domain is a modular unit that folds independently of the expression platform.

Ultimately, the ligand-bound or -unbound status of the aptamer domain is interpreted through the expression platform, which is responsible for exerting an influence upon gene expression. The view of riboswitch domains as modular elements is further supported by the fact that aptamer domains are highly conserved amongst various organisms (and even between...
kings, as is observed for the TPP riboswitch), whereas the expression platform varies in sequence, structure, and in the mechanism by which expression of the appended open reading frame is controlled. For example, ligand binding to the TPP riboswitch of the tenA mRNA of Bacillus subtilis causes transcription termination. This expression platform is distinct in sequence and structure compared to the expression platform of the TPP riboswitch in the thim mRNA from Escherichia coli, wherein TPP binding causes inhibition of translation by an SD blocking mechanism. The TPP aptamer domain is easily recognizable and of near identical functional character between these two transcriptional units, but the genetic control mechanisms and the expression platforms are very different.

Aptamer domains for riboswitch RNAs typically range from ~70 – 170 nucleotides in length (Scheme 1). These large sizes were somewhat unexpected given that in vitro evolution experiments identified a wide variety of small-molecule-binding aptamers, which are considerably shorter in length and structural intricacy (Scheme 1). Although the reasons for the substantial increase in complexity and information content of the natural aptamer sequences relative to artificial aptamers remain to be proven, this complexity is most likely required to form RNA receptors that function with high affinity and selectivity. Apparent dissociation constants ($K_d$) for the ligand–riboswitch complexes range from low nanomolar to low micromolar values. It is also worth noting that some aptamer domains, when isolated from the appended expression platform, exhibit improved affinity for the target ligand over that of the intact riboswitch (typically ~10- to 100-fold improvement). Presumably, there is an energetic cost in sampling the multiple distinct RNA conformations required by a fully intact riboswitch RNA, which is reflected by a loss in ligand affinity. Since the aptamer domain must serve as a molecular switch, this might also add to the functional demands on natural aptamers and help explain their more sophisticated structures.

3. Riboswitch Regulation of Transcription Termination in Bacteria

Bacteria primarily make use of two methods for termination of transcription. Certain genes incorporate a termination signal that is dependent upon the Rho protein, while others make use of Rho-independent terminators (intrinsic terminators) to destabilize the transcription elongation complex. The latter RNA elements are composed of a GC-rich stem-loop followed by a stretch of 6 – 9 uridyl residues. Intrinsic terminators are widespread throughout bacterial genomes and are typically located at the 3′-termini of genes or operons. Interestingly, an increasing number of examples of intrinsic terminators are being identified within 5′-UTRs.

Amongst the wide variety of genetic regulatory strategies employed by bacteria there is a growing class of examples wherein RNA polymerase responds to a termination signal within the 5′-UTR in a regulated fashion. Under certain conditions the RNA-polymerase complex is directed by external signals either to perceive or to ignore the termination signal. Presumably, one of at least two mutually exclusive mRNA conformations results in the formation or disruption of the RNA structure that signals transcription termination. A trans-acting (intermolecularly acting) factor, which in some instances is an RNA and in others is a protein, is generally required for receiving a particular intracellular signal and subsequently stabilizing one of the RNA conformations. In contrast, riboswitches offer a direct link between RNA structure modulation and the metabolite signals that need to be interpreted by the genetic control machinery. A brief overview of the FMN riboswitch from a B. subtilis mRNA is provided below to illustrate this mechanism.

3.1. A natural aptamer for FMN

A highly conserved RNA domain, referred to as the RFN element, was identified in bacterial genes involved in the biosynthesis and transport of riboflavin and FMN. This element is required for genetic manipulation of the ribDEAHT operon (hereafter, ribD) of B. subtilis, as mutations resulted in a loss of FMN-mediated regulation. These data led to the proposal that either a protein-based FMN sensor or FMN itself interacts with the RFN element in order to repress ribD gene expression. Although RNA sequences that specifically bind FMN have been identified through directed-evolution experimentation, they exhibit no obvious resemblances to the RFN element (Scheme 1), a fact suggesting that FMN recognition by the RFN element would be achieved by a mechanism distinctive to that of the engineered aptamers.

3.1.1. Structural probing reveals FMN-mediated RNA folding modulation

Each internucleotide linkage in an RNA polymer is susceptible to spontaneous hydrolysis by an $S_2$-like mechanism, wherein the $\dot{2}^\prime$-oxygen atom attacks the adjacent phosphorus center, thereby leading to chain cleavage. For optimal speed, this reaction requires a 180° orientation between the attacking nucleophile, the phosphorus center, and the 3′-oxygen leaving group (in-line conformation). Nucleotides that are base paired, or otherwise structurally constrained, are typically incapable of adopting this configuration and therefore display low rates of spontaneous cleavage. In contrast, nucleotides that are structurally unrestrained exhibit much higher rates of spontaneous cleavage. These observations have been exploited in a structural-probing method, referred to as ‘in-line probing’, which establishes the relative rates of spontaneous cleavage for a given RNA polymer and correlates this with secondary- and tertiary-structure models.

To assess whether the RFN element of ribD was responsive to FMN, a fragment of the corresponding 5′-UTR was 5′-$\gamma$-P labeled and incubated in the absence and presence of FMN, and the resulting fragments were analyzed by polyacrylamide gel electrophoresis (PAGE). Interestingly, patterns differ between reactions that are incubated either with or without FMN, a fact signifying that there is a structural rearrangement of the RNA upon FMN binding to ribD. The spontaneous cleavages of certain nucleotide positions located within interhelical regions of

the RFN element become significantly reduced in the presence of FMN; this suggests that these nucleotides are involved in forming an FMN–RNA complex, which forces structural constraints upon the RNA (Scheme 2). It is this type of structural modulation that can be harnessed by the expression platform for allosteric modulation of gene expression.

Additional evidence for direct binding of FMN by the ribD RFN element was generated by enzymatic probing. Oligonucleotides predicted to anneal with the RFN element were added to ribD transcripts in the presence and absence of FMN, and the resulting mixtures was digested with RNase H (which specifically cleaves RNA:DNA heteroduplexes) and analyzed by PAGE. A significant portion of transcripts bind certain oligonucleotides in the absence of FMN, but not in the presence of FMN, a fact indicating that FMN stabilizes a structural rearrangement of ribD transcripts that in turn prevents annealing of the oligonucleotide.

3.1.2. Affinity and specificity of the FMN–ribD complex

If the RFN element serves as an aptamer for FMN, it should exhibit characteristics of a saturable receptor that has some ability to discriminate against related ligands. To obtain apparent $K_d$ values for FMN, in-line probing assays were repeated with trace amounts of ribD RNA and increasing concentrations of FMN. The ligand concentration that correlates with half-maximal modulation of RNA structure should reflect the apparent $K_d$ value. These experiments indicate that the ribD RNA contains a saturable ligand binding site that exhibits an apparent $K_d$ value of $\approx 5 \text{ nM}$. Furthermore, the RNA discriminates against riboflavin (the dephosphorylated form of FMN) by approximately three orders of magnitude. This exceptional ligand specificity of the ribD mRNA is surprising since the aptamer must generate a binding pocket for FMN that makes productive interactions with a phosphate group. An interesting aspect of future structure analyses will be to determine how RNA, a highly anionic polymer, is capable of preferentially recognizing the negatively charged phosphate group of FMN.

3.2. FMN-induced transcription termination

3.2.1. In vitro transcription termination mediated by an FMN riboswitch

The relative amounts of the major transcription products for the ribD leader region were examined by in vitro transcription with T7 RNA polymerase or Bacillus subtilis RNA polymerase. The ribD leader region contains a typical intrinsic terminator just upstream of the ribD coding region. Interestingly, transcripts that terminated at the intrinsic terminator are specifically induced by FMN, in the absence of additional protein factors. Furthermore, mutations in the RFN element abrogate this phenomenon. The left half of the terminator sequence is proposed to form alternative base-pairing interactions with a portion of the RFN element, thereby forming an antitermination element. Sequence

Scheme 2. Regulation of the B. subtilis ribD mRNA by FMN. A) In-line probing reveals ligand-dependent structure modulation with FMN: internucleotide linkages identified with red circles exhibit decreased amounts of spontaneous cleavage when ribD is incubated in the presence of FMN (indicating an increase in structural stabilization for these nucleotides) relative to incubation in the absence of FMN. Yellow circles identify linkages that exhibit consistently high levels of scission, which indicates that they are not modulated by presence of FMN. B) Model for the mechanism of ribD regulation. The ribD mRNA adopts an antitermination conformation in the absence of FMN. Increased levels of FMN stabilize an RFN–FMN complex that permits formation of the terminator structure. C) Chemical structure and apparent dissociation constants for riboflavin and FMN.
Alterations of the intrinsic terminator eliminate FMN-induced termination while alterations in the antiterminator result in constitutive termination.\cite{5, 56} Taken together, these observations are consistent with a mechanistic model wherein FMN directly interacts with ribD transcripts during conditions of excess FMN. Complex formation subsequently induces transcription termination within the 5′-UTR (Scheme 2), which precludes gene expression by preventing the downstream coding regions from being transcribed. During conditions of limited FMN, an antiterminator structure is formed within the ribD nascent transcript, which allows for synthesis of the complete mRNA.  

3.2.2. FMN-mediated control of transcription termination in vivo

The molecular details of riboswitch-mediated transcription termination are likely to be more complex than this rather simplistic model implies. For example, given that the ‘decision’ to form the terminator or antiterminator conformation occurs only once during transcription, the regulatory mechanism is likely to rely on precise transcriptional kinetics as well as the appropriate RNA folding pathways. Moreover, the kinetics of FMN interacting with the RNA receptor are presumed to be a critical factor. Although the affinity that the RNA has for FMN is exceptionally strong compared to engineered aptamers, it is possible that the kinetics of ligand association might be the more important determinant of genetic regulation. Indeed, all of these parameters are likely to conspire together in order to exert appropriate control over the intrinsic terminator. If true, then the kinetic parameters determined for FMN binding and for FMN-induced transcription termination in vitro might be an imperfect reflection of the function of riboswitches in vivo. Specifically, the impact of transcription speed and of the various cellular conditions might lead to significant differences in the performance characteristics of riboswitches.

3.3. Control of transcription termination by other riboswitches

Intrinsic terminators can be identified by computer-assisted search algorithms.\cite{20} By using such bioinformatic analyses, it is possible to identify a subset of riboswitch RNA that are predicted to contain an intrinsic terminator and an alternate antiterminator structural element.\cite{22, 37–41} Therefore, the results described above for the FMN riboswitch may be indicative of the mechanisms used by many other riboswitch RNAs. Indeed, SAM- and TPP-dependent riboswitches have been demonstrated to exert control over termination by formation of mutually exclusive intrinsic terminator and antiterminator structures.\cite{5–7} Furthermore, mutations that disrupt and subsequently restore helices within the SAM riboswitch aptamer result in loss and restoration, respectively, of SAM binding. Concurrently, these mutations also result in disruption or restoration of SAM-induced transcription termination in accordance with ligand-binding function.\cite{27} It is also possible, and perhaps even likely, that some riboswitches will exert control over transcription termination signals that differ appreciably from classical intrinsic terminators.

4. Riboswitch Regulation of Translation Initiation in Bacteria

An alternative mechanism of genetic control by riboswitches is the modulation of translation initiation. Unlike transcription termination, the entire mRNA could be synthesized by RNA polymerase, but expression would be prevented by the riboswitch until the metabolite concentration reached a certain level. In most instances, we observed that riboswitches prevent translation initiation in the presence of high concentrations of target metabolite. However, it is certainly possible that allosteric modulation of riboswitch structures could lead to translation activation. The regulatory mechanism of translation control is briefly described below for a TPP riboswitch from \textit{E. coli}.

4.1. A natural aptamer for TPP

A conserved RNA element, referred to as the \textit{thi} box, was identified within 5′-UTRs of mRNAs that are responsible for thiamine biosynthesis and transport.\cite{41, 42} Genetic experiments confirmed that this structural element is required for thiamine-dependent regulation of \textit{Rhizobium meliloti} thiamine biosynthesis genes,\cite{42} yet no regulatory factor had been identified through classical genetic experimentation. Therefore, it was possible that the \textit{thi} box might serve as a portion of a riboswitch that responds to thiamine or its derivatives.

In \textit{E. coli}, thiamine biosynthesis and transport genes are primarily located within three operons and four single genes,\cite{43} wherein each operon is preceded by a \textit{thi} element. To begin to assess the regulatory properties of these sequences, the leader regions for the \textit{thiMD} and \textit{thiCEFSGH} operons were utilized to construct transcriptional and translational fusions to a \textit{lacZ} reporter gene.\cite{20} Addition of exogenous thiamine results in repression of the \textit{lacZ} reporter gene in \textit{E. coli}. Results from these data demonstrate that the \textit{thiM} gene is regulated primarily at the level of translation while the \textit{thiC} leader region confers both transcriptional and translational regulation to the \textit{lacZ} reporter.

4.1.1. Direct binding of thiamine pyrophosphate by \textit{E. coli} mRNAs

As described above for the FMN aptamer, direct binding of TPP to the \textit{thiM} and \textit{thiC} leaders was demonstrated by in-line probing assays.\cite{22} The addition of thiamine, thiamine monophosphate (TP), or TPP leads to structural rearrangement of the \textit{thiM} RNA, particularly in the region encompassing the \textit{thi} element (Scheme 3). Significantly, TPP, which is typically the bioactive form of thiamine, exhibits the best affinity of the ligands, with an apparent \textit{K}_D value of 500 nM, while TP and thiamine exhibit apparent \textit{K}_D values of 3 and 40 μM, respectively. In-line probing assays of RNAs resembling the \textit{thiC} leader reveal even more dramatic discrimination between thiamine and its phosphorylated forms, with a difference of more than 1,000-fold exhibited between binding of thiamine and TPP. These data are consistent with genetic experiments that suggested that TPP synthesis is required for regulation.\cite{44, 45} Also, this system
provides another example of a natural RNA aptamer that makes productive contacts to phosphate groups.

4.1.2. Confirmation of TPP binding by equilibrium dialysis

RNAs resembling the thiM leader region were synthesized and placed into one side of a two-chamber equilibrium dialysis apparatus, in which the compartments are separated by a dialysis membrane with a molecular-weight cut off of 5000 daltons. ³H-thiamine was preferentially retained within the thiM-containing chamber when allowed to equilibrate between chambers.³¹ This effect could be eliminated by providing excess unlabeled thiamine, but could not be reversed when supplemented with oxythiamine, a close chemical analogue of thiamine. Additionally, a mutated version of thiM was unable to shift ³H-thiamine into the RNA-containing chamber. Together, these data are indicative of the formation of stable thiM±thiamine complexes, wherein the sequence of the RNA and the chemical form of the ligand are critical for maximal binding affinity.

4.2. Binding of thiamine derivatives correlates with structural modulation

Close inspection of in-line probing data for thiM reveals two surprising patterns of structural modulation. First, the relative rates of spontaneous fragmentation between reactions containing either thiamine or TPP differ within an internal loop of the thi element (Scheme 3). Nucleotides in this region adopt an increase in structural order in the presence of TPP but not with thiamine, a fact implying that this region is somehow involved in the formation of a pyrophosphate-recognition pocket. Secondly, the region of the SD sequence is the only portion outside of the thi element that becomes structurally modulated in the presence of TPP.

Specifically, the SD sequence exhibits a significant decrease in spontaneous cleavage relative to reactions lacking TPP; this suggests that the SD is converted into a more structurally constrained form upon binding of TPP. This idea is consistent with a mechanism (Scheme 3) whereby in the absence of TPP the SD has a significant degree of single-stranded character and is accessible for translation initiation. An anti-SD sequence is proposed to interact with an anti-anti-SD sequence within the TPP aptamer under these conditions. In contrast, during conditions of excess TPP, a TPP±RNA complex is formed that disrupts the base pairing of the anti-SD sequence. The anti-SD sequence is then free to interact directly with the SD, decrease the single-stranded character of the region, and subsequently decrease the efficiency of translation initiation. Preliminary site-directed mutagenesis of the thiM mRNA supports this overall model.³¹ Specifically, mutations that disrupt TPP binding also disrupt regulation of translation for thiM±lacZ fusions, while mutations that alter the anti-SD sequence affect regulation but do not affect TPP binding. Thus, binding of thiamine correlates with both the structural accessibility of the SD and the translation efficiency in vivo.

4.3. Control of translation initiation by other riboswitches

Bioinformatics analyses have suggested that molecular mechanisms similar to that of thiM also might be recurrent amongst riboswitch RNAs. Specifically, anti-SD and anti-anti-SD structures have been proposed for several riboswitch classes, including FMN,³² lysine,³¹ TPP,³¹ coenzyme B₁₂,¹, ⁴⁰ and SAM.³¹ In general, riboswitches from Gram-negative organisms seem to favor expression platforms that exert control over translation, while
Riboswitches from Gram-positive bacteria appear to predomi-
nately use expression platforms that control transcription ter-
mination. The latter might reflect a greater reliance upon
multigene transcriptional units in Gram-positive organisms,
which could be more efficient to preclude transcription of long
operons when the gene products are unnecessary.

Biochemical evidence for riboswitch-mediated control over
translation initiation has also been obtained for FMN and
coenzyme B$_2$, riboswitches.[1, 3] FMN binding to a riboswitch that
regulates the B. subtilis ypaA gene results in alteration of the SD
structural context, in a similar manner to what was observed for
thiM.[10] Interestingly, this genetic control element has also been
proposed to regulate ypaA transcription,[46] although the leader
region does not contain an obvious intrinsic terminator struc-
ture. Binding of coenzyme B$_2$, to the E. coli btuB riboswitch has
also been demonstrated to correlate with regulation of trans-
lation in vivo. However, this RNA differs from the thiM paradigm
as the mechanism for control over translation efficiency, which
has not yet been elucidated, appears to be derived from
something other than the formation of alternate, mutually
exclusive stem-loop structures (anti-SD and anti-anti-SD).

Preliminary data also indicate that certain riboswitch RNAs
exert control over transcription and translation by using the
same RNA sequence.[47] For this class of riboswitches, the SD
sequence is contained within an intrinsic terminator. Therefore,
the formation of the terminator structure also enacts formation
of a SD-sequestering structure. In total, all of these observations
suggest that the thiM and ribD riboswitches represent useful
paradigms for riboswitch-mediated control of translation and
transcription, respectively. However, there is likely to be a greater
variety of molecular mechanisms utilized by riboswitch RNAs for
control of gene expression. Indeed, TPP riboswitches that must
be employing different mechanisms of control have been
identified in several plant and fungal species.[14] The placement
of these RNAs near splice sites in some instances and in the 3’-
UTR in others is suggestive of TPP-responsive control over
splicing and mRNA stability or expression, respectively.

5. Early Origins?
The FMN, TPP, lysine, and coenzyme B$_2$, riboswitch RNAs are
widespread among evolutionarily distant microorganisms, a fact
that implies there is an ancient origin for these RNA genetic
elements.[16, 23, 40, 41] SAM, guanine, and adenine riboswitches are
also represented in numerous different genera, although they
appear to be primarily limited to Gram-positive bacteria, with a
few Gram-negative bacteria as exceptions.[2, 39, 48] In all instances,
the structural and sequence conservation of riboswitch classes is
limited to the aptamer domain (Scheme 1). This is not unex-
pected given that the aptamer RNA must preserve its capability
to bind the target chemical, which has not been significantly
modified through evolution. In contrast, there is considerable
sequence and structural diversity between expression platforms,
even between riboswitches of the same class and within the
same organism. Together, these data hint that the ligand-
binding properties of riboswitch aptamer domains have been
maintained throughout expansive evolutionary timescales.

Furthermore, the ligands for riboswitch RNAs have been
proposed to be functional relics from a hypothetical RNA-based
world, in which RNA polymers provided all the necessary
catalytic and genomic functions of the earliest organisms.[49, 50]
Therefore, it is tempting to speculate that, as cofactor-binding
RNAs, the aptamer domains from riboswitches may have been
useful in the context of an RNA-based world for some of the
earliest forms of genetic control, for allosteric modulation
of ribozymes, or as part of ribozymes that utilized the ligands as
essential cofactors.

6. Riboswitches as Drug Targets and Genetic Tools
Riboswitches are utilized for control of numerous genes involved
in the biosynthesis and transport of prokaryotic enzymatic
cofactors. At least 69 genes, which represents nearly 2% of the
total genomic content of B. subtilis, are under the control of
riboswitch RNAs (Table 1); this exemplifies the extensive use of
riboswitch RNAs for genetic control in prokaryotes.[39] Many
riboswitch-mediated genes are expected to be essential under
most growth conditions. Interference with riboswitch function is
then predicted to result in dramatic destabilization of vital
metabolic pathways and, perhaps, cessation of growth. There-
fore, it seems likely that compounds that closely resemble the
target metabolites will bind to riboswitch RNAs and cause a
decrease in gene expression. If this analogue-induced disruption
of gene expression is sufficient, then such compounds might be
candidates for antimicrobial applications.

There is clear precedence for the targeting of RNAs with small-
molecule drugs,[51] the most obvious example being that of
ribosomal RNA. Several other bacterial-specific RNAs have been
explored as candidates for small-molecule drug interaction;
however, the approach relies upon screening large chemical
libraries for those chemicals that fortuitously interact with the
RNA of interest, even though the RNA itself does not naturally
form a binding pocket for small organic molecules. Riboswitch
RNAs therefore may exhibit an advantage in antimicrobial
development given that they serve as a receptor for small-
molecule ligands, much like their protein receptor counterparts.
The continued exploration of the molecular recognition of target
ligands by riboswitch RNAs will ultimately reveal whether ‘blind
spots’, that can be exploited for the development of antimicro-
bials, exist within the molecular recognition landscape.

In addition to their potential as targets for chemical inhibition,
a detailed understanding of the mechanisms utilized by natural
riboswitch RNAs may lead to the development of novel genetic
control elements. Numerous aptamer RNA sequences have been
identified that interact with a wide variety of small organic
molecules.[15] It is reasonable to expect that engineered ribos-
witches could be generated that respond to nonbiological, or
otherwise metabolically inert, compounds. The range of uses for
such genetic control elements will then be limited only by the
range of human imagination and the speed at which engineered
aptamers and expression platforms could be built.

Further exploration of riboswitch mechanism, character, and
distribution in biological systems will require a combination of
genomic, bioinformatic, and biochemical techniques. It is possible that the riboswitches observed to date are of ancient origin, and thus, their evolutionary distribution and functions might reflect this possibility. Further examination of the kinetic and molecular recognition properties will also reveal whether riboswitches are equal to the task of genetic control, as compared to protein factors, or whether we are observing the last vestiges of an ancient but imperfect form of genetic regulation machinery. Given the fundamental metabolic processes that riboswitches control in modern organisms, further studies hold considerable promise for revealing insights into how organisms orchestrate complex genetic networks.

**Keywords:** aptamers · gene expression · riboswitches · RNA · transcription

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Transcriptional unit</th>
<th>Predicted gene function(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>lysine</td>
<td>lySC</td>
<td>aspartokinase II</td>
</tr>
<tr>
<td>FMN</td>
<td>ypaA</td>
<td>putative flavin transporter</td>
</tr>
<tr>
<td>coenzyme B$_D$</td>
<td>ribD-ribe-ribBA-ribH</td>
<td>riboflavin biosynthesis</td>
</tr>
<tr>
<td>TTP</td>
<td>thyC</td>
<td>biosynthesis of thiamine pyrimidine moiety</td>
</tr>
<tr>
<td></td>
<td>tenA1-thiX1-thiY1-thiZ1-thiE2-thiD</td>
<td>thiamine biosynthesis</td>
</tr>
<tr>
<td></td>
<td>ykof-ykoA-ykoD-ykoC</td>
<td>unknown; similar to iron transport proteins</td>
</tr>
<tr>
<td>guanine</td>
<td>yxA</td>
<td>unknown; putative thiamine transporter</td>
</tr>
<tr>
<td></td>
<td>xpt-pxuX</td>
<td>similar to acetylornithine deacetylase</td>
</tr>
<tr>
<td>adenine</td>
<td>pbaG</td>
<td>similar to pyrimidine nucleoside transport</td>
</tr>
<tr>
<td></td>
<td>purE-purB-purC-purS-purQ-purL-purF-purM-purN-purH-purD</td>
<td>xanthine permease</td>
</tr>
<tr>
<td>SAM</td>
<td>ydhL</td>
<td>hypoxanthine/guanine permease</td>
</tr>
<tr>
<td></td>
<td>yIU</td>
<td>purine biosynthesis</td>
</tr>
<tr>
<td></td>
<td>met-metC</td>
<td>unknown</td>
</tr>
<tr>
<td></td>
<td>ykT-ykxS</td>
<td>putative methylene tetrahydrofolate reductase</td>
</tr>
<tr>
<td></td>
<td>ykrW-ykrX-ykrY-ykrZ</td>
<td>methionine biosynthesis</td>
</tr>
<tr>
<td></td>
<td>cysH-cysP-sat-cysC-ylnD-ylnE-ylnF</td>
<td>5'-methylthioadenosine recycling pathway</td>
</tr>
<tr>
<td></td>
<td>yoaD-yoaC-yoaB</td>
<td>5'-methylthioadenosine recycling pathway</td>
</tr>
<tr>
<td></td>
<td>metE</td>
<td>cysteine biosynthesis</td>
</tr>
<tr>
<td></td>
<td>metK</td>
<td>unknown</td>
</tr>
<tr>
<td></td>
<td>yucC-yucB-yucA</td>
<td>methionine synthase, B$_D$-independent</td>
</tr>
<tr>
<td></td>
<td>yxA</td>
<td>S-adenosylmethionine synthetase</td>
</tr>
<tr>
<td></td>
<td>yxG</td>
<td>unknown ABC transporter</td>
</tr>
<tr>
<td></td>
<td>yxH</td>
<td>unknown</td>
</tr>
</tbody>
</table>

[a] Gene nomenclature is derived from the SubList database [54] except for metC and metE, which are recent designations [56]. Functional roles for ypaA, [57] ysaL [58], ykrTS [56], and ykrWXyz [56] have recently been proposed.

---


[40] A. Nahvi, J. E. Barrick, M. S. Ebert, R. R. Breaker, unpublished results.

Received: June 11, 2003 [M685]