The POWER of Riboswitches

Discovering relics from a lost world run by RNA molecules may lead to modern tools for fighting disease

By Jeffrey E. Barrick and Ronald R. Breaker

A mystery surrounding the way bacteria manage their vitamins piqued our interest in the fall of 2000. Together with growing evidence in support of a tantalizing theory about the earliest life on earth and our own efforts to build switches from biological molecules, the bacterial conundrum set our laboratory group at Yale University in search of an answer. What we found was a far bigger revelation than we were expecting: it was a new form of cellular self-control based on one of the oldest types of molecule around—ribonucleic acid, or RNA.

Long viewed as mostly a lowly messenger, RNA could have considerable authority, as it turned out, and sophisticated mechanisms for asserting it. Although the workings of this newfound class of RNA molecules that we dubbed riboswitches are still being characterized, it is already clear that they may also provide novel ways of fighting human diseases. Many pathogenic bacteria rely on riboswitches to control aspects of their own fundamental metabolism, for instance.

That this ancient form of self-regulation persists in modern organisms attests to its importance. Bacterial cells are astonishingly adaptive and self-reliant chemical factories dedicated to making one final product: endless copies of themselves. But only strains that have been able to maintain this hurried chain of descent in the face of cutthroat competition for resources in changing environments have survived this long.

Inventory Control

A BACTERIUM'S ABILITY to craft the hundreds of elaborate molecules required to replicate itself in as little as 20 minutes starts with the double-stranded DNA genome that every living organism faithfully copies from generation to generation. This operating manual is written in the four-nucleotide DNA alphabet of the nitrogenous bases adenine, thymine, cytosine and guanine, which are strung onto an alternating sugar-phosphate backbone. As much as 90 percent of the DNA in a typical bacterium is dedicated to coded instructions for assembling the protein machinery that accelerates and organizes the chemical steps of metabolism neces-

SLIM FILMS

sary to build a new cell from scratch.

On the cellular factory floor, that process begins when RNA polymerase enzymes latch onto the genomic DNA and start copying portions of its text into the chemically similar form of messenger RNA (mRNA) molecules. Bacterial cells are in such a hurry that after one polymerase has scarcely begun reading the DNA message and transcribing it, another polymerase is pressing eagerly against it to begin the next mRNA copy. Most messages encode a single protein, although some, known as operons, describe how to make an entire suite of operationally related proteins. RNA is less chemically stable than DNA, and the bacterial cell treats these multiple mRNA transcripts like paper photocopies. Unused mRNAs are rapidly shredded and recycled so that only fresh work orders get distributed to ribosomes, the factory's protein-building machinists.

Ribosomes, too, are in a rush, typically lining up like boxcars of a train to start reading and executing the mRNA instructions even before the polymerase enzyme has completely finished the transcript. They chug along the mRNA track, decoding each successive triplet of nucleotides into a specific amino acid and adding that to a growing chain. As this protein emerges from the ribosome, it wraps around itself into a complicated three-dimensional structure, and a new piece of machinery is now ready to float off and begin work.

The cell relies in particular on two categories of protein to keep its chemical production humming smoothly: transporters, which shuttle raw materials, and enzymes, which accelerate their transformation through successive steps in the dizzying cycles and pathways of metabolism. Bacteria are careful not to waste resources by making superfluous infrastructure, however, so they have evolved control mechanisms that can short-circuit the transfer of work orders for this equipment in response to changes in nutrient needs and availability. Scientists' understanding of how those cellular supervisors function first raised the mystery of vitamin management.

Bacteria typically employ a number of proteins that constantly check the current stocks of various raw materials and adjust the number of transporters and enzymes allocated to different production lines. The Lac repressor in the gut bacterium *Escherichia coli,* for example, is a protein complex that blocks access to DNA blueprints, both for a transporter that pumps the sugar lactose into the cell and for an enzyme that cleaves lactose apart so that it can be used as fuel, until they are needed. When lactose concentration rises above a certain threshold, the Lac complex lets go of the DNA template, allowing transcription of the genes to commence.

A similar regulatory mechanism relies on protein supervisors that decide what to do with mRNA strands as they are being copied from genomic DNA. In the soil bacterium *Bacillus subtilis,* a protein complex with the acronym TRAP controls one operon encoding enzymes for synthesizing the amino acid tryptophan and another describing a tryptophan transporter. When TRAP senses that these proteins are not needed, it wraps the leading end of their mRNA instructions tightly around itself. This prevents a ribosome from recognizing a valid site on the transporter transcript to start translation. Sequestration of its leader causes the synthesis mRNA strand to fold into a hairpin shape, held together by nucleotides binding to one another, that prematurely terminates transcription of the message [*see box on opposite page*].

In addition to such equipment for regulating the manufacture of basic cellular protein machinery, bacteria carry around a large toolbox for making more exotic chemicals. We humans must obtain the nutrients we call vitamins from what we eat, for example, whereas bacteria know how to assemble them from scratch. Many of the more complex vitamins are actually versions of "coenzymes," which, as the name hints, are small molecules that cooperate with protein enzymes. They are specialty tools, akin to pneumatic nail guns and diamond drill bits, with powerful chemical functions. Epic metabolic pathways are involved in constructing coenzymes from raw materials, and naturally thrifty bacteria strictly control this expensive synthesis by shutting it down when coenzymes are not in demand.

By the late 1990s scientists investigating exactly how manufacture of certain coenzymes was regulated in bacteria recognized a molecular pattern reminiscent of the TRAP and Lac repressor

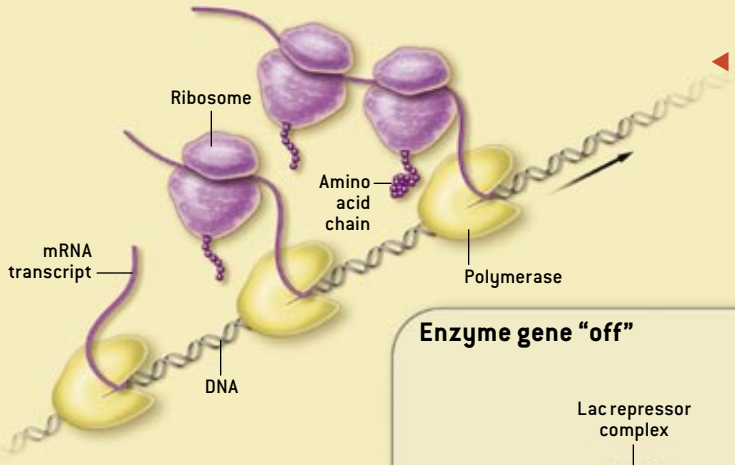
These photocopies were folding like possessed origami and CHOOSING THEIR OWN FATES.

Overview/*Flipping Ancient Switches*

- Regulation of gene activity in cells is usually the job of protein supervisors, but certain bacteria employ RNA messengers to oversee some valuable cellular infrastructure.
- Forms of RNA with proteinlike powers lend support to the idea of a primordial world ruled by RNA.
- Newly discovered riboswitches are a group of RNA molecules that carry messages transcribed from DNA while also making supervisory decisions about whether those instructions should be executed.
- Riboswitches regulate many fundamental processes in microbes, making them potential targets for new antimicrobial drugs.

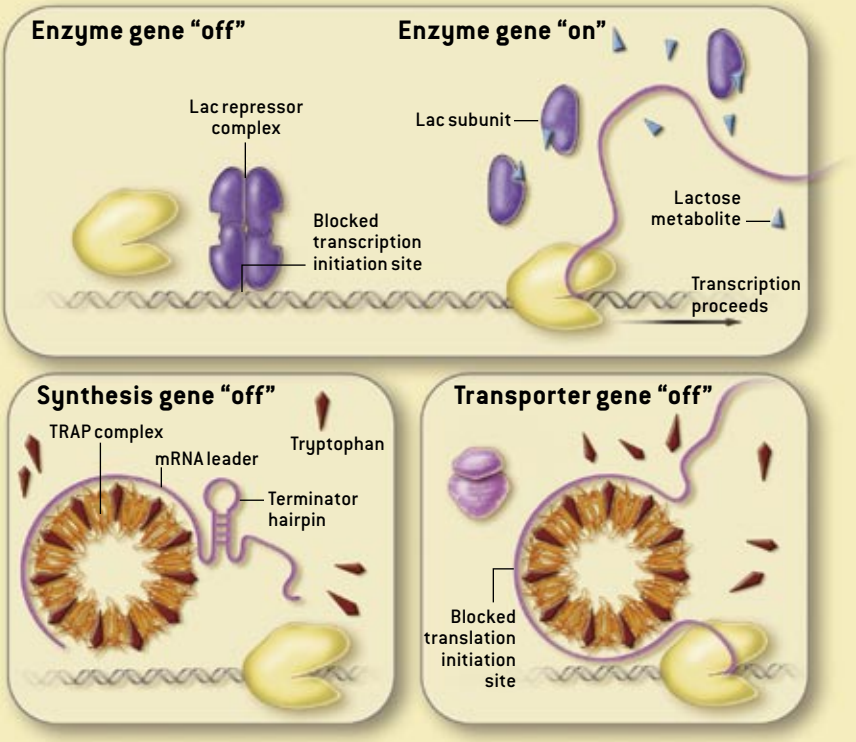
PROTEIN SUPERVISORS IN THE CELLULAR FACTORY

To coordinate and optimize manufacture of the parts bacteria need to survive and replicate, the cells typically employ protein supervisors. These can repress production of equipment until they sense it is needed and the raw materials for making it are available. Knowledge of their mechanisms helped to reveal the existence of riboswitches.



THE FACTORY FLOOR

Fast-moving assembly lines turn out cellular equipment based on instructions encoded in DNA genes (*left*). Polymerase enzymes move along the DNA strand, transcribing a gene into a messenger RNA (mRNA) copy. Ribosomes latch onto the mRNA as it emerges and begin translating its message into an amino acid chain that will fold itself into a finished protein.



INVENTORY MANAGEMENT

Supervisor proteins regulate a bacterium's manufacture of basic parts through a variety of mechanisms (*right*).

The Lac repressor complex (*top*) turns "off" a gene encoding a lactose-cleaving enzyme by blocking polymerase from accessing the DNA when lactose is absent. When lactose is high, one of its metabolites binds to clefts in Lac subunits, causing them to let go of the DNA and turning the gene "on."

A TRAP complex regulates genes involved in the synthesis and transport of the amino acid tryptophan by interfering with their mRNA copies in two ways. When tryptophan is present, TRAP wraps the leading end of a tryptophan-synthesis mRNA around itself, causing part of the message strand to assume a hairpin shape that prematurely terminates transcription (*bottom left*). TRAP also sequesters the leader of mRNA for a tryptophan transporter, blocking ribosomes from accessing a translation initiation site (*bottom right*).

systems. Yet as their attempts to identify the supervisory proteins responsible for sensing each coenzyme and controlling mRNA transcription or translation in response drew a blank, a deepening mystery emerged: if not through hypothetical protein supervisors, how was the cell's machinery measuring the levels of these nutrients? The unexpected answer arose from the work of researchers studying apparently unrelated applications for RNA molecules. To understand how, one must briefly revisit the ribosome.

RNA World Legacy

PROTEINS MAY BE the wheels, cogs, chutes and conveyor belts that transport and transform raw materials into new cells, but not all of the factory's essential equipment is made of protein. Most notably, the ribosome has a core consisting of the very same nucleotides that make up the mRNA messages it reads. But although ribosomal RNA (rRNA) starts out as a ticker-tape transcript of a DNA blueprint, unlike mRNA it contains no instructions for making something else. Instead the rRNA immediately crumples itself into a defined shape, within which certain nucleotide bases bind to one another, much like a terminator hairpin.

Ribosomal RNA folds on a much grander scale, involving several subunits that are further hardened in places by subtle chemical modifications. Protein staples and struts reinforce its crevices and coat its surface. But atomic-resolution structure studies have revealed that the core of the ribosome, responsible for catalyzing the formation of new bonds between amino acids, is made exclusively of RNA.

TAMI TOLPA

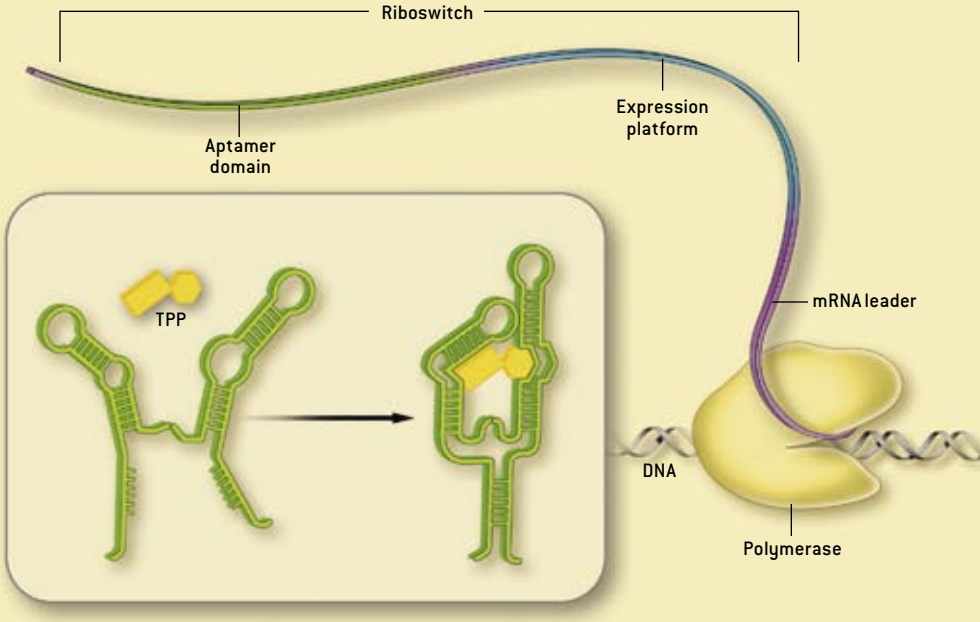
This recent confirmation of an RNA structure with the powers of a protein catalyst was exciting to anyone familiar with a theory about early life advanced in the late 1970s by Harold White III of the University of Delaware. He had observed that many important coenzymes have curious RNA components within their chemical structures. Adenosylcobalamin (coenzyme B_{12}), for example, contains an entire RNA nucleotide, and thiamine pyrophosphate (coenzyme B_1) carries around a piece of sugar-phosphate backbone. These nucleotide bits seem to act as handles for proteins to grasp, and White theorized that they could be vestigial traces from a primordial time when protocells had not yet evolved modern DNA storage or protein synthesis. Instead RNA performed double duty as the information storage molecule and the biopolymer capable of folding into metabolic machines and performing the complex work that today is generally the province of proteins.

By the early 1980s two "living" examples of such ancient RNA elements had been discovered. One of them, RNase P, is an RNA molecule in bacteria that is able to cleave raw RNA transcripts. Another breakthrough identified fascinating RNA sequences that edit themselves out of a longer mRNA transcript, achieving their self-cleavages through a series of chemical bond exchanges. Sidney Altman of Yale University and Thomas R. Cech of the University of Colorado at Boulder received the 1989 Nobel Prize in Chemistry for these separate findings, which demonstrated that RNA molecules—previously seen as only passive messages—could fold into complex three-dimensional structures and accelerate chemical reactions, just like protein enzymes. Collectively, such RNA enzymes, including ribosomes, are termed ribozymes.

In the early 1990s research tools for manipulating biomolecules outside of living cells had matured enough for investigators to experiment with creative uses of RNA's newfound ability to fold itself into complex and functional shapes. In part, scientists were seeking to test the versatility of RNA and thus the plausibility of the "RNA world" hypothesis; they were also looking for new biotechnology applications for ribozymes. Our own group's participation in these pursuits is what eventually led us to look beyond proteins for the mysterious regulators of bacterial coenzyme production.

SELF-DETERMINING SWITCHES

A newfound form of cellular regulation relies on certain RNA copies of genes to supervise themselves. Riboswitches are segments within the leading end of a messenger RNA transcript that are able to gauge the cell's need for the protein encoded by the rest of their message, then rearrange their own shape to control whether that protein is manufactured. Riboswitches therefore have two important domains: an aptamer that senses a specific metabolite (*below*) and an expression platform that affects the mRNA's fate by undergoing one of many possible structural reconfigurations (*right*).



METABOLITE SENSING
An aptamer for the coenzyme thiamine pyrophosphate (TPP) assumes a defined shape (*left*) as it exits the polymerase. When TPP is present, the aptamer binds to it, grasping the molecule tightly (*right*).

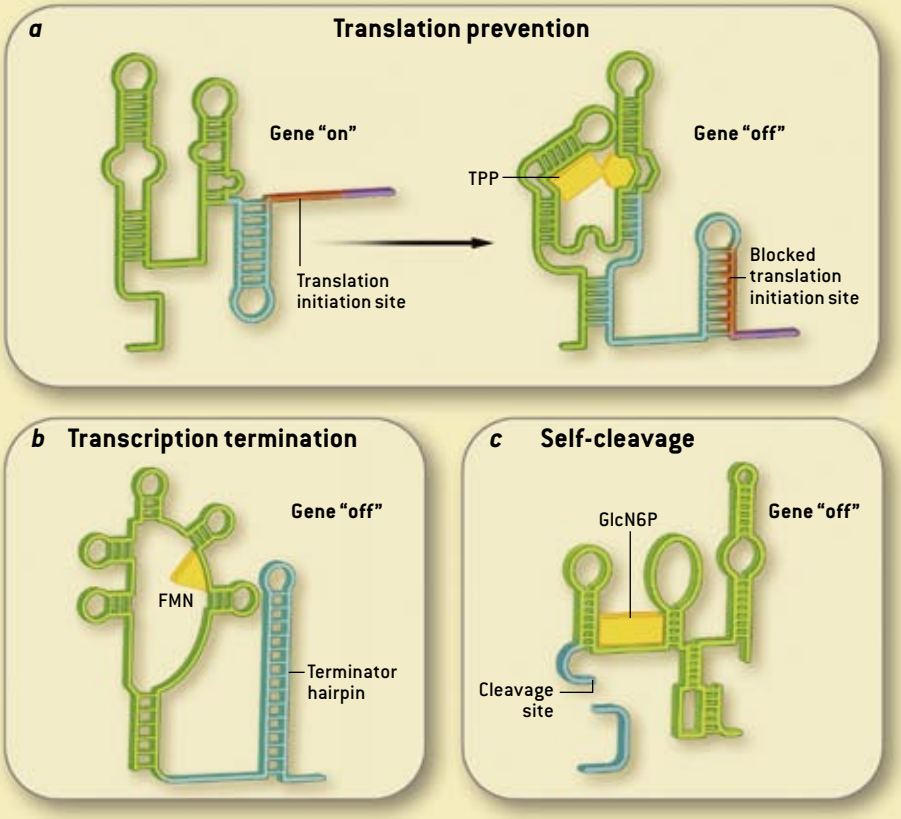
Natural Sensors

THE LABORATORIES of Larry Gold at U.C. Boulder, Gerald Joyce of the Scripps Research Institute and Jack W. Szostak of Massachusetts General Hospital developed a method of test-tube evolution that enabled them to subject trillions of synthetic RNA sequences to a Darwinian test that the "fittest" molecules would pass. Using this test-tube evolution, Szostak's group quickly discovered a variety of short RNA structures that could bind tightly to adenosine triphosphate (ATP), as well as many organic dyes, amino acids and antibiotics.

Szostak christened all these RNA molecules born in laboratories "aptamers," a term derived from the Latin *aptus*, meaning "fitted." And despite their unnatural origins, many aptamers possessed a quality that is more important in a biological context than just binding their target molecule tightly: they reject molecules with closely related structures.

Our laboratory set out to exploit this high selectivity by designing a sensor made of RNA. The plan was to create an aptamer capable of recognizing a target molecule by binding to it and to join

TAMI TOLPA



RIBOSWITCH RESPONSES

Riboswitches employ a variety of strategies to control protein manufacture. When TPP is absent, for example, the expression platform can leave a translation initiation site open to ribosomes, allowing expression of the gene's instructions to remain "on" (*a, left*). When TPP is bound by the aptamer, the expression platform can form a hairpin that blocks translation, turning the gene "off" (*a, right*). A riboswitch sensing the coenzyme flavin mononucleotide (FMN) forms a terminator hairpin that halts transcription of its message by polymerase (*b*). An unusual ribozyme triggered by glucosamine-6-phosphate (GlcN6P) self-destructs by cleaving itself (*c*).

that to a second RNA segment that could signal the event with a visible readout. For the latter role, we chose the "hammerhead" ribozyme. Named for its distinctive-looking structure, it is one of the simplest and most effective natural self-cleaving ribozymes known. We could, for example, attach a fluorescent tag to one end of the hammerhead strand and a so-called quencher group that dampens the fluorescence in close proximity to each other within the RNA's folded structure. Once the aptamer end of our apparatus found and bound the target molecule, self-cleavage by the hammerhead would separate the quencher group from the fluorescent tag, and the molecule would light up as if a lampshade had been removed.

RNA proved so adept at this sensor function that we were later able to develop aptamer-coupled ribozymes capable of sensing and reporting the presence of a wide variety of molecules. Our collection of sensors could be arrayed on a tiny chip and used to accurately detect many different chemical compounds simultaneously, even in a complicated mixture.

Indeed, the ease with which we could create RNAs that sensed small molecules and transduced that binding into a purposeful rearrangement of their own structures made us wonder whether natural evolution had created similar RNAs. Ribozymes from the RNA world were clearly still performing critical tasks in modern organisms. Might there be undiscovered sequences for other important RNA machines hiding in modern genomes?

We started scouring the scientific literature for hints pointing to natural aptamers and found only tantalizing references to noncoding RNA sequences known to be somehow important for cellular regulation. Then our search brought us to the mystery of bacteria and their vitamins. We came across mentions of a protein, BtuB, which is a part of the apparatus for importing coenzyme B_{12} into the *E. coli* bacterium. The mRNA transcript that encodes the BtuB protein begins with a massive leader of 240 noncoding nucleotides, and its extraordinary length was our first clue that it might have an unusual function. Another research group had also already shown that production of BtuB protein was inhibited when B_{12} concentrations in the cell were high. Yet no protein sentry that sensed B_{12} had been discovered.

From the previously published work of others, we knew that the presence of B_{12} somehow prevented ribosomes from binding to BtuB mRNA. One experiment had also hinted that some kind of structural change in the mRNA leader sequence was occurring in the presence of B_{12}. Could it be that the long BtuB RNA leader contained a natural B_{12}-binding aptamer that acted to regulate expression of the instructions encoded in its own gene?

We used a technique called in-line probing to map the parts of the BtuB RNA message that were becoming more structured or more flexible in the pres-

THE AUTHORS

JEFFREY E. BARRICK and *RONALD R. BREAKER* investigated the diversity and importance of riboswitches together in Breaker's laboratory at Yale University. Barrick is now a postdoctoral fellow at Michigan State University, where he is studying the evolution of bacteria as well as self-replicating computer programs. Breaker's group is continuing to explore the nature and uses of nucleic acids, in part by creating designer gene-control elements made of RNA and developing antibiotics to target natural riboswitches.

ence of B_{12} and found, most notably, that a new twist was formed near the beginning of the BtuB mRNA coding region. This structure could explain the inhibition of ribosome binding. The RNA itself seemed to be sensing B_{12} and regulating its transport in the same manner that TRAP regulates the tryptophan transporter message in *B. subtilis*—by preventing the ribosome from translating it. We therefore named this RNA molecule that was able to toggle gene expression from on to off a "riboswitch."

As we were investigating the BtuB leader, another case of unexplained regulation also caught our attention. Previous research had determined that the mRNAs encoding synthesis enzymes and transporters for coenzyme B_1 in diverse groups of bacteria all contained a common stretch of RNA sequence and that mutations in this sequence disrupted the normal repression of these genes in cells that had accumulated sufficient B_1. In *E. coli,* the mRNA of an operon for two synthesis enzymes has a leader containing the sequence near the site where translation of the first protein starts. We were able to show that B_1 induced a structural change in this mRNA such that the ribosome binding site was tightly zippered up. We then determined that a smaller 91-nucleotide domain within the leader could bind to B_1. Like our artificial sensors, this natural riboswitch consisted of a separate aptamer domain linked to a functional "response" sequence that allowed it to regulate whether B_1 would be produced.

Thus, we had found at least two messenger RNAs with the remarkable ability to monitor cellular conditions and make their own decisions about whether the protein machines they encoded were necessary, without intervention by protein supervisors. These paper photocopies were not passive messages; they were folding like possessed origami and choosing their own fates. And those two proved to be more than curiosities. Natural RNA switches that responded to a variety of other fundamental cellular metabolites were latent in the scientific literature and quickly identified by members of our labs and other research groups.

A sequence common to relatives of *B. subtilis* turned out to be a riboswitch that recognizes the coenzyme *S*-adenosylmethionine (SAM). An RNA element known to occur in messages directing the synthesis and transport of the coenzyme flavin mononucleotide (B_2) was another riboswitch. A section of mRNA thought to encode a protein monitor for lysine in *E. coli* was in fact a piece of a complex lysine aptamer that regulated synthesis of this amino acid in a broad range of bacteria. Riboswitches were a widespread form of genetic control.

Reverse-Engineering Riboswitches

A DOZEN CLASSES of riboswitches, defined by their aptamer structures, have been identified so far, and although they vary in certain features and mechanisms, a few general principles have emerged. Riboswitches are messenger RNA transcripts capable of regulating their own gene's expression by controlling whether the message they contain is translated into a protein or destroyed without ever being read by a ribosome. The riboswitch makes this call by monitoring the cell's need for the protein it encodes through its ability to sense a target metabolite and then altering its own structural configuration in response. A riboswitch thus contains two important segments: its metabolite-sensing aptamer domain and its regulatory "expression platform" sequence.

The aptamer serves as a complex receptor for one specific small-molecule metabolite, and in all members of a class the core aptamer structure is the same, even in evolutionarily distant organisms. Riboswitch expression platforms, which can include part of the aptamer domain, contain the sequences that rearrange their own structure to affect gene expression [*see box on pages 54 and 55*]. The B_{12} and B_1 riboswitches we first discovered both have expression

Tempting Targets

Many bacteria, including the human pathogens listed here, employ riboswitches to control the activity of their own genes. Agents that trigger those riboswitches might therefore serve as new antibiotics, particularly if the drugs disrupt the function of genes essential to an organism's virulence or survival. The number of riboswitch classes found in each organism and the number of genes known to be regulated by riboswitches are shown below. Asterisks indicate that at least one vital gene is regulated by a riboswitch.

Human Bacterial Pathogen	Riboswitch Classes	Genes Regulated
Acinetobacter baumannii	4	6
Bacillus anthracis	9	82
Brucella melitensis	5	21*
Enterococcus faecalis	7	17
Escherichia coli	4	15*
Francisella tularensis	4	8
Hemophilius influenzae	5	15*
Helicobacter pylori	1	2
Listeria monocytogenes	9	49
Mycobacterium tuberculosis	3	13
Pseudomonas aeruginosa	3	27
Salmonella enterica	3	34*
Staphylococcus aureus	8	30*
Streptococcus pneumoniae	5	19
Vibrio cholerae	5	13
Yersinia pestis	3	11

SOURCE: KENNETH F. BLOUNT AND RONALD R. BREAKER *Yale University*

platforms that prevent translation initiation by configuring themselves to hide sequences the ribosome needs to recognize a valid work order, for example. Other instances of riboswitches containing these same aptamers have expression platforms that cause premature termination of mRNA transcription by forming a terminator hairpin.

As our group learned more about riboswitches, we began to appreciate how carefully evolution has balanced the gears and springs that animate their mechanisms. For example, metabolite recognition in the cells must occur within the mere seconds that it takes polymerase to stream out an mRNA leader and ribosomes to bind to it and begin translation. Thus, the speed of metabolite binding, not necessarily the strength, is critical for determining whether a riboswitch can sense its target. Timing sequences between the aptamer and expression platform, which cause polymerase to briefly stall, are sometimes necessary to introduce delays that give the aptamer enough time to capture a metabolite and properly rearrange its expression platform.

When we started scanning bacterial genomes looking for new examples of riboswitches, we found still more surprises. In the *B. subtilis* genome alone, we identified eight new sequences with the hallmarks of riboswitches. One of these was a riboswitch with a double aptamer that acted to turn gene expression on rather than off. Another, as it turned out, was not just a riboswitch but a metabolite-triggered ribozyme. Rather than undergoing a structural reconfiguration, this molecule's expression platform self-cleaved—in essence, self-destructing before its message could be translated.

Just one of the riboswitch classes discovered to date has been seen in multicellular organisms; the rest, as far as we know, are found only in bacteria. The genomes of higher organisms have more complicated genetic regulation than bacteria, and the route from blueprints to protein is also more circuitous. Instead of tidy mRNA photocopies, first-draft gene transcripts often have huge chunks of noncoding text, known as introns, which must be spliced out before the message is translated into protein. We found a riboswitch on the cutting-room floor.

The coenzyme B_1 aptamer occurs in the sequences of introns within thiamine synthesis operons in many fungi and plants, including rice. When bound to B_1, these riboswitches appear to cause the RNA structure around the intron junctions to be rearranged, preventing splicing from proceeding. Although the details are not clear, this may target the entire message for the rubbish bin or prevent it from being moved to the correct part of the cell to be translated.

Intriguingly, a known antifungal drug has also been found to bind the B_1 riboswitch. Evidence suggests that it thereby tricks a fungus into thinking that it has enough B_1, repressing synthesis of more. Because the fungus does not actually have this important nutrient, its growth is slowed and eventually it can die from the deficiency. As this example illustrates, riboswitches are such vital regulators of critical nutrient supplies in a range of microbes that they also make tempting targets for new antibiotics.

More than a dozen human pathogens are now known to rely on riboswitch regulation of several important metabolites [*see box on opposite page*]. Many researchers are working to find molecules that can fool bacterial riboswitch aptamers into mistaking them for a natural metabolite and thereby trigger a gene regulatory response that would be deleterious to the cells.

Some research groups are also exploring the idea of using artificial riboswitches to control genes inside living cells—for example, in the context of gene therapy. The goal is to design an on-off switch triggered by a benign druglike molecule and incorporate that into a therapeutic gene. The construct could then be inserted into a patient's cells and regulated by having the person take a pill containing the molecule that activates the designer riboswitch. As with antibiotic applications, this use of riboswitches is still in early stages of investigation.

The general feeling of surprise and excitement that was inspired by ribozymes and the efforts to apply those ancient molecules to modern uses is renewed by the very existence of riboswitches. Only bits and pieces of the lost RNA world seem to be with us today, but these RNA devices with complex mechanisms and regulatory roles have tenaciously hung on in modern organisms. We cannot help but wonder whether riboswitches are the last vestiges of the RNA world that will be revealed or if other primordial molecules are still employed by the metabolic factories or administrative offices in modern cells—perhaps even our own human cells—and awaiting discovery. SA

Only bits and pieces of the LOST RNA WORLD seem to be with us today.

MORE TO EXPLORE

The Origin of Life on the Earth. Leslie E. Orgel in *Scientific American,* Vol. 271, No. 4, pages 76–83; October 1994.

Thiamine Derivatives Bind Messenger RNAs Directly to Regulate Bacterial Gene Expression. Wade Winkler, Ali Nahvi and Ronald R. Breaker in *Nature,* Vol. 419, pages 952–956; October 31, 2002.

Metabolite-Binding RNA Domains Are Present in the Genes of Eukaryotes. Narasimhan Sudarsan, Jeffrey E. Barrick and Ronald R. Breaker in *RNA,* Vol. 9, No. 6, pages 644–647; June 2003.

Riboswitches as Antibacterial Drug Targets. Kenneth F. Blount and Ronald R. Breaker in *Nature Biotechnology* (in press).