Review

Use of a coenzyme by the glmS ribozyme-riboswitch suggests primordial expansion of RNA chemistry by small molecules

Adrian R. Ferré-D’Amaré*
Laboratory of RNA Biophysics and Cellular Physiology, National Heart, Lung and Blood Institute, 50 South Drive, MSC-8012, Bethesda, MD 20892-8012, USA

The glmS ribozyme-riboswitch is the first known example of a naturally occurring catalytic RNA that employs a small molecule as a coenzyme. Binding of glucosamine-6-phosphate (GlcN6P) activates self-cleavage of the bacterial ribozyme, which is part of the mRNA encoding the metabolic enzyme GlcN6P-synthetase. Cleavage leads to negative feedback regulation. GlcN6P binds in the active site of the ribozyme, where its amine could function as a general acid and electrostatic catalyst. The ribozyme is pre-folded but inactive in the absence of GlcN6P, demonstrating it has evolved strict dependence on the exogenous small molecule. The ribozyme showcases the ability of RNA to co-opt non-covalently bound small molecules to expand its chemical repertoire. Analogue studies demonstrate that some molecules other than GlcN6P, such as L-serine (but not D-serine), can function as weak activators. This suggests how coenzyme use by RNA world ribozymes may have led to evolution of proteins. Primordial cofactor-dependent ribozymes may have evolved to bind their cofactors covalently. If amino acids were used as cofactors, this could have driven the evolution of RNA aminoaconlylation. The ability to make covalently bound peptide coenzymes may have further increased the fitness of such primordial ribozymes, providing a selective pressure for the invention of translation.

Keywords: catalytic RNA; enzymatic cofactor; RNA world

1. INTRODUCTION

The glmS ribozyme is a catalytic RNA derived from a ligand-dependent self-cleaving mRNA domain conserved among Gram-positive bacteria (reviewed in [1]). In these organisms, the 5′-untranslated region (5′-UTR) of the mRNA encoding the enzyme glucosamine-6-phosphate (GlcN6P) synthetase self-cleaves when it binds to GlcN6P [2]. Self-cleavage of the 5′-UTR leads to degradation of the mRNA, resulting in negative feedback regulation of the protein enzyme by its product metabolite [3]. GlcN6P synthetase is an essential enzyme that catalyzes the first committed step in the metabolic pathway that leads to synthesis of the bacterial cell wall. The activity is universally present, but is catalysed by two types of protein enzymes distributed in different taxa [4]. Eukaryotes and Gram-negative bacteria have a GlcN6P synthetase that is allosterically controlled by its reaction product. Gram-negative bacteria such as Escherichia coli also regulate expression of the enzyme at the mRNA level using small antisense RNAs. The GlcN6P synthetase of Gram-positive bacteria, in contrast, is not an allosteric enzyme. Thus, the regulation of protein expression at the level of mRNA by the ribozyme-riboswitch domain is likely to be important [1]. Indeed, it has been shown that replacement of the ribozyme by a catalytically compromised mutant results in abrogation of sporulation by Bacillus subtilis [3].

In vitro, the self-cleaving domain of the glmS mRNA can be engineered to function as a multiple-turnover catalyst, that is, a ribozyme. Deletion analysis showed that the ribozyme is comprised of approximately 150 nucleotides (nt) and that a single nucleotide 5′ of the cleavage site suffices for maximum GlcN6P-induced activity. Furthermore, approximately 75 nt RNA constructs (with one nt 5′ of the cleavage site) are responsive to GlcN6P, but exhibit reduced catalytic activity [2]. Structure determination of the full-length glmS ribozyme [5] demonstrated that it is composed of a doubly pseudoknotted approximately 75 nt core domain that encompasses the cleavage site and the GlcN6P-binding site, and a peripheral element that appears to stabilize the core, fully explaining the biochemical results (figure 1).

2. MECHANISM OF GLUCOSAMINE-6-PHOSPHATE ACTIVATION OF THE RIBOZYME

GlcN6P could activate the glmS ribozyme by two mechanisms. The small molecule could function as an allosteric effector, its binding to an allosteric site leading to a conformational rearrangement that activates the ribozyme. Alternatively, the small molecule could function as a catalytic cofactor or coenzyme, binding to the active site of the ribozyme and providing...
GlcN6P binds to the RNA with its amine group in van der Waals contact with the scissile phosphate; that is, the GlcN6P-binding site is the active site of the ribozyme, not an allosteric effector site [5]. The functional importance of the amine group of GlcN6P for glmS ribozyme activity is underscored by studies using the isosteric sugar, glucose-6-phosphate (Glc6P). This compound, which differs from GlcN6P only in having a hydroxyl rather than an amine group at position 2 of the pyranose ring, is not an activator of the ribozyme [12]. Cocrystal structures show that it binds precisely in the same location as GlcN6P. As expected from the foregoing, the conformation of the Glc6P-bound ribozyme is identical to those of the ribozyme in other functional states [6]. Biochemical studies show that Glc6P competes with GlcN6P for binding to the glmS ribozyme, functioning in effect as an antagonist [12]. Since the ribozyme is pre-folded but completely inactive in the absence of GlcN6P, the most parsimonious conclusion is that this compound functions as a catalytic cofactor of the RNA.

3. PLASTICITY OF NATURAL AND ARTIFICIAL GENE-REGULATORY RNAs

The glmS ribozyme is also a riboswitch, that is, an mRNA domain that controls gene expression in cis in response to the intracellular concentration of a small molecule metabolite or second messenger (reviewed in [13,14]). However, its mechanism of action and its structural rigidity make it an atypical riboswitch. Five distinct mechanisms of gene regulation have been documented for riboswitches. First, the majority of known bacterial riboswitches function by transcriptional attenuation. The ligand-binding (or ‘aptamer’) domain of the riboswitch and a ρ-independent terminator stem-loop compete for folding, with ligand binding altering the efficiency of formation of the latter. (Aptamers are RNAs evolved in vitro to bind to specific ligands [15].) Second, many bacterial riboswitches function at the level of translation initiation. In these, the ligand-bound conformation of the ligand-binding domain (the ‘aptamer’ domain) of the riboswitch competes with another that either occludes or exposes the Shine–Dalgarno element. Third, ligand binding by some riboswitches results in alternative splicing of introns. The thiamine pyrophosphate (TPP)-responsive thi-box riboswitch is the only riboswitch thus far described in eukaryotes. It is present in some introns in algae, fungi and plants, and it controls alternative splicing by exposing splice junctions or splicing enhancers depending on TPP concentration (reviewed in [16]). In bacteria, group I self-splicing introns have been described [17] that select between two alternative splice sites based on binding of cyclic di-γ-glutamate (c-di-GMP). Fourth, an S-adenosylmethionine (SAM)-responsive riboswitch has been described that initially functions through transcriptional attenuation, but in which the prematurely terminated transcript induced by ligand binding also functions in trans as an antisense RNA [18]. Finally, the glmS riboswitch is the first known example of a riboswitch that functions by ligand-induced self-cleavage.
Whereas the first four mechanisms of riboswitch action require that the riboswitch (comprised of aptamer domain and the ‘expression platform’, the RNA sequences that interface with the transcription, translation or splicing machinery) adopt at least two mutually exclusive conformations, the glmS riboswitch is rigid and fully assembled prior to GlcN6P binding, and it is only its catalytic activity that switches (by approximately a million-fold).

4. LIGAND SELECTIVITY OF THE glmS RIBOZYME
Riboswitches and artificial aptamers that recognize their ligands by partially or wholly enveloping them are known to undergo folding transitions upon ligand binding (e.g. [19–25]). Conversely, the rigid and pre-folded glmS ribozyme binds GlcN6P in an open, solvent-accessible pocket (figure 1) [5]. Analogue studies [12] underscore the importance of three of the functional groups of GlcN6P that interact with the glmS ribozyme: the anomeric hydroxyl, the amine and the phosphate. A variety of compounds that present vicinal amine and hydroxyl groups are weak activators of the glmS ribozyme. Such (presumably adventitious) activators include glucosamine (GlcN), L-serine, serinol, Tris and ethanolamine (figure 2). Importantly, the apparent pKα of the ribozyme reaction tracks the pKα of the amine group of the activating compound [12]. Thus, the pKα of the reaction rises from 7.9 to 8.7 when GlcN is replaced with serinol (the pKα values of their amine groups are 7.8 and 8.8, respectively). This confirms the importance of the amine group of the activator, suggested by the intimate contact between the amine of GlcN6P and the scissile phosphate [5,6,8]. However, analogues also point to the importance of the vicinal hydroxyl group, and to the precise stereochemical relationship between the amine and the hydroxyl groups. Thus, ammonia and methylamine are not activators, and d-serine, in which the stereochemical arrangement of the two functional groups is opposite that of GlcN6P (and L-serine) is not an activator [12]. The ability of GlcN or glucosamine-6-sulphate to serve as weak activators of the glmS ribozyme also indicates that the phosphate of GlcN6P is not chemically essential, even though it is important for increased binding affinity to the RNA. In addition to interacting with the RNA through outer-sphere coordinated metal ions, the phosphate of GlcN6P receives a hydrogen bond from the N1 imine of riboswitch residue G1 [5]. It was found that replacing G1 with purines with a non-protonated N1 position (such as A, 2-aminopurine, or dimethyladenine) resulted in riboswitches more strongly activated by GlcN than GlcN6P, presumably because of the clash between the phosphate of GlcN6P and the unprotonated N1 of these purines [6].

RNA cleavage proceeds through a concerted (SN2) transesterification in which the 2′-OH of the nucleotide preceding the scissile phosphate (residue −1) attacks the phosphorus of the scissile phosphate. A transition state featuring a pentacovalent phosphorus is resolved by departure of the 5′-oxo group of residue +1, and formation of a 2′,3′-cyclic phosphate (figure 3). A variety of strategies are employed by RNA to catalyse this reaction. The HDV ribozyme employs a water molecule chelated by a magnesium ion to function as a specific base to deprotonate the nucleophilic 2′-OH, and the N3 amine of a cytosine residue (C75) to serve as a general acid, to protonate the leaving group [27,28]. In order that it functions effectively as a general acid, the pKα of the N3 of C75 is perturbed by over 2 pH units from its undisturbed value of approximately 4.2. Indeed, C75 is responsible for most of the catalytic rate enhancement achieved by the HDV ribozyme (figure 4) [29]. The hairpin ribozyme has two nucleobases, G8 and A38, positioned to serve as general base and general acid in the cleavage reaction [30,31]. This ribozyme, unlike the HDV or glmS ribozymes, also catalyzes the reverse ligation reaction. For catalysis of the ligation, the catalytic role of the nucleotides has to be reversed (because of the principle of microscopic reversibility). Raman spectroscopic studies of crystalline hairpin ribozymes show that the pKα of A38 is perturbed by nearly 2 pH units [32]. In addition, crystallographic studies show that G8 and A38 have the ability to preferentially bind to the transition state of the reaction, thus lowering the activation energy by hydrogen bonding [33]. The hammerhead ribozyme appears to use yet another strategy, in which N1 of G12 functions as a general base to deprotonate the nucleophilic 2′-OH of the cleavage reaction, and the 2′-OH of G8 functions as a general acid [34]. Like the hairpin ribozyme, the hammerhead ribozyme also catalyses ligation during which the role of the catalytic functional groups must be reversed. Importantly, it can be shown that near-neutral pKα values are not a strict condition for functional groups to serve as general acid–base catalysts [35,36].
The location of the amine group of GlcN6P in the active site of the glmS ribozyme is analogous to that of the N3 of C75 in the HDV ribozyme and A38 in the hairpin ribozyme [5,30,37]. This suggests that the small molecule would function as a general acid catalyst (figure 4). The simplest interpretation of the catalytic inactivity of the glmS ribozyme in the absence of GlcN6P is that the small molecule is responsible for all the catalytic power of the ribozyme. However, mutational analysis does not support this. All known glmS ribozyme isolates have a G residue at position 40 (Thermoanaerobacter tengcongensis numbering scheme). G40 is positioned with its N1 imine within hydrogen bonding distance of the nucleophilic 2'-OH of A – 1. Thus, it is possible that G40 functions as a general base in catalysis, or that it serves to orient the nucleophilic 2'-OH during the reaction. Mutation of G40 to A completely abrogates the activity of the ribozyme, underscoring its importance. Structure determination of the G40A mutant reveals that a ribozyme folds into a structure indistinguishable from that of the wild-type with, most surprisingly, GlcN6P bound in precisely the same position as in the active ribozyme [7]. This demonstrates that while GlcN6P is necessary for glmS ribozyme activity, it is not sufficient. It also indicates that whatever the precise role of G40 in catalysis might be, it becomes manifest only when GlcN6P binds the ribozyme. Thus, the active site of the glmS ribozyme is not simply a rigid, passive scaffold for GlcN6P binding. Rather, the RNA and the small molecule appear to tune each other’s chemical properties to give rise to catalysis. Two molecular dynamics studies reached conflicting conclusions regarding the potential catalytic role of G40 [38,39]. The exact nature of this coupling between the glmS ribozyme active site and GlcN6P remains unresolved.
The mechanism of action of the COFACTORS?

7. AMINO ACIDS AS PRIMORDIAL RIBOZYME

chemical similarity to nucleotides (amine sugars or which can use cofactors that do not have obvious
does not fold concomitant with ligand binding and
which are remnants of the active sites of such ancestral
ribozymes.) [48].

Noller [47] has suggested that the fact that many coenzymes
are remnants of the active sites of such ancestral
ribozymes. Structural and biochemical studies of riboswitches
and aptamers show that many of these RNAs undergo
large conformational rearrangements concomitant
with ligand binding. Likewise, the plasticity of RNA
is a recurrent theme in RNA–protein interactions
(e.g. [45,46]). Noller [47] has suggested that the fact that many peptide-binding RNAs display large ligand-driven folding transitions may hint at the primordial driving force for the invention of translation. Specifically, peptides could have first served as facilitators for the folding of primordial ribozymes, and translation evolved to make synthesis of such peptides more efficient and reproducible. The successful design of synthetic aptazymes, artificial catalytic RNAs that require binding of a small molecule (by an aptamer domain) in order to achieve catalytic activity, demonstrates that ligand-induced folding of RNA can be functionally coupled to the catalytic competence of ribozymes [48].

8. COVALENT ATTACHMENT OF COFACTORS TO RNA AND THE EVOLUTION OF TRANSLATION

Others have speculated that tRNA aminoacylation originated either as a means of conferring a replicative advantage to a primordial tRNA-like genomic tag [52] or as a form of post-transcriptional RNA modification that would confer new structural or functional capabilities to the nucleic acid [53]. If, as suggested above, primordial metabolism was catalysed by a set of ribozymes some of which evolved dependence on amino acids functioning as coenzymes, then RNA aminoacylation (and the subsequent evolution of translation) would have arisen as a consequence of the evolutionary pressure to improve the efficiency of such ribozymes by covalent attachment of their cofactors. In vitro selection experiments demonstrate that even fairly simple ribozymes can carry out aminoacylation with high specificity and regioselectivity (e.g. [54]).

In summary, the glmS ribozyme is a small-molecule-dependent catalytic RNA that is widespread among Gram-positive bacteria, where it regulates a key metabolic pathway. It employs the small molecule

Phil. Trans. R. Soc. B (2011)
RNA world in which translation evolved. This use of a free amino acid as a coenzyme by a ribozyme suggests how the use of amino acids in modern biology may have derived from the primordial use of these small molecules as coenzymes in the RNA world, and that the specific set of amino acids used throughout contemporary biology may reflect the ‘basis set’ of amino acid cofactors shared among the ribozymes of the RNA world in which translation evolved.

The author thanks past and current members of the Ferré-D’Amare laboratory for their many contributions, and in particular D. Klein for his work on the glmS ribozyme. This research was supported by the intramural programme of the National Heart, Lung and Blood Institute, National Institutes of Health.

REFERENCES


