Structure Determination of Catalytic RNAs and Investigations of Their Metal Ion-Binding Properties

Michèle C. Erat and Roland K.O. Sigel*

Abstract: Naturally occurring RNA molecules exhibit many unexpected and fascinating properties in living cells such as protein synthesis and transport, regulation of metabolic functions, and catalytic cleavage reactions. To understand this functional diversity, a detailed knowledge of RNA structure and metal ion-binding properties is crucial. In our research group, we address these problems by combining various biochemical, analytical and spectroscopic techniques. A large part of our work is devoted to the structure determination of catalytic RNA molecules, i.e. ribozymes, by NMR. Based on the three-dimensional structure, further experiments are carried out to understand in detail the effects of different metal ions on the local and global structure, as well as catalysis itself.

Keywords: Affinity constants · Metal ions · NMR · Ribozymes · RNA structure

1. Introduction

The research in our laboratory focuses on the interaction of metal ions with RNA, i.e. which roles different M⁺⁺ have on folding, structure, and catalytic properties of ribonucleic acids. Like DNA, RNA is a polyanion under physiological conditions, with every bridging phosphate group adding one negative charge to the overall charge. It follows that nucleic acids cannot occur without metal ions being present. RNA fulfills many functions in all living organisms and new roles of these fascinating molecules are discovered every year. The best known RNAs are certainly mRNA and tRNA, but also the ribosome itself is mainly composed of RNA.

Other functions carried out by RNAs include protein transport, splicing, i.e. the removal of non-coding intron sequences, the regulation of metabolic functions, and even the regulation of the cell cycle itself [1]. Some of these RNAs are not only ‘passively’ exhibiting their function by ‘just’ binding to other biomolecules, but perform also catalytic reactions. These so-called ribozymes were discovered nearly 25 years ago [2][3]. Nowadays eleven different classes of naturally occurring ribozymes are known, two of them having been discovered only last year [4][5].

Our research concentrates on two classes of the naturally occurring ribozymes: (i) riboswitches and (ii) group II intron ribozymes. The emphasis of this short review is on our structural and analytical NMR studies on group II intron ribozymes; our riboswitch work is only briefly indicated in the following Section.

2. Riboswitches

Riboswitches are highly conserved RNA sequences occurring for example in the 5'-UTR (see Abbreviations) of certain mRNAs. They bind small metabolites like FMN, SAM, TPP, guanine, adenine, and coenzyme B₁₂ with a high degree of affinity and specificity [6][7]. Upon binding of the corresponding metabolite, the riboswitch sequence presumably undergoes structural changes [8] and in some cases even cleavage of the RNA strand occurs [4]. Both cleavage and structural changes inhibit binding of the 5'-UTR to the ribosome, and thus prohibit translation to the corresponding protein. As the protein(s) expressed by a mRNA is always involved in either biosynthesis or transport of the metabolite that binds to the riboswitch, this new class of RNA molecules constitutes an elegant way of feedback control within living cells.

In our group, we work with the so-called B₁₂ riboswitch. This riboswitch consists of an about 200 nt long and highly conserved sequence located in the 5'-UTR of mRNAs associated with the metabolism and transport of coenzyme B₁₂ and vitamin B₃₂, respectively [9]. Our investigations encompass structural studies as well as binding studies of coenzyme B₁₂ and its derivatives to this riboswitch sequence in the presence of different metal ions.

3. Group II Intron Ribozymes

The main efforts of our research are presently devoted to group II intron ribozymes. These large molecular machines are self-splicing introns, i.e. capable of cutting themselves out of the precursor RNA and in a second step joining the adjacent exons to yield the final mRNA – all without the aid of proteins (Fig. 1). In addition, these amazing molecules can insert themselves again into RNA and even DNA. The occurrence, reaction mechanisms, and the different aspects of metal ion binding to group II introns have lately been reviewed [10][11], thus only a few important points are summarized below:
3.1. General Remarks on Group II Intron Ribozymes

Group II introns are defined (i) by their secondary structure consisting of six domains radiating out from a central wheel, (ii) by a set of highly conserved nucleotides and (iii) by tertiary contacts spread throughout the whole intron (Fig. 2). These catalytic RNAs are found in organellar genes of plants, fungi and lower eukaryotes as well as in bacteria [10]. Some group II introns contain an ORF encoding for a maturation protein with a reverse transcriptase and a nucleotide excision domain [10]. Together with the maturase, these introns act as mobile genetic elements that can insert themselves into DNA at homologous and even heterologous sites.

Based on phylogeny, group II introns are thought to be very old and to have played an important role in eukaryotic evolution [10]: Including nuclear introns, the transposable LINE elements and spliceosomal RNA, as much as 30% of the human genome might stem from ancestral group II introns [12]. Group II introns still show a rather close sequence relationship to the spliceosome, the eukaryotic splicing machinery that consists of five RNAs and >100 proteins and is responsible for the accurate removal of all introns. It is surprising to see that despite their evident importance in evolution and their direct role in the correct expression of proteins in many organisms, relatively little is known about the structure and catalytic mechanisms of group II intron ribozymes.

To perform the self-splicing reaction, group II introns fold into a compact three-dimensional structure. Out of the six domains, only domains 1, 5, and 6 (D1, D5, D6) together with the linker regions are essential for the complete splicing reaction: D1 binds the exon sequences and provides the scaffold for the three-dimensional structure, D5 is viewed as the catalytic center, and D6 harbors the branch adenosine, the 2’-OH of which serves as the nucleophile in the first step of splicing (Fig. 1). To perform catalysis, group II introns have a very specific need for metal ions as cofactors. Monovalent (usually Na+ or K+) as well as divalent metal ions (usually Mg2+) are intrinsically tied to all RNAs for charge compensation. Without metal ions, the high negative charge of the sugar-phosphate backbone would prevent any higher-order folding of RNA molecules. In addition, at least for the larger ribozymes like the group II introns, more and more experimental evidence suggests a direct involvement of divalent metal ions in catalysis [11]. The little that is known about metal ion binding in group II introns has been reviewed [11] and is therefore only briefly described in the next Section.

3.2. General Aspects of Metal Ion Binding to RNA

Metal ion binding to RNA molecules differs considerably from binding of metal ions in metalloproteins [13]:

(i) Generally, proteins do not carry a large charge. Some domains may be charged, but folding to the correct three-dimensional structure is predominantly determined by a hydrophobic collapse. In contrast, RNAs are highly negatively charged, as they contain one phosphodiester per nucleotide. It follows that folding can only occur in the presence of suitable counter ions, which are predominantly alkaline ions as well as Mg2+ as these are most abundant in living cells.

(ii) Proteins consist of 20 amino acids. Among them, only four (Cys, Asp, Glu, His) comprise the majority of all metal ion binding sites. In contrast, neglecting all post-transcriptional modifications, RNA consists of only four nucleotides, each of which contains several possible metal ion-coordinating sites. RNA-metal ion binding is usually weak, and as a consequence often poorly defined. Nevertheless, RNA can achieve a surprisingly high specificity with respect to the sites of metal coordination as well as the kind of metal ion involved: For example, small amounts of Ca2+ efficiently compete with Mg2+ ions and inhibit the splicing reaction of the group II intron al5γ in vitro splicing [14][15].

(iii) Depending on their metabolic function, proteins bind and use a large variety of different metal ions, ranging from alkaline and alkaline earth metal ions to numerous other metal ions [16]. In contrast, in naturally occurring RNAs, only Na+, K+ and Mg2+ have so far been identified as natural cofactors.

(iv) Mg2+ is a kinetically labile metal ion. Thus, if an excess of possible coordinating atoms is present, Mg2+ may not be fixed to the same atoms at all times. In addition, inner-sphere and outer-sphere, i.e., direct and water-mediated binding is possible.

Taken together, the four points mentioned above lead to the following two ma-
for the detection of $M^{2+}$ binding sites in RNA: (a) How to detect the few specific binding sites within an ocean of unspecifically bound ions of the same kind? (b) How to detect $K^+$ or $Mg^{2+}$, both of which are spectroscopically silent?

We apply a combined approach of methods from different research fields to address and answer these questions regarding the role of metal ions in structure and folding of RNAs [15][17–19].

4. NMR Structure Determination

In order to investigate metal ion binding to RNAs, structures at atomic resolution are needed. We follow a dual approach to achieve this challenging goal, one being X-ray crystallography (in collaboration with Dr. Eva Freisinger in our Institute) and the other NMR spectroscopy. NMR spectroscopy is a powerful tool to study the structure, tertiary interactions, dynamics and requirements for metal ions of RNA molecules. Usually, RNAs of a size between 20 and 40 nucleotides are studied by NMR. Nevertheless, recently a 30 kDa dimeric RNA structure (86 nt) has been solved at high resolution setting the future direction of NMR structure determinations of nucleic acids [20]. The fact that this size corresponds to only about one third of what is achieved with proteins nowadays can be attributed to the small number of protons present, to the lack in diversity of residues and consequently to the heavy overlap of resonances (Fig. 3A).

As an entire group II intron is too large to be structurally examined by NMR, we investigate individually the different domains involved in catalysis [17][19]. First, the desired RNA is synthesized in large amounts by *in vitro* transcription, as described in our other contribution in this CHIMIA issue [21]. For a complete structure determination by NMR, several samples of 0.5–1 mM RNA in 220 μl 100% $D_2$O or 90% $H_2$O/10% $D_2$O are needed (both natural abundance as well as isotopically $^{13}$C, $^{15}$N enriched RNA).

First, the best conditions for obtaining high-quality NMR spectra need to be determined by optimizing factors such as the RNA sequence, temperature and pH as well as concentrations of monovalent and possibly divalent metal ions. Also the RNA concentration can be crucial, because on the one hand, higher concentrations lead to shorter measurement times, but on the other hand they can also lead to dimerization of a hairpin sequence, giving a double-stranded structure. Dimers can often be removed by gel filtration [21] or circumvented by dilution of the RNA into a large volume (*e.g.* 100 ml) of water, followed by heating to 95 °C for up to one hour and subsequent slow cooling to room temperature and concentration of the RNA by ultrafiltration. In this initial phase of a structure determination mainly 1D spectra in $D_2$O and 90% $H_2$O are recorded to obtain a first indication on the sample quality.

Once optimized conditions are found, 2D $^1$H–$^1$H-NOESY spectra in $D_2$O (Fig. 3) and 90% $H_2$O are recorded. In these spectra Nuclear Overhauser crosspeaks (NOE) are observed between protons that are closer in space than 5–6 Å, yielding distance constraints that are subsequently needed for structure calculation. A key step in assigning all resonances is the so-called ‘sequential walk’ (Fig. 3B): The aromatic nucleobase-protons are close in space not only to their intra-residue ribose H1’, but also to the H1’ of the upstream 5’-nucleotide (Fig. 3C). Thus, the connectivity of nucleotides along a RNA strand can be established. Several recent reviews give a good overview on the exact methodology of resonance assignments for RNAs [22–26]. For full assignment of all peaks and verification thereof, a series of 2D $^1$H–$^1$H-NOESY spectra in $D_2$O as well as in $H_2$O are recorded at different temperatures and mixing times. Other special multidimensional NMR experiments give further information: $J_{\text{RDC}}^\text{N}N$ COSY experiments reveal the existence of hydrogen bonds, residual dipolar couplings (RDCs) yield orientational restraints [27], and HCCH-COSY/TOSY experiments are needed to identify the spin system within the sugar residues. For the latter experiments, $^{15}$C, $^{15}$N-enriched RNA is required. At the University of Zürich, we are in the lucky position to have access to an excellent NMR facility with high-field NMR machines ranging from 500–700 MHz, including cryoprobes to enhance sensitivity, enabling us to take advantage of the full diversity of possible experiments.

Once a set of distance and conformational restraints is collected, structure calculation is started from an initial structure by a fully computational simulated annealing protocol (X-PLOR NIH package [28]), yielding a first set of structures. These output structures are analyzed for distance and conformational violations, the input files corrected and supplemented with further restraints, before a next round of simulated annealing is performed (Scheme). This cycle is repeated until at least 10% of the output structures converge (*e.g.* 20 out of 200 calculated structures) [17][19]. Usually, a r.m.s.d. of 1–3 Å for all heavy atoms of such an ensemble of 20 structures is obtained. This r.m.s.d. is larger than for comparable protein structures, but one should keep in mind that RNA is not as compact as a protein and thus much more flexible.

After the determination of the three-dimensional structure of a RNA molecule, we focus on its binding properties to metal ions and their effect on local structures.

5. Investigation of $M^{n+}$ Binding to RNA

As mentioned above, the metal ions usually associated with nucleic acids ($Na^+$, $K^+$, and $Mg^{2+}$) are spectroscopically silent and thus difficult to detect. Nevertheless, these
are several ways to examine metal ion binding to RNA. Recently we have shown that group II introns possess numerous specific Mn\(^{+} \) binding sites within the catalytic core of the folded RNA molecule by partially replacing Mg\(^{2+}\) with Tb(\(^{III}\)) [18][29]. Tb(\(^{III}\)) replaces Mg\(^{2+}\) at its binding sites and leads to a hydrolytic cleavage of the phosphate sugar backbone, when bound in the minor groove. If Mg\(^{2+}\) is replaced with Fe\(^{2+/3+}\), cleavage occurs through a radical mechanism (‘Fenton cleavage’) and thus irrespective of the binding site [30].

A further way to detect metal ion binding are so-called thio-rescue experiments. Here, one of the non-bridging phosphate oxygens is replaced by a sulfur. If Mg\(^{2+}\) binding to this site is essential for catalysis (or folding), the activity of the ribozyme is reduced due to the lower affinity of Mg\(^{2+}\) towards the softer sulfur ligand [31]. Upon addition of more thiophilic metal ions like Cd\(^{2+}\) or Zn\(^{2+}\), which coordinate readily to sulfur, ribozyme activity may be restored [32]. Indeed, such experiments are not unequivocal and often rescue is observed also with metal ions that are hardly more thiophilic than Mg\(^{2+}\), like Mn\(^{2+}\), but which have a higher overall affinity [31]. To gain a better understanding on the effect of sulfur within RNA oligonucleotides on structure, hydrogen bonding, and metal ion binding, we perform NMR studies on such modified RNAs [33].

NMR spectroscopy enables us to directly monitor Mg\(^{2+}\) binding to RNA [17][19]. Addition of increasing amounts of MgCl\(_2\) to a RNA solution leads to changes in chemical shifts of specific protons (Fig. 4B). These changes can either be caused by direct Mg\(^{2+}\) binding at a neighboring atom, or by a slight structural change in local geometry upon Mg\(^{2+}\) binding nearby. By fitting the experimental data to a 1:1 binding model, affinity constants for Mg\(^{2+}\) binding can be calculated [17][19]. We could show that RNA hairpins like D5 [29] or D6 [17] bind 3–5 Mg\(^{2+}\) ions at specific sites with a rather high affinity. Interestingly, some of these sites are very close to the catalytically important nucleotides implying a possible catalytic role of these ions. Considering the charge repulsion between metal ions close in space, it is remarkable that nucleic acids have the ability to bind metal ions to several sites in close proximity, as we could show by potentiometric pH titrations [34].

Metal ions not only bind passively to RNA but can lead to a substantial decrease in pK\(_a\) values of nucleotide functional groups [35], a change in hydrogen bonding properties [35], or perhaps even a structural change. We investigate these various effects by fluorescence spectroscopy [36], Fig. 4. Solution structures and metal ion binding properties of catalytic domains of group II intron ribozymes. A) Schematic secondary structure of a group II intron together with the solution structures of D5 [19] and part of D6 [17], as solved in our laboratory (PDB codes 1R2P and 2AHT). B) Change of the chemical shift of a H\(_2\) proton within D6 upon titration with MgCl\(_2\). The experimental data is fit to a 1:1 binding model to give the affinity constant \(K_A = 112 \pm 26 \text{ M}^{-1}\) (log \(K_A = 2.05 \pm 0.10\)).
biochemical activity assays [14][15], potentiometric pH titrations [37], as well as by NMR [17][19]. By applying this combined approach we make use of the individual advantages that each of these techniques offers. Ideally, we thereby achieve a full characterization of the role of specific metal ions in the structure and function of our RNA.

6. Concluding Remarks

In contrast to the traditional view of ribonucleic acids as mere working copies of the genome, rapidly increasing experimental evidence attributes a large variety of specific cellular functions to RNA, including protein transport and synthesis, metabolic regulation, and even catalytic reactions [1]. The polyanionic character of RNA makes charge compensation indispensable. This is usually accomplished by weak and unspecific binding of monovalent and divalent metal ions to the phosphate-sugar backbone. On the other hand, RNA employs distinct metal ions to stabilize higher order structures and to act directly as cofactors in RNA-catalysis [13]. These metal ions, usually Mg2+, bind with an exceptionally high degree of specificity and affinity to their assigned sites.

As indicated above, the research in our laboratory focuses on the characterization of these metal ions and their roles in folding and function in two specific classes of naturally occurring RNAs: riboswitches and group II intron ribozymes. This review summarizes our recent NMR studies with the group II intron ribozyme a5y. We have solved the solution structures of domain 5 (D5) [19] comprising the catalytic center as well as of a shortened construct of the branch-point domain 6 (D6) [17]. The preceding pages provide an overview of the different steps that lead to the structure determination of these two RNA hairpins by NMR.

Subsequent to the structure determination, we are studying the effect of different divalent metal ions on these two important domains and their ability to carry out their function as well as to involve themselves in tertiary interactions with other parts of the intron. To this end, we apply a wide range of techniques, including NMR, X-ray crystallography, fluorescence, UV and CD spectroscopy, biochemical assays and potentiometric pH-titrations to gain insight into the diverse roles that metal ions can play in RNA structure and function. With our work, we hope to shed light on the role of metal ions in RNA and to emphasize the necessity to integrate the fields of chemistry and structural biology for gaining a better understanding of these fascinating molecules.

Abbreviations
Asp, aspartic acid; Cys, cysteine; EBS, exon binding site; FMN, flavin mononucleotide; Glu, glutamic acid; His, histidine; IBS, intron binding site; mRNA, messenger RNA; ORF, open reading frame; r.m.s.d., root mean square deviation; SAM, S-adenosylmethionine; TPP, thiamine pyrophosphate; tRNA, transfer RNA; 5′-UTR, 5′-untranslated region.

Acknowledgements
We thank Dr. Eva Freisinger for careful reading, helpful comments and discussions regarding this manuscript. R.K.O. thanks all the enthusiastic members of his research group: Michele Erat, MayaFurler, SofiaGallo, Dr. Bernd Knobloch, Daniela Kruschel, Miriam Steiner, and Veronica Zelenay for their devoted work and valuable contributions. Financial support by the Swiss National Science Foundation (SNF-Förderungsförpasser PP-002-68733/1) is gratefully acknowledged.

Received: August 23, 2005