Archives of Biochemistry and Biophysics xxx (2017) 1–15

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

# Applications of NMR to structure determination of RNAs large and small

Ravi P. Barnwal [1], Fan Yang, Gabriele Varani*

*Department of Chemistry, University of Washington, Seattle 98195, WA, USA*

ARTICLE INFO

*Article history:*
Received 18 January 2017
Received in revised form
31 May 2017
Accepted 4 June 2017
Available online xxx

*Keywords:*
RNA
Isotope labeling
Structure
NMR spectroscopy

ABSTRACT

Nuclear magnetic resonance (NMR) spectroscopy is a powerful tool to investigate the structure and dynamics of RNA, because many biologically important RNAs have conformationally flexible structures, which makes them difficult to crystallize. Functional, independently folded RNA domains, range in size between simple stem-loops of as few as 10–20 nucleotides, to 50–70 nucleotides, the size of tRNA and many small ribozymes, to a few hundred nucleotides, the size of more complex RNA enzymes and of the functional domains of non-coding transcripts. In this review, we discuss new methods for sample preparation, assignment strategies and structure determination for independently folded RNA domains of up to 100 kDa in molecular weight.



## 1. Introduction

RNA has emerged as an increasingly important player in gene expression and its regulation, but the often elongated and only partially ordered structures responsible for its function make crystallization challenging. NMR of RNA is also challenging because proton density is much smaller than in proteins [1–4], and the chemical diversity of its monomeric units much reduced, making spectral overlap more severe [1,5]. RNA resonances also have shorter transverse relaxation times, since most protons in RNA are $^{13}$C-attached and the shape more anisotropic than globular proteins [3]. These challenges become increasingly severe for larger RNAs (>50–70 nt). As a consequence, the total number of RNA structures (excluding 51 DNA/RNA hybrids) deposited into the PDB accounts for only ~1% of the total PDB depositions, while the human genome codes for many more functional RNAs than proteins. Of all RNA structures in the PDB, 484 have a size of <155 nucleotides (nts) and were determined by NMR, including one from solid-state NMR spectroscopy [6] (Fig. 1). Remarkably, this constitutes about 40% of all RNA structures reported in the PDB, while x-ray crystallography accounts for about 55% of the total. It must be noted, of course, that structures determined by x-ray crystallography are often larger and much more complex and therefore structurally richer.

Understanding the structural basis for the growing diversity of essential biological function of RNA requires an increase in the pace of high-resolution structure determination. In this review, we describe approaches to unravel the structure of folded RNAs that take advantage of recent developments in isotopic labeling schemes, higher magnetic field spectrometers and cryogenically cooled probes. We limit this review to techniques apt to investigate RNAs of up to about 200–300 nucleotides in size; larger RNAs can be studied by NMR, but formal structure determination with an acceptable density of experimental constraints remains an unmet challenge for RNAs greater than 100 nts. Many RNA-binding proteins recognize short single stranded RNAs of <10–15 nucleotides; or stem-loop or internal loop structures of 20–50 nts in size; most RNAs studied by NMR and deposited into the PDB belong to this group. Catalytically active small ribozymes, riboswitches and other important regulatory RNAs are less than 50 nts in size as well. More complex ribozymes and independently folded transcripts, such as tRNA, 5S RNA or 7SK RNA, and independent domain of non-coding transcripts as well, remain within the 200–300 nucleotide cutoff of this review.

## 2. Sample preparation

RNA can be prepared efficiently either by chemical synthesis or various *in vitro* enzymatic transcription protocols and purified by

* Corresponding author.
*E-mail address:* varani@chem.washington.edu (G. Varani).
[1] Present address: Department of Biophysics, Panjab University, Chandigarh 160014 India.

http://dx.doi.org/10.1016/j.abb.2017.06.003
0003-9861/

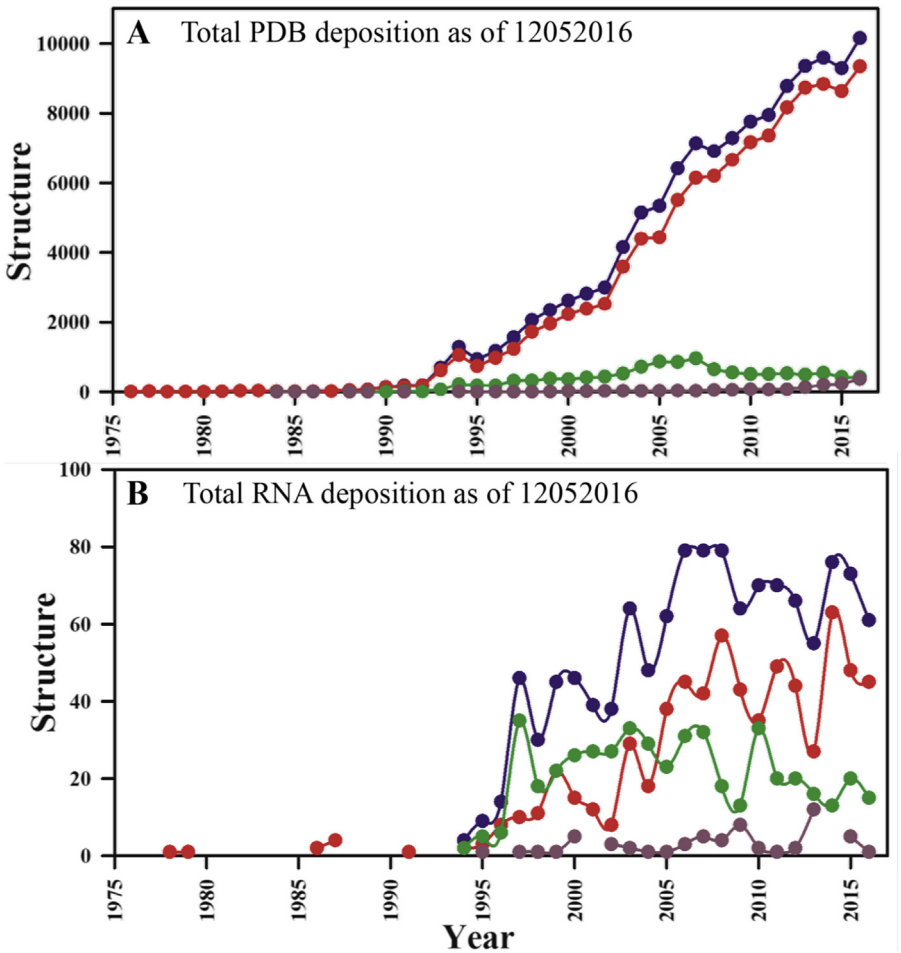


**Fig. 1. PDB statistics of deposited RNA structures as of Dec 2016**: (A) Total number of PDB depositions per year and (B) total RNA structures deposited per year. Here, all reported coordinates are shown in blue, X-ray structures are in red, while NMR structures are in green. Structures determined with other techniques (e.g. electron microscopy; EM) are shown with dark brown filled circles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

either polyacrylamide gel electrophoresis (PAGE) or, preferably, chromatographic methods [7,8]. Denaturing PAGE has been in use for many years [9,10]. It is very robust and provides nucleotide level resolution. The band corresponding to the RNA is excised from the gel and the RNA is extracted by electroelution, ethanol precipitation and extensive dialysis to refold the RNA and remove any leftover acrylamide. More recently, a set of chromatographic approaches has also been reported, which includes reverse-phase high-performance liquid chromatography (HPLC), anion-exchange HPLC, size exclusion chromatography and affinity based chromatography [11–14]. The chromatographic methods are faster, reduce sample losses and retain the RNA in its native conformation throughout, but it is not always possible to obtain sufficient purification of desired products from aborts.

### 2.1. Synthesis of small RNAs

#### 2.1.1. Chemical synthesis

Small RNAs of up to 20 nts can be chemically synthesized and studied using two dimensional NMR. However, isotopic labeling may be required for highly repetitive sequences. Wenter et al. [15] reported a unique method for selective $^{13}C$ labeling of RNA for the structure determination of protein-RNA complexes; however, their phosphoramidites are not commercially available [15]. One of the advantages of this approach for RNA synthesis is position selective labeling which is discussed at the end of the sample preparation section.

#### 2.1.2. In-vitro transcription

RNAs of 20–60 nts can be prepared efficiently by *in vitro enzymatic* transcription [10] with T7, T3 or SP6 RNA polymerases (the first being the most commonly used) [16]. Transcription of RNAs of this size can be performed with chemically synthesized single-stranded or double stranded DNA templates comprising a polymerase promoter region. Two common promoter sequences for T7 are used in this laboratory, called class II and class III [17]; other RNA polymerases (T3 and SP6), require different promoter sequences. *In vitro* transcription based on T7 RNA polymerase often ends with 3′-inhomogeneity, which can be greatly reduced by incorporation of ribozyme sequences in the template in *cis* and *trans* [18–20] or by chemically incorporating 2′-*O*-$CH_3$ nucleotides at the end of the template to reduce jitter of the polymerase [21,22].

#### 2.1.3. Small ribozymes

An alternative approach places hammerhead (HH) ribozymes (or other self-cleaving small ribozymes) at the 5′-end of the DNA template to obtain efficient co-transcriptional cleavage during *in-vitro* transcription (Fig. 2A) [18] to release the desired sequence. Hammerhead ribozymes are a class of small self-cleaving/catalytic RNAs that perform self-cleavage at a well-established site within a

Please cite this article in press as: R.P. Barnwal, et al., Applications of NMR to structure determination of RNAs large and small, Archives of Biochemistry and Biophysics (2017), http://dx.doi.org/10.1016/j.abb.2017.06.003