doi:10.1016/j.jmb.2009.08.001 J. Mol. Biol. (2009) 393, 717–734

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

# A Method for Helical RNA Global Structure Determination in Solution Using Small-Angle X-Ray Scattering and NMR Measurements

**Jinbu Wang[1]†, Xiaobing Zuo[1]†, Ping Yu[1,2], Huan Xu[1], Mary R. Starich[3], David M. Tiede[4], Bruce A. Shapiro[5], Charles D. Schwieters[6] and Yun-Xing Wang[1]***

[1]*Protein–Nucleic Acid Interaction Section, Structural Biophysics Laboratory, National Cancer Institute at Frederick, National Institutes of Health, Frederick, MD 21702, USA*

[2]*SAIC-Frederick, Inc., National Cancer Institute at Frederick, National Institutes of Health, Frederick, MD 21702, USA*

[3]*Office of the Chief, Structural Biophysics Laboratory, National Cancer Institute at Frederick, National Institutes of Health, Frederick, MD 21702, USA*

[4]*The Chemical Sciences and Engineering Division, Argonne National Laboratory, Argonne, IL 60439, USA*

[5]*Center for Cancer Research Nanobiology Program, National Cancer Institute at Frederick, National Institutes of Health, Frederick, MD 21702, USA*

[6]*Division of Computational Bioscience, Center for Information Technology, National Institutes of Health, Bethesda, MD 20892, USA*

We report a "top–down" method that uses mainly duplexes' global orientations and overall molecular dimension and shape restraints, which were extracted from experimental NMR and small-angle X-ray scattering data, respectively, to determine global architectures of RNA molecules consisting of mostly A-form-like duplexes. The method is implemented in the G2G (from global measurement to global structure) toolkit of programs. We demonstrate the efficiency and accuracy of the method by determining the global structure of a 71-nt RNA using experimental data. The backbone root-mean-square deviation of the ensemble of the calculated global structures relative to the X-ray crystal structure is 3.0±0.3 Å using the experimental data and is only 2.5±0.2 Å for the three duplexes that were orientation restrained during the calculation. The global structure simplifies interpretation of multidimensional nuclear Overhauser spectra for high-resolution structure determination. The potential general application of the method for RNA structure determination is discussed.

Published by Elsevier Ltd.

*Corresponding author. E-mail address: wangyu@ncifcrf.gov.

† J.W. and X.Z. contributed equally to this work.

Present address: H. Xu, Gynecology and Obstetric Hospital, Fudan University, 419 Fangxie Road, Shanghai 200011, People's Republic of China.

Abbreviations used: RDC, residual dipolar coupling; DRO, discrete relative orientation; SAXS, small-angle X-ray scattering; SA, simulated annealing; COSY, correlated spectroscopy; NOE, nuclear Overhauser enhancement; NOESY, NOE spectroscopy; PDDF, pair distance distribution function; IPAP, in-phase, antiphase; WAXS, wide-angle X-ray scattering; NSD, normalized spatial discrepancy.

0022-2836/$ - see front matter. Published by Elsevier Ltd.

*Received 11 June 2009;*
*received in revised form*
*4 August 2009;*
*accepted 4 August 2009*
Available online
8 August 2009

***Edited by M. F. Summers*** *Keywords:* RNA structure; SAXS; NMR; duplex orientation; RDC waves

## Introduction

Among the greatest advances in biology today are the discoveries of the various roles played by RNA in biological functions. RNA is an active participant in the regulation of gene expression by interference[1] or by riboswitches,[2] in the processing of RNA introns,[3] in the maintenance of chromosome ends by telomerase,[4] and in protein synthesis by ribosome.[5] RNA function is encoded in its dynamics and structures,[6,7] and determination of RNA structures remains a major goal in contemporary biology. Currently, despite significant advances in X-ray crystallography and solution NMR, structure determination of any given medium- to large-sized RNA molecule with a complex fold remains a daunting task. This is because of the great difficulty in growing crystals and/or obtaining phase information in X-ray crystallography and severe size constraints on structure determination by solution NMR spectroscopy.

The prevailing approach for structure determination of RNA in solution is a "bottom–up" approach, similar to the approach used for determining protein structures,[8] despite vast differences in both structural features and chemical compositions between these two types of biomacromolecules. In RNA, very high structural similarities among its basic building units lead to very similar chemical shift environments; thus, a very narrow chemical shift dispersion and severe NMR signal overlap, making it extremely difficult, if not impossible, to extract sufficient local and global structural information to construct global structures of medium- and large-sized RNAs.[9,10] Consequently, the current bottom–up approach runs into a size barrier.

A survey of RNA X-ray crystal structures with resolution better than 3.0 Å in the Protein Data Bank reveals that A-form-like duplexes are the most predominant building blocks in RNA structures and that the A-form conformation in terms of the sugar-phosphate backbones is very conserved (Table 1 and Table S1). The other A-form structural parameters also vary in narrow ranges, except for the base-pair tilt, which heavily depend on the base-pair types (Table 1). Thus, RNA duplexes (stems) depicted in a secondary structure can be treated as A-form-like and can be generated from RNA databases as approximate initial structures with acceptable accuracy. Therefore, in determining a global architecture of an RNA molecule that mostly consists of duplexes, there are essentially two problems: (1) orientations and phases (the rotation of duplex around the helical axis) of duplexes and (2) the relative positions of the duplexes in a global sense. Once the global relative orientations, phases, and relative positions of these duplexes are determined, so is the approximate global structure of the RNA. We used residual dipolar coupling (RDC)–structural periodicity correlation to derive the discrete relative orientations (DROs) of duplexes in terms of the polar angles of a duplex axis, $\Theta$ and $\Phi$, and phase angle $\rho_0$[11] (Fig. 1a–c) and utilized the small-angle X-ray scattering (SAXS)-derived molecular envelope to identify the correct combination of duplex orientations and approximate relative positions of the duplexes. The programs ORIENT, BLOCK, and PACK in the G2G (from global measurement to global structure) toolkit calculates DROs, generates duplex coordinates from an RNA structure database library, and packs duplexes together, respectively. In the next step, the starting structures that are generated by G2G are subjected to rigid-body simulated annealing (SA) refinement, where the orientations and phases of duplexes and the overall shape of the RNA are restrained. We demonstrated the accuracy and efficiency of the method by determining the global structure of the adenine riboswitch (riboA) (71 nt) using experimental RDC and SAXS data. The structure of riboA with a different sequence (see Materials and Methods), previously determined using X-ray crystallography,[12] was used as a benchmark for the G2G method. It is noteworthy that combined use of RDCs and SAXS data to refine a known structure of tRNA has been reported in the literature.[13]

## Results

### The G2G method

The four DROs of a chiral molecule (a domain or a whole protein or RNA) are usually derived using the singular value decomposition method if the coordinates are known[14] or using periodicity–RDC correlation[11] if the structure is that of periodic structural elements, as described briefly in the following. Duplexes belong to a unique class of chiral molecules that consist of repetitive structural units that are arranged in certain periodic patterns. This structural periodicity is reflected in and correlated with geometrical measurements, such as dipolar coupling or RDC of particular types of spin pairs.[11,15] DROs are derived using an efficient nonlinear least-squares optimization routine to fit the RDC–structural periodicity correlation to experimental imino RDCs of duplexes.[11] The general RDC equation in terms of