Methods xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Methods

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

Applications of PLOR in labeling large RNAs at specific sites

Yu Liu [a,*], Ping Yu [b], Marzena Dyba [b], Rui Sousa [c], Jason R. Stagno [a], Yun-Xing Wang [a]

[a] Protein-Nucleic Acid Interaction Section, Structural Biophysics Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD 21702, USA
[b] Structural Biophysics Laboratory, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, MD 21702, USA
[c] Department of Biochemistry, University of Texas Health Science Center, San Antonio, TX 78229, USA

ARTICLE INFO

Article history:
Received 2 February 2016
Received in revised form 15 March 2016
Accepted 22 March 2016
Available online xxxx

Keywords:
PLOR
RNA
Specific labeling

ABSTRACT

Incorporation of modified or labeled nucleotides at specific sites in RNAs is critical for gaining insights into the structure and function of RNAs. Preparation of site-specifically labeled large RNAs in amounts suitable for structural or functional studies is extremely difficult using current methodologies. The position-selective labeling of RNA, PLOR, is a recently developed method that makes such syntheses possible. PLOR allows incorporation of various probes, including $^{2}D/^{13}C/^{15}N$-isotopic labels, Cy3/Cy5/Alexa488/Alexa555 fluorescent dyes, biotin and other chemical groups, into specific positions in long RNAs. Here, we describe in detail the use of PLOR to label RNAs at specific segment(s) or discrete sites.

Contents

1. Introduction

Knowledge about RNA structures and dynamics is fundamentally important to understand RNA's functions. First of all, high-resolution structures of macromolecules are invaluable for better understanding their functions in biological systems. NMR and X-ray crystallography are the two main methods used to solve atomic-resolution structures of proteins and RNAs, and account for >97% of the RNA structures in the Protein Data Bank. Specific labeling of RNAs offers several advantages for both of these methods. Selective isotopic labeling of RNAs, especially those larger than 50 nt, can greatly alleviate NMR peak overlap, thereby simplifying spectra interpretation to allow extraction of structures and dynamics information that would otherwise be difficult to obtain [1–8]. With the availability of high-field NMR spectrometers, the use of selectively isotopically labeled RNAs is now one of the primary strategies for RNA structure determination. In X-ray crystallography, the calculation of electron density maps requires determination of the phases of the diffraction data, which is usually done either by molecular replacement (MR) with a known, homologous structure, or by anomalous phasing by incorporating heavy atoms into the crystal [9–16]. Since known, homologous structures are rarely available for RNAs, *de novo* phasing of RNA crystal data can be achieved through site-specific incorporation of heavy-atom-labeled nucleotides. Secondly, smFRET is a very powerful

Abbreviations: NMR, nuclear magnetic resonance; PLOR, position-selective labeling of RNA; T7 RNAP, T7 RNA polymerase; nt, nucleotides; smFRET, single molecule Förster resonance-energy transfer; SPE, solid-phase extraction; PCR, polymerase chain reaction; NOESY, nuclear Overhauser spectroscopy; TROSY, transverse relaxation optimized spectroscopy.

* Corresponding author.
E-mail address: liuy13@mail.nih.gov (Y. Liu).

http://dx.doi.org/10.1016/j.ymeth.2016.03.014
1046-2023/

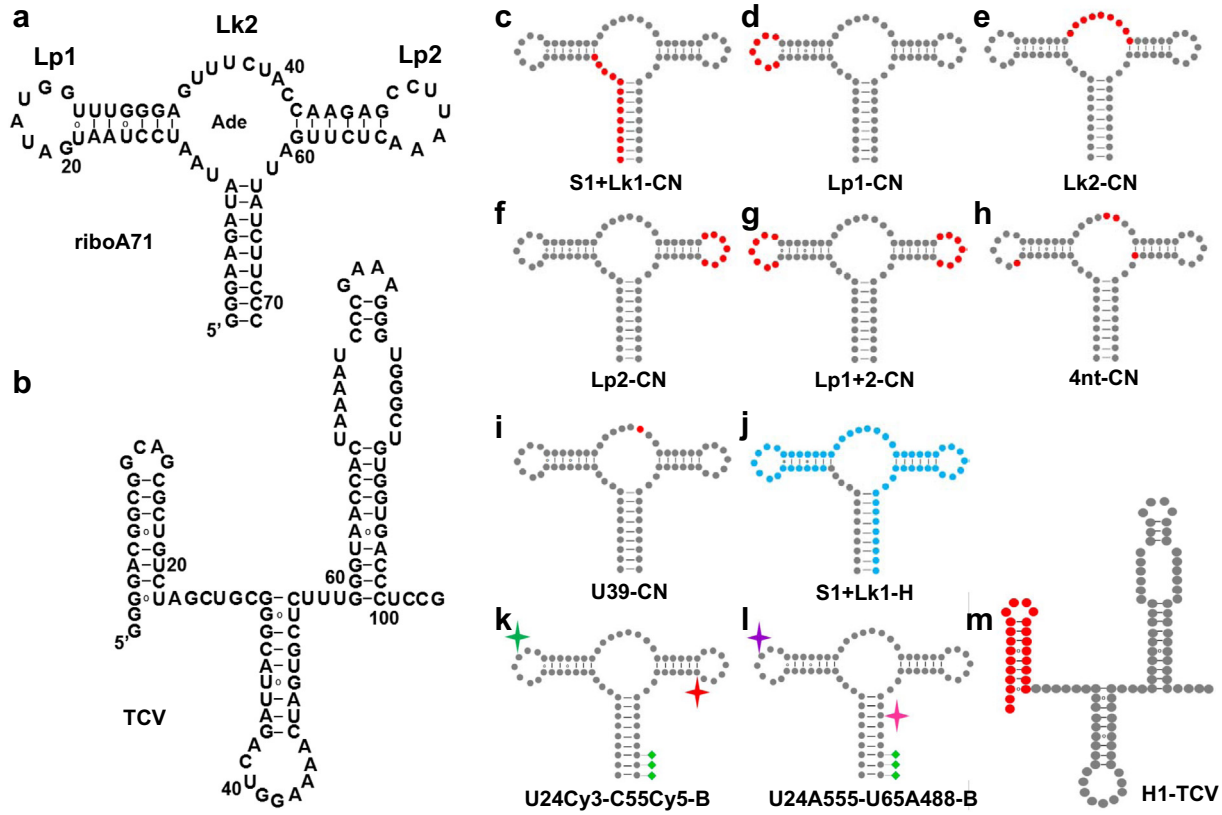

Fig. 1. Site-specific labeling of RNA by PLOR. a–b, Secondary structure of a, riboA71,and b, TCV. c–m, riboA71 and TCV with specific labels. Non-labeled residues are shown as grey dots. $^{13}C^{15}N$- and ^{2}D-isotopic residues are represented by red and blue dots, respectively, in c (S1 + Lk1-CN-riboA71), d (Lp1-CN-riboA71), e (Lk2-CN-riboA71), f (Lp2-CN-riboA71), g (Lp1 + 2-CN-riboA71), h (4nt-CN-riboA71), i (U39-CN-riboA71), j (S1 + Lk1-H-riboA71), and m (H1-TCV). The fluorescent dyes, Cy3 and Cy5, in k (U24Cy3-C55Cy5-B-riboA71) are represented by green and red stars, respectively. The fluorescent dyes, Alexa 488 and Alexa 555, in l (U24A555-U65A488-B-riboA71) are represented by pink and purple stars, respectively. The green diamonds in k and l represent biotin groups.

tool to study conformation space and dynamics. RNA labeled with fluorophores offers the capability to characterize dynamics of individual molecules [17–19] Lastly, RNA molecules are used as biosensors and labeled RNAs with various probes have broad applications in disease diagnosis, substance detection and molecular imaging [20–22].

The most widely used methods to introduce labels into RNAs during synthesis are *in vitro* transcription and solid-phase chemical synthesis. *In vitro* transcription has been routinely used to prepare RNAs in large quantity, in which nucleotide types are homogeneously labeled, but is not readily adapted to allow labeling of nucleotides by position [7,23,24]. Chemical synthesis provides greater flexibility in labeling RNA at specific positions. In theory it is possible to use the chemical synthesis method to generate small RNAs selectively labeled at any position, provided that all labeling chemical reagents are commercially available. However, its primary limitations are size and cost [23–26]. Neither of these two methods is practical for preparing multi-milligram amounts of large RNAs with position-specific isotope labels for NMR.

Recently, we reported a method, PLOR, to synthesize selectively-labeled RNA on a milligram scale [27]. We demonstrated PLOR by synthesizing RNAs with various labeling schemes and labeling reagents. Those RNAs include a specifically labeled 104-nt TCV RNA (Fig. 1) from the 3′ untranslated region of turnip crinkle virus [28–30], and a series of 71-nt riboA71 RNAs (Fig. 1) that form the aptamer domain of an adenine riboswitch [29,31–33]. Isotope-labeled (^{13}C, ^{15}N, and ^{2}D), fluorescent (Cy3, Cy5, Alexa 488, and Alexa 555), or biotinylated derivatives of riboA71 were also generated by PLOR, in which the labeled nucleotides were incorporated into a single helix (Fig. 1c and j), a single hairpin (Fig. 1d and f), an internal loop (Fig. 1e), multiple hairpins (Fig. 1g), multiple discrete positions (Fig. 1h, k, and l), or, a single position (Fig. 1i), and subsequently used in NMR and smFRET studies.

2. Theory

In vitro transcription catalyzed by T7 RNA polymerase has a slow, unstable initiation phase, followed by fast chain and highly processive elongation at about 200 nucleotides per second [34,35]. It has been shown that the elongation complex composed of DNA:RNA:T7 RNAP is extremely stable once it is formed, and that transcription by this ternary complex can be paused and resumed [36]. PLOR exploits this ability to pause and resume transcription by the elongation complex by omitting and adding back the NTPs required to proceed pause positions. Polymerase 'walking' is not unique to T7 RNA polymerase, such 'walking' has been observed for other RNA polymerases to study RNAP mechanism [37]. PLOR combines aspects of both liquid-phase transcription and solid-phase chemical synthesis. The DNA template in PLOR is coupled to a solid support, which allows stepwise buffer and NTP changes. The PLOR reaction is divided into three stages: initiation, elongation, and termination (Fig. 2) [27]. The transcription reaction is initiated by gentle mixing of T7 RNAP, NTPs and bead attached templates in a reaction vessel. One or more types of NTPs are excluded in the reaction mixtures, which causes the RNA polymerase to pause at the positions where it requires the missing NTP(s). After removing the residual NTPs by SPE and rinsing the beads thoroughly, a new NTP mix is added to allow the transcription