Secondary Structure for RNA Aptamers Binding to Guanine-Rich Sequence in the 5'-UTR RNA of N-Ras Oncogene

Bongrae Cho
Department of Applied Chemistry, Cheongju University, Cheongju 28503, Korea.
E-mail: brcho@cju.ac.kr
(Received December 16, 2020; Accepted January 8, 2021)

ABSTRACT. RNA molecules which bind to the G-rich sequence in the 5'-UTR RNA which plays an important role in expression of N-ras, were selected. The secondary structures of five selected RNA aptamers including primer sequence were found by the CLC RNA workbench ver. 4.2 program (www.clcbio.com) and investigated with RNA structural probes such as RNase T1 which has specificity for a G in single-stranded region, RNase V1 specific for double strand and nuclease S1 specific for single strand. The generalized secondary structure model was proposed and characterized. It was composed of a central long double strand region flanked by single strand region at both end sides. The double strand region had an internal single-strand region and bulges. The single strand loop in the right side was composed of four or five nucleotides.

Key words: Secondary structure, RNA aptamer, guanine-rich sequence RNA, RNA structural probe

INTRODUCTION

Guanine-rich sequences observed in the terminal segments of eukaryotic genomes form the four-stranded topology called G-quadruplexes. Guanine-rich sequences in the 5'-UTR of human genome, were identified. Sequence of 5'-GGGAGGGGCGGGUCUGGG-3' containing four guanine-tracts in the 5'-UTR of oncogenic N-ras is thought to be one of G-quadruplex-forming elements in the 5'-UTRs of human genome.

To design the therapeutic RNAs inhibiting translation level of N-ras, we selected RNA aptamers interacting with the G-rich sequence in the 5'-UTR RNA of N-ras from an RNA pool containing 30 randomized nucleotides. Optimally predicted secondary structures of five selected RNA aptamers - 11-21-5, 11-21-30, 11-30-32, 11-30-36 and 11-48-2 - including primer sequence were found by the CLC RNA workbench ver. 4.2 program (www.clcbio.com). Their secondary structures were investigated with RNA structural probes such as RNase T1 which has specificity for a G in single-stranded region, RNase V1 specific for double strand and nuclease S1 specific for single strand, which have been widely used for probing RNA structure in solution. In this study, the determined secondary structures of five RNA aptamers were analyzed and compared. The new secondary structure model which can be commonly applied to them was proposed and characterized.

EXPERIMENTAL

Preparation of RNA

RNA aptamers were synthesized by in vitro transcription with T7 RNA polymerase and eluted from the gel with the crush and soak method. The 5' terminal phosphate of resulting RNAs were removed with CIP (calf intestinal alkaline phosphatase) and then labeled at the 5' end with T4 polynucleotide kinase. 5' End-labeled RNAs were purified with PAGE (polyacrylamide gel electrophoresis).

Enzymatic Cleavage Reaction

5'-terminal radiolabeled RNA aptamers were heated in binding buffer (30 mM Tris-acetate, pH 7.5, 60 mM Mg(OAc)₂, 120 mM K(OAc), and 120 mM NH₄(OAc)) at 37°C for 20 min and cooled to room temperature. Then, 0.1-1 unit of nuclease S1, 0.001-0.01 unit of RNase V1, or 0.1-1 unit of RNase T1 was added to the above mixture, which was incubated for 25 min at RT. An additional 1 mM ZnCl₂ for nuclease S1 cleavage was added to the reaction mixture. The cleavage products were separated on denaturing 15% polyacrylamide gel. Cleavage patterns were visualized with autoradiography.
RESULTS AND DISCUSSION

RNA molecules binding to the G-rich sequence in the 5'-UTR RNA which plays an important role in expression of N-ras, were selected from a pool (Fig. 1). After 11th round of selection, the selected RNA aptamers were reverse-transcribed into cDNA, amplified by PCR and cloned into pGEM-T easy vector for sequance analysis. The sequences of the selected RNAs were shown in Fig. 2. The consensus sequences of the selected RNA aptamers were searched with CLUSTAL W (1.83) multiple sequence alignment. The consensus sequence GGGAUCCGCAUGCAAGCUUA were found between the selected RNAs.

Secondary structures of selected RNA aptamers were optimally predicted with theoretical method like CLC RNA workbench ver. 4.2 program and supplemented with RNA structural probes. Phosphodiester bonds cleaved by RNA structural probes for analysis of secondary structures were shown in Table 1 and Fig. 3. Base G1 was conserved in the secondary structures of five RNA aptamers - 11-21-5, 11-21-30, 11-30-32, 11-48-2 and 11-30-36. Phosphodiester bonds of sequence G2-G-A-C6 in an RNA aptamer 11-30-36, C6-C7 in an RNA aptamer 11-30-32 and C7-G8 in two RNA aptamers - 11-30-32 and 11-48-2 were cleaved by RNase V1. Phosphodiester bonds of G33-G34 in three RNA aptamers - 11-30-32, 11-30-36 and 11-48-2, G34-A35 in an RNA aptamer 11-30-32, C37-C38 in two RNA aptamers - 11-30-32 and 11-48-2, and C38-G39 in an RNA aptamer 11-30-36 were also cleaved by RNase V1. These cleavage results by double-strand specific RNase V1 suggested that two sequences G2GAUCC7 and G33GAUCC38 formed

Table 1. The cleavage of phosphodiester bonds by single-strand specific and double-strand specific nuclease

<table>
<thead>
<tr>
<th>Cleavage</th>
<th>Enzymes</th>
<th>Aptamers</th>
<th>Structural Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>C7-G8</td>
<td>V1</td>
<td>11-30-32, 11-48-2</td>
<td>ds</td>
</tr>
<tr>
<td>G8-C9</td>
<td>V1</td>
<td>11-30-32, 11-48-2</td>
<td>ds</td>
</tr>
<tr>
<td>G12-C13</td>
<td>T1, V1, S1</td>
<td>11-30-36, 11-30-32, 11-48-2</td>
<td>ds or ss</td>
</tr>
<tr>
<td>A15-G16</td>
<td>V1</td>
<td>11-30-36, 11-30-32, 11-48-2</td>
<td>ds</td>
</tr>
<tr>
<td>G16-C17</td>
<td>T1</td>
<td>11-30-36, 11-30-32, 11-48-2</td>
<td>unstable ds</td>
</tr>
<tr>
<td>U18-U19</td>
<td>V1, S1</td>
<td>11-30-36, 11-30-32, 11-48-2</td>
<td>stacked ss</td>
</tr>
<tr>
<td>U19-A20</td>
<td>V1, S1</td>
<td>11-30-36, 11-30-32, 11-48-2</td>
<td>stacked ss</td>
</tr>
<tr>
<td>A20-C21</td>
<td>V1, S1</td>
<td>11-30-36, 11-30-32, 11-48-2</td>
<td>stacked ss</td>
</tr>
<tr>
<td>G23-C24</td>
<td>T1</td>
<td>11-30-36, 11-48-2</td>
<td>unstable ds</td>
</tr>
<tr>
<td>U32-G33</td>
<td>V1, S1</td>
<td>11-30-32, 11-48-2</td>
<td>unstable ds</td>
</tr>
<tr>
<td>G33-G34</td>
<td>V1</td>
<td>11-30-36, 11-30-32, 11-48-2</td>
<td>ds</td>
</tr>
<tr>
<td>C37-C38</td>
<td>V1</td>
<td>11-30-32, 11-48-2</td>
<td>ds</td>
</tr>
<tr>
<td>G39-C40</td>
<td>T1</td>
<td>11-30-36, 11-30-32, 11-48-2</td>
<td>ss</td>
</tr>
<tr>
<td>C40-A41</td>
<td>V1, S1</td>
<td>11-30-36, 11-30-32, 11-48-2</td>
<td>stacked ss</td>
</tr>
<tr>
<td>A41-U42</td>
<td>V1, S1</td>
<td>11-30-36, 11-30-32, 11-48-2</td>
<td>stacked ss</td>
</tr>
<tr>
<td>U42-G43</td>
<td>V1, S1</td>
<td>11-30-36, 11-30-32, 11-48-2</td>
<td>stacked ss</td>
</tr>
<tr>
<td>G43-C44</td>
<td>T1, S1</td>
<td>11-30-36, 11-30-32, 11-48-2</td>
<td>ss</td>
</tr>
</tbody>
</table>

Only phosphodiester bonds cleaved in more than two RNA aptamers are shown. Sss indicate single stranded and dss indicate double stranded.

Figure 1. (a) Guanine-rich sequence used for in vitro selection. (b) Scheme of experimental strategy for SELEX. RNA aptamers interacting with a G-rich sequence RNA were selected with G-rich sequence RNA-attached affinity column.

Figure 2. Sequences of the randomized region in selected RNAs. The consensus sequences are underlined in bold letters.
stable base pairs. Bases in this double strand region were well conserved in five RNA aptamers.

Base G8 existed as a bulge in three RNA aptamers - 11-21-5, 11-21-30 and 11-30-36 because U32 was deleted. But this base was thought to pair with U32 to form G8:U32 base pair in two RNA aptamers - 11-30-32 and 11-48-2.

Phosphodiester bond of C9-A10 in an RNA aptamer 11-30-32 and two bonds of sequence A10-U11 in an RNA aptamer 11-48-2 were cleaved by RNase V1. And phosphodiester bonds of U30-G-U32 in an RNA aptamer 11-30-36 were also digested by RNase V1. These cleavage results by RNase V1 meant that two sequences C9AU11 and G29UG31 formed base pairs in four RNA aptamers - 11-21-5, 11-21-30, 11-30-36 and 11-48-2. But A10U11 was thought to exist as a bulge in an RNA aptamer 11-30-32 because the counterpart sequence G29U30 was deleted.

Phosphodiester bond of G12-C13 in an RNA aptamer 11-30-36 was cleaved by RNase T1, the same bond in an RNA aptamer 11-30-32 was cleaved by both RNase T1 and RNase V1, and the same bond in an RNA aptamer 11-48-2 was cleaved by both RNase T1 and nuclease S1. Bond of A27-C28 in an RNA aptamer 11-30-36 was cleaved by RNase V1. These results suggested that base G12 paired with U28 to form G12:U28 base pair in two RNA aptamers - 11-21-5 and 11-21-30, and with C28 to form G12:C28 base pair in an RNA aptamer 11-30-32. And base C13 paired with G27 to form C13:G27 base pair in three RNA aptamers - 11-21-5, 11-21-30 and 11-30-32. Bases A14 and C26 existed as bulges in these three RNA aptamers. But two sequences G12CA14 and C26GU28 were thought to exist as single strand regions in two RNA aptamers - 11-30-36 and 11-48-2.

Phosphodiester bond of A15-G16 in three RNA aptamers - 11-30-32, 11-30-36 and 11-48-2 was cleaved by RNase V1 and bond of G16-C17 in three RNA aptamers - 11-30-32, 11-30-36 and 11-48-2 was cleaved by RNase T1. Bond of G23-C24 in two RNA aptamers - 11-30-36 and 11-48-2 was weakly digested by RNase T1. So sequence A15GC17 was thought to pair with G23CC25 in two RNA aptamers - 11-30-36 and 11-48-2, with G23CU25 in an RNA aptamer 11-21-5, with G23UC25 in an RNA aptamer 11-21-30, and with C23AU25 in an RNA aptamer 11-30-32 to form the double strand region. But base pair C17:G(A)23 was thought to be unstable according to the result of accessibility of RNase T1 to this base pair.

Phosphodiester bond of U18-U19 in two RNA aptamers - 11-30-32 and 11-30-36 was cleaved by both RNase V1 and nuclease S1, bond of U19-A20 in three RNA aptamers - 11-30-32, 11-30-36 and 11-48-2 was cleaved by both RNase V1 and nuclease S1, bond of A20-C21 in two RNA aptamers - 11-30-32 and 11-30-36 was cleaved by both RNase V1 and nuclease S1, and bond of C21-A22 in an RNA aptamer 11-48-2 was cleaved by nuclease S1. Seeing that phosphodiester bonds in U18UANN22 were cleaved by single-strand specific nuclease S1, this sequence U18UANN22 was thought to be located in single strand loop. And it was thought that bases in this loop must be stacked from intramolecular interaction because they were also susceptible to double-strand specific RNase V1. The size of the single strand loop was composed of four or five nucleotides. The single strand loop of an RNA aptamer 11-30-32 had sequence of UUAC, the loop of an RNA aptamer 11-21-30 had sequence of UUAG, the loop of two RNA aptamers - 11-21-5 and 11-30-36 had sequence of UUACU, and the loop

---

**Figure 3.** Secondary structure for RNA aptamers which bind to the G-rich sequence in the 5'-UTR RNA of N-ras. Squares indicate nucleotides on double helical regions that form base-pairs and circles indicate nucleotides on single-strand regions. Two nucleotides in squares show that one of them can occur on this position. Arrows indicate the sites cleaved by each enzyme.
of an RNA aptamer 11-48-2 had sequence of UUACA.

Phosphodiester bond of G39-C40 in three RNA aptamers - 11-30-32, 11-30-36 and 11-48-2 was cleaved by RNase T1, bond of C40-A41 in two RNA aptamers - 11-30-32 and 11-30-36 was cleaved by both RNase V1 and nuclease S1, bond of A41-U42 in two RNA aptamers - 11-30-32 and 11-30-36 was cleaved by both RNase V1 and nuclease S1, bond of U42-G43 in three RNA aptamers - 11-30-32, 11-30-36 and 11-48-2 was cleaved by both RNase V1 and nuclease S1, bond of G43-C44 in three RNA aptamers - 11-30-32, 11-30-36 and 11-48-2 was cleaved by both RNase T1 and nuclease S1, and bond of G47-C48 in an RNA aptamer 11-48-2 was cleaved by RNase T1. So sequence G39CAUGCAAGCUU50 was thought to be a single strand region. And the bases in this single strand region must be stacked from intramolecular interaction because they were also susceptible to RNase V1.

From above cleavage results with theoretically predicted model, the generalized secondary structure for RNA aptamers which bind to the G-rich sequence in the 5'-UTR RNA of N-ras was proposed in Fig. 3. It had a central long double strand region flanked by single strand region at both end sides. We anticipate that this model can be employed to design the candidates of therapeutic agents inhibiting translation level of N-ras oncogene. Footprinting assay will be applied to get the information for contact points between G-rich RNA ligand and aptamer RNA. And binding assay will also be performed to know whether aptamer RNA does not bind to UTR of other genes but UTR of N-ras.

CONCLUSION

In summary, we proposed the secondary structure model for RNA aptamers which bind to the G-rich sequence in the 5'-UTR RNA of N-ras. The model was composed of a central long double strand region flanked by single strand region at both end sides. The double strand region had an internal single-strand region and bulges. The single strand loop in the right side was composed of four or five nucleotides. The bases in this model were considerably conserved. This secondary structure model is expected to be used to give the information necessary to get higher order structure with X-ray crystallography or NMR (nuclear magnetic resonance) spectroscopy.

Acknowledgments. This work was supported by a 2020 research grant of Cheongju University.

REFERENCES