Probing RNA Structure and Function by Nucleotide Analog Interference Mapping

The chemistry of phosphorothioate interference has been expanded into a method to identify the specific chemical groups that are important for RNA activity (Eckstein, 1985; Gaur and Krupp, 1993; Strobel and Shetty, 1997). This approach, termed nucleotide analog interference mapping (NAIM), makes it possible to simultaneously, yet individually, test the contribution of a particular functional group at almost every position within the RNA molecule in an assay that is as simple as RNA sequencing. In a NAIM experiment the smallest mutable unit is not the base pair, but rather the individual functional groups that comprise the nucleotide. Because the deletion or modification of particular functional groups within an RNA can severely affect its activity, this approach makes it possible to determine the chemical basis of RNA structure and function (Ryder et al., 2000).

Using phosphorothioate-tagged nucleotide analogs, all of the positions on the RNA can be simultaneously characterized. NAIM involves a four-step process: (1) the phosphorothioate-tagged nucleotide analog is randomly incorporated into the RNA by in vitro transcription; (2) the active members of the RNA population are selected from inactive transcripts; (3) the phosphorothioate linkage is cleaved by iodine treatment; and (4) the sites of analog substitution detrimental to activity are identified by gel electrophoresis and autoradiography. The approach is applicable to any RNA that can be transcribed in vitro and where active variants can be selected in some fashion. RNA functions include catalysis, protein or ligand binding, folding, and others.

STRATEGIC PLANNING

Phosphorothioate-Tagged Nucleotide Analog Triphosphates (NTPαS)

NAIM relies on the use of nucleotides in which one of the non-bridging α-phosphate oxygen atoms has been replaced with a sulfur (Eckstein, 1985). This modification does not drastically alter the RNA structure in the vast majority of cases, and any effects resulting from the sulfur are readily controlled. The NTPαS is randomly incorporated into the RNA molecule by in vitro transcription. The sulfur facilitates the direct detection of sites sensitive to analog incorporation because the phosphorothioate linkage can be selectively cleaved with iodine without affecting any of the other bonds within the RNA (Gish and Eckstein, 1988). The RNA fragments are separated by denaturing gel electrophoresis to identify the extent of analog incorporation and magnitude of interference at each position within the transcript.

NAIM uses both parent and analog phosphorothioate-tagged nucleotides (Ryder et al., 2000). Any effects that arise from the phosphorothioate tag are controlled by comparison to the band intensity from the four parental 5′-O-(1-thio)nucleotide triphosphates (AαS, CαS, GαS, and UαS). The phosphorothioate substitution introduces an additional chiral center into the nucleotide. Synthesis of the phosphorothioate-tagged nucleotides results in a diastereomeric mixture of R_P and S_P isomers, of which only the S_P diastereomer is recognized by RNA polymerase (Chamberlin et al., 1983; Griffiths et al., 1987). The transcription reaction proceeds with inversion of configuration at the α-phosphate, resulting in RNAs with R_P phosphorothioate substitutions. It is possible to purify the S_P diastereomer of the triphosphate by HPLC (Fischer et al., 1999), but the R_P diastereomer is not recognized by the polymerase, nor does it inhibit the transcription reaction; therefore, its inclusion in the in vitro transcription reaction has no significant consequence.

Contributed by Jesse C. Cochrane and Scott A. Strobel

Analog nucleotides contain modifications to the sugar or the base in addition to the phosphorothioate tag (Ryder et al., 2000). The range and variety of modifications are summarized in Table 6.9.1. Any analog that is amenable to transcriptional incorporation by RNA polymerase, even at low efficiency, can be used in this approach.

To date, twelve adenosine analogs have been utilized in NAIM (Conrad et al., 1995; Hardt et al., 1996; Ortoleva-Donnelly et al., 1998; Soukup et al., 2002; Jones and Strobel, 2003; Schwans et al., 2003). Eight analogs modify the nucleotide base and four modify the ribose sugar. The base analogs include purine riboside (PurαS), N-methyladenosine (m^{6}AαS), 7-deazaadenosine (7dAαS), diaminopurine riboside (DAPαS), 2-aminopurine riboside (2APαS), formicin (FormαS), 3-deazaadenosine (c^{3}AαS), and 8-azaadenosine (n^{8}AαS). The ribose sugar analogs all modify the 2'-OH group and include 2'-deoxyadenosine (dAαS), 2'-deoxy-2'-fluoroadenosine (^{F}AαS), 2'-O-methyladenosine (^{OMe}AαS), and 2'-deoxy-2'-thioadenosine (^{SH}AαS). Each of these analogs provides specific information about the chemical basis of RNA activity at almost every incorporated position in the transcript. PurαS, 2APαS, and m^{6}AαS measure the effect of modifications to the N6 exocyclic amine of adenosine. The base analog 7dAαS replaces the N7 nitrogen with a C-H group. Interference with this analog indicates an important major groove contact to the ring nitrogen. Additionally, 7dAαS, FormαS, and n^{8}AαS can be used to probe important protonation events. DAPαS and 2APαS both add an additional amine to the C2 position of adenosine. dAαS interference identifies the 2'-OH groups important for RNA function, while ^{F}AαS delineates the role these 2'-OH groups play as either hydrogen bond donors or hydrogen bond acceptors.

Seven guanosine analogs have been utilized in NAIM (Hardt et al., 1996; Strobel and Shetty, 1997; Basu et al., 1998; Kazantsev and Pace, 1998; Schwans et al., 2003). Four analogs modify the base functional groups of G, including inosine (IαS), N^2-methylguanosine (m^{2}GαS), 7-deazaguanosine (7dGαS), and 6-thioguanosine (s^{6}GαS). Three analogs, 2'-deoxyguanosine (dGαS), 2'-O-methylguanosine (^{OMe}GαS), and 2'-deoxy-2'-thioguanosine (^{SH}GαS), modify the 2'-OH of the ribose ring. IαS and m^{2}GαS both modify the N2 exocyclic amine of G. m^{2}GαS substitution is isoenergetic with G in the context of G·C, G·U, and G·A base pairs, and m^{2}G interference has only been observed at sites of tertiary hydrogen bonding (Rife et al., 1998). The base analog 7dGαS, like 7dAαS, replaces the N7 nitrogen with a C-H group. s^{6}GαS also modifies a major groove functional group. It introduces an oxygen-to-sulfur substitution at the O6 keto group of G. dGαS, like dAαS, replaces the 2'-OH with a proton.

Thirteen pyrimidine analogs have been used, including seven cytidine and six uridine analogs (Ryder and Strobel, 1999b; Szewczak et al., 1999; Oyelere et al., 2002; Schwans et al., 2003). These include seven analogs that modify the 2'-OH group of the ribose: 2'-deoxyuridine (dUαS), 2'-deoxy-2'-fluorouridine (^{F}UαS), 2'-O-methyluridine (^{OMe}UαS), 2'-deoxy-2'-thiouridine (^{SH}UαS), 2'-deoxycytidine (dCαS), 2'-O-methylcytidine (^{OMe}CαS), and 2'-deoxy-2'-thiocytidine (^{SH}CαS). This series of analogs can be used to address the same types of questions as outlined for the A analogs. Two additional uridine derivatives have been used in NAIM, 5-methyluridine (m^{5}UαS) and pseudouridine ($\Psi\alpha$S). m^{5}UαS introduces steric bulk in the major groove and probes for possible hydrophobic contacts to the base. $\Psi\alpha$S changes the 5 position of U to a nitrogen and is a C-linked nucleoside. Four cytidine analogs have been synthesized for use in NAIM experiments, 5-fluorocytidine (f^{5}CαS), 6-azacytidine (n^{6}CαS), pseudoisocytidine (ΨiCαS), and zebularine (ZαS). All four of these analogs have altered pK$_a$ values.

Some of the phosphorothioate-tagged nucleotide analog triphosphates are available from a variety of commercial sources. TriLink Biotechnologies sells the parental and 2'-deoxy-phosphorothioates, Amersham-Pharmacia sells two 2'-deoxy-phosphorothioates, and

Table 6.9.1 Phosphorothioate-Tagged Nucleotides for NAIM

Analog δTPαS	[δTPαS] (mM)[a]	[NTP] (mM)[b]	Polymerase (WT or Y639F)	Description	Reference
AαS (S$_p$ isomer only)[c]	0.05	1.0	WT		Christian and Yarus (1992)
7dAαS (S$_p$)	0.05	1.0	WT	Replaces N7 with a carbon	Ortoleva-Donnelly et al. (1998)
m^{6}AαS (S$_p$)	0.1	1.0	WT	Replaces a hydrogen on exocyclic amine with methyl	Ortoleva-Donnelly et al. (1998)
DAPαS (S$_p$)	0.025	1.0	WT	Replaces H2 with exocyclic amine	Strobel and Shetty (1997)
PurαS (S$_p$)	2.0	1.0	WT	Replaces exocyclic amine with a hydrogen	Ortoleva-Donnelly et al. (1998)
2APαS (S$_p$)	0.5	1.0	WT	Replaces exocyclic amine with hydrogen and H2 with exocyclic amine	Ortoleva-Donnelly et al. (1998)
FormAαS	0.1	1.0	WT	Replaces N9 with carbon and C8 with nitrogen	Ryder et al. (2001)
c^{3}AαS	2.0	0.5	Y639F	Replaces N3 with carbon	Ryder et al. (2001)
n^{8}AαS	0.5	1.0	WT	Replaces C8 with nitrogen	Ryder et al. (2001)
dAαS (S$_p$)	0.8	1.0	Y639F	Replaces 2′-OH with hydrogen	Ortoleva-Donnelly et al. (1998)
^{OMe}AαS (S$_p$)	2.0	0.2	Y639F	Replaces hydrogen on 2′-OH with methyl	Ortoleva-Donnelly et al. (1998)
^{F}AαS (S$_p$)	0.25	1.0	Y639F	Replaces 2′-OH with fluorine	Ortoleva-Donnelly et al. (1998)
^{SH}AαS	0.4	1.0	Y639F	Replaces 2′-OH with 2′-SH	Schwans et al. (2003)
GαS (S$_p$ isomer only)[c]	0.05	1.0	WT		Christian and Yarus (1992)
7dGαS (S$_p$)	0.05	1.0	WT	Replaces N7 with a carbon	Kazantsev and Pace (1998)
m^{2}GαS (S$_p$)	0.1	1.0	Y639F	Replaces a hydrogen on exocyclic amine with methyl	Ortoleva-Donnelly et al. (1998)
IαS (S$_p$)	0.1	1.0	WT	Replaces exocyclic amine with hydrogen	Strobel and Shetty (1997)
s^{6}GαS (S$_p$, 4 mM Mn^{2+})	0.25	1.0	WT	Replaces O6 with sulfur	Basu et al. (1998)
dGαS (S$_p$)	0.25	1.0	Y639F	Replaces 2′-OH with hydrogen	Szewczak et al. (1998)
^{OMe}GαS (S$_p$)	2.0	0.1	Y639F	Replaces hydrogen on 2′-OH with methyl	(unpub. observ.)
^{SH}GαS	0.5	1.0	Y639F	Replaces 2′-OH with 2′-SH	Schwans et al. (2003)
CαS (S$_p$ isomer only)[c]	0.05	1.0	WT		Christian and Yarus (1992)
f^{5}CαS	1.0	1.0	WT	Replaces C5 hydrogen with fluorine	Oyelere and Strobel (2000)
n^{6}CαS	1.0	0.5	WT	Replaces C6 with nitrogen	Oyelere and Strobel (2000)
ΨiCαS	0.1	1.0	WT	Replaces N1 with carbon and C5 with nitrogen	Oyelere and Strobel (2000)
ZαS	1.0	0.5	Y639F	Replaces N1 with carbon and exocyclic amine with hydrogen	Szewczak et al. (1999)

continued

Table 6.9.1 Phosphorothioate-Tagged Nucleotides for NAIM, *continued*

Analog δTPαS	[δTPαS] (mM)[a]	[NTP] (mM)[b]	Polymerase (WT or Y639F)	Description	Reference
dCαS (S$_p$)	0.75	1.0	Y639F	Replaces 2′-OH with hydrogen	Ryder and Strobel (1999b)
^{OMe}CαS	2.0	0.05	Y639F	Replaces hydrogen on 2′-OH with methyl	Conrad et al. (1995)
^{SH}CαS	0.4	1.0	Y639F	Replaces 2′-OH with 2′-SH	Schwans et al. (2003)
UαS (S$_p$ isomer only)[c]	0.05	1.0	WT		Christian and Yarus (1992)
m^{5}UαS	1.0	1.0	WT	Replaces H5 with methyl	Ryder and Strobel (1999a)
ΨαS	0.1	1.0	WT	Replaces C5 with amine	Ryder and Strobel (1999a)
dUαS (S$_p$)	0.25	1.0	Y639F	Replaces 2′-OH with hydrogen	Szewczak et al. (1998)
^{OMe}UαS	2.0	0.1	Y639F	Replaces hydrogen on 2′-OH with methyl	Conrad et al. (1995)
^{F}UαS	0.25	1.0	Y639F	Replaces 2′-OH with fluorine	Szewczak et al. (1998)
^{SH}UαS	0.4	1.0	Y639F	Replaces 2′-OH with 2′-SH	Schwans et al. (2003)

[a]Final (1×) concentration of phosphorothioate-tagged analog in transcription reaction.

[b]Final (1×) concentration of untagged, unmodified nucleotide in same transcription reaction.

[c]Phosphorothioate-tagged parent nucleotide (NTPαS) used for control transcript.

Ambion sells the parental phosphorothioates. Currently, there are 18 phosphorothioate-tagged nucleotide triphosphates available from Glen Research. Parental nucleotides are provided as the purified S$_P$ diastereomer, while nucleotide analogs are typically a diasteromeric mixture.

Many more nucleotide analogs can be utilized in a NAIM experiment than are available commercially. In the vast majority of cases, it is relatively straightforward to synthesize the phosphorothioate-tagged triphosphate beginning with the unprotected nucleoside (Fischer et al., 1999; Ryder et al., 2000). A wide variety of nucleoside analogs are commercially available from sources such as Sigma, Berry & Associates, or RI Chemical. In the first step, the nucleoside is converted to the 5′-*O*-(1,1-dichloro-1-thio)phosphorylnucleoside using PSCl$_3$. In the second step, the nucleotide is converted to the cyclotriphosphate intermediate by the stepwise addition of tributylammonium pyrophosphate, and the triphosphate is generated by hydrolysis with aqueous triethylammonium bicarbonate. A detailed synthesis with improved overall yields has been published by Fischer et al. (1999).

RNA Transcripts for NAIM

RNA molecules with randomly incorporated phosphorothioate nucleotides are readily prepared using run-off transcription from a DNA template, either a cleaved plasmid or a synthetic oligonucleotide. A Support Protocol presents a brief method for transcription and purification. Complete protocols for these activities are found elsewhere in this manual and in *Current Protocols in Molecular Biology*. (For digestion of plasmid DNA with restriction enzymes, see *CPMB UNIT 3.1*; for in vitro transcription of RNA, see *CPMB UNIT 3.8*.) For the most part, the analogs are incorporated into the RNA by the standard procedure for run-off transcription (see *CPMB UNIT 3.8*), except that a phosphorothioate-tagged analog triphosphate is included in the reaction. Depending upon the analog, it may also be necessary to make slight adjustments to the standard transcription cocktail, such as reduction of the non-tagged nucleotide triphosphate concentration, inclusion of

Mn^{2+} in the buffer, or use of a mutant polymerase. These adjustments are summarized in Table 6.9.1 for each of 36 analogs.

The conditions described in Table 6.9.1 should be viewed as a guideline. These values were derived from transcription of the *Tetrahymena* group I intron, but each RNA molecule has slightly different conditions under which it is efficiently transcribed. This necessitates transcriptional optimization of the RNA and optimization of the incorporation of individual analogs during the initial stages of the project. Phosphorothioate incorporation levels should be kept at ~5% (one analog incorporated per 20 occurrences of that residue in the sequence). The nature of the modification affects the concentration at which the nucleotide must be included to achieve this level of incorporation. Incorporation levels <5% make it difficult to detect the cleavage signal, while over-incorporation can cause multiple cleavage events within a single RNA. This is manifested as a phosphorothioate cleavage ladder in which the bands at the bottom of the gel are substantially more intense than those closer to the top.

The wildtype form of the T7 RNA polymerase is incapable of incorporating several analogs, most notably the 2′-deoxy nucleotides. Several of the transcriptionally resistant analogs can be incorporated using the Y639F mutant of the T7 RNA polymerase, a variant originally reported by R. Sousa and co-workers for its ability to synthesize DNA polymers (Sousa and Padilla, 1995; Padilla and Sousa, 1999). This mutant is able to incorporate several analogs with modifications in the minor groove (Soukup et al., 2002). It is sold commercially by Epicentre Technologies under the name T7 R&DNA Polymerase.

Radiolabeling RNA Transcripts and Substrates

See *UNIT 6.1* for both 3′-end-labeling using [^{32}P]pCp and T4 RNA ligase, and 5′-end-labeling using [γ-^{32}P]ATP and T4 polynucleotide kinase. The protocols outlined in *UNIT 6.1* utilize calf intestinal alkaline phosphatase to dephosphorylate the 5′ ends of the RNA transcripts prior to 5′-end-labeling. RNAs without a 5′-phosphate can also be successfully generated by including guanosine in the transcription reaction. A guanosine concentration is chosen that is two to three times that of the GTP concentration, up to the saturating guanosine concentration of 20 mM, to maximize the number of molecules with a free 5′-OH.

Choice of Selection Method and Expanded Applications

NAIM requires an assay to distinguish active from inactive variants within a substituted RNA population. In general, two types of selection method have been employed. The first is to physically separate active RNAs by gel mobility shift, column chromatography, filter binding, or denaturing polyacrylamide gel electrophoresis (Boudvillain and Pyle, 1998; Kazantsev and Pace, 1998; Basu and Strobel, 1999; Batey et al., 2000; Szewczak et al., 2002). This approach has been used in NAIM experiments to study protein-RNA interactions, RNA-RNA interactions, RNA folding, and ribozyme cleavage activity. The second approach is specific to analysis of ribozymes, wherein the inherent reactivity of the RNA is used to selectively radiolabel or change the size of active members in the population (Strobel and Shetty, 1997; Ryder et al., 2001).

The Basic Protocols presented here outline the methodology used initially to define the important functional groups within an RNA molecule (Ryder and Strobel, 1999a). The first three protocols present methods for selecting RNA based on ligand binding activity (see Basic Protocol 1), ribozyme activity (see Basic Protocol 2), or RNA structure or folding (see Basic Protocol 3). Basic Protocol 4 describes how the sites of interference are identified after RNA selection is complete.

Chemical and Enzymatic Probes for Nucleic Aicd Structure

6.9.5

Four Alternate Protocols are also presented that expand upon the information gained from the basic NAIM experiment. These include: (1) identification of tertiary contacts within a folded RNA by nucleotide analog interference suppression (NAIS; see Alternate Protocol 1; Szewczak et al., 1998); (2) identification of residues with a functionally important pK_a perturbation (see Alternate Protocol 2; Oyelere et al., 2002); (3) identification of functional groups likely to coordinate monovalent or divalent metal ions (see Alternate Protocol 3; Basu and Strobel, 2001); and (4) quantitation of the free energy contribution made by individual functional groups for ligand affinity (see Alternate Protocol 4; Cochrane et al., 2003).

Many of the experiments suggested here can be optimized for the RNA of interest. The selectivity of the procedure can be adjusted using various salt concentrations, ligand concentrations, and temperatures as desired. Each case must be considered on an individual basis. In the experiments outlined below, a starting point for selectivity is provided, but this number is just a recommendation that provides a balance between being too permissive and providing a reasonable signal for NAIM analysis.

SELECTION FOR LIGAND BINDING

There are many ways in which selection for ligand binding can be accomplished. In this protocol, filter binding allows for the selective retention of molecules that are able to perform a particular binding activity. Column chromatography, in which the ligand has been immobilized on beads, can be used for the same purpose. Native gel electrophoresis (see Basic Protocol 3) is another way in which RNA molecules that are able to bind ligand can be separated from those that cannot.

The selectivity of the ligand binding assay can be adapted to the peculiarities of any RNA-ligand system. Obviously, variations in the concentration of the RNA or the ligand will result in different levels of selectivity in the assay. A high RNA concentration and low ligand concentration will lead to the most selective conditions, while a high ligand concentration and low RNA concentration may mask all interferences. A ligand concentration approximately equal to the K_d for the complex and an RNA concentration ten-fold lower than the ligand results in ~50% of the RNA molecules being bound in a noncompetitive selection, which is ideal for the first sets of interference experiments.

Filter binding is a relatively straightforward method for separating molecules that are able to bind ligand from those that are not (Wong and Lohman, 1993). In this approach, the ligand must be able to bind efficiently to a nitrocellulose filter. The RNA is incubated with a ligand and then applied to a nitrocellulose membrane. The RNA is eluted from the filter, cleaved with iodine, and the fragments separated by gel electrophoresis. The specific activity and concentration of the RNA will dictate the volume of RNA-ligand mixture that will need to be applied to the membrane. The minimum activity required for visualization in a NAIM experiment is 5000 cpm in the final loaded sample.

Materials

Wash buffer (see recipe)
Transcribed ribozyme RNA containing phosphorothioate-tagged analogs (see Support Protocol and Strategic Planning)
Stock of ligand
500 mM Tris/HEPES buffer, pH 7.5 (see recipe)
2 M KCl
1 M $MgCl_2$
0.1 M dithiothreitol (DTT)

**Nucleotide
Analog
Interference
Mapping**

6.9.6

Supplement 17

0.1% Igepal C-680
1 mg/mL tRNA
Filter elution buffer (see recipe)
Ethanol (EtOH), cold

Nitrocellulose membrane (Fisher Scientific)
Hybond-N+ nylon filter (Amersham Pharmacia Biotech)
90°C heating block
1.5-mL microcentrifuge tubes
Whatman paper
Glass plate
Platform rocker

Additional reagents and equipment for denaturing filter binding assays (see *CPMB UNIT 12.8*) and denaturing gel electrophoresis (see *APPENDIX 3B*)

1. Soak nitrocellulose and Hybond-N+ filters in wash buffer.

   The Hybond-N+ filter retains RNA not bound to ligand.

2. Heat the labeled RNA for 1 min at 90°C, then place on ice for at least 3 min.

3. In a 1.5-mL tube incubate:

   $0.1 \times K_d$ labeled RNA
   $1 \times K_d$ ligand
   20 mM Tris/HEPES buffer, pH 7.5
   200 mM KCl
   10 mM MgCl$_2$
   1 mM DTT
   0.01% Igepal C-680
   0.1 mg/mL tRNA.

   Incubations should reach equilibrium. For many RNA-ligand complexes, 1 hr at room temperature is reasonable.

4. Assemble a filter binding apparatus as diagrammed in *CPMB UNIT 12.8*. Place a piece of Whatman paper directly on the bottom plate, then the Hybond-N+ membrane and then the nitrocellulose membrane. Place the top plate on the apparatus and clamp it.

5. Pipet 40 μL of the reaction mixture (from step 3) into the wells of the filter binding apparatus.

   For an experiment with ~10 pM RNA, 10,000 cpm/pmol and a ligand concentration around the K_d, twelve wells are needed to have sufficient signal in the final stages of the experiment.

6. Wash each well with 100 μL wash buffer.

7. Remove the top plate of the filter binding apparatus. Allow the filters to air dry for 1 to 2 min.

8. Remove the nitrocellulose filter and place it on a glass plate. Cut out the individual dot blots from the filter.

9. Place the blots into a 1.5-mL microcentrifuge tube and add 700 μL filter elution buffer. Rock the tube for 1 hr to allow the RNA to elute from the filter.

10. Remove the elution buffer from the tube and divide into two equal fractions.

11. Ethanol precipitate the RNA by adding 3 vol cold EtOH and incubating at $-80°C$ for 15 min. Centrifuge 15 min at $17,000 \times g$, $25°C$. Save the pellet and allow the RNA to air dry for 5 to 10 min.

12. Determine the positions of interference (see Basic Protocol 4).

ACTIVITY SELECTION FOR RIBOZYMES

Active ribozymes can be selected from an RNA population based on a change in the size of the molecule or transfer of the radiolabeled substrate onto itself (Cech, 1990). A general set of conditions for performing a ribozyme catalysis reaction is presented here, but it should be recognized that some molecules are more reactive than others and may require other co-factors, higher or lower reaction temperatures, or longer incubation times to achieve between 20% and 40% reacted ribozyme.

To use a ribozyme in a NAIM experiment, it must be self-catalytic either in ligation or cleavage (Jones et al., 2001). The main difference between these two activities is that the ribozyme must be labeled in the case of cleavage, while either the ribozyme or the substrate may be labeled in the case of the ligation. Ribozymes and substrates may be labeled at either the 5′ or 3′ end; however, during the course of the reaction, the label must either be retained by the ribozyme or be transferred to the ribozyme.

Materials

Transcribed ribozyme RNA containing phosphorothioate-tagged analogs (see Support Protocol and Strategic Planning)
500 mM HEPES buffer, pH 7.0
10 mM $MgCl_2$
10 mM $Mn(OAc)_2$
Labeled substrate molecule (optional; see Strategic Planning)
2× formamide loading buffer (FLB; *APPENDIX 2A*)
Gel elution buffer (see recipe)
Ethanol (EtOH), cold

0.65-mL microcentrifuge tubes
$37°$ and $50°C$ heating blocks
Autoradiography film
Platform rocker

Additional reagents and equipment for denaturing gel electrophoresis (see *APPENDIX 3B*)

1. In a 0.65-mL microcentrifuge tube, mix:

5 μL 1.0 μM RNA transcript
1 μL 500 mM HEPES buffer, pH 7.0
3 μL 10 mM $MgCl_2$
1 μL 10 mM $Mn(OAc)_2$.

Incubate reaction 10 min at $50°C$.

This is the RNA folding step. Alternative temperatures, incubation times, and salt concentrations should also be used as appropriate for the RNA of interest.

2. For ligating ribozymes, add 10 μL of preheated ($37°C$) 10 nM substrate dissolved in 50 mM HEPES buffer, pH 7.0, 3 mM $MgCl_2$, and 1 mM $Mn(OAc)_2$.

3. Incubate the reaction at 37°C for a time sufficient for 20% to 40% of the ribozyme to react.

   To determine the length of time required to achieve 20% to 40% reactivity, monitor the reaction by taking aliquots at time intervals over several minutes. Electrophorese the time points on a denaturing polyacrylamide gel and determine the fraction reacted over time.

4. Quench the reaction in 2 vol FLB.

   If the label was transferred from the substrate to the ribozyme, the sample can be used directly to determine the sites of interference. Otherwise, it is necessary to physically separate the active from inactive fraction based upon the size of the RNA using the radio-label to visualize the active population.

5. For all experiments in which the initial label was on the ribozyme, separate the reaction products from the substrates by denaturing polyacrylamide gel electrophoresis (see APPENDIX 3B).

6. Expose the gel to autoradiography film and excise the RNA from the gel by physically cutting the products from the gel. Elute RNA in 700 μL gel elution buffer by rocking for 2 to 4 hr at room temperature.

   Shorter incubation times are used to minimize the extent of RNA degradation. Lower temperatures and inclusion of detergents (1% SDS) in the elution buffer can also be used toward this goal.

7. Ethanol precipitate the RNA by adding 3 vol cold EtOH and incubating 15 min at −80°C. Centrifuge 15 min at $17,000 \times g$, 25°C. Save the pellet and allow the RNA to air dry for 5 to 10 min.

8. Determine the extent of interference at all positions (see Basic Protocol 4).

SELECTION FOR STRUCTURE OR FOLDING

Selecting for structured RNA molecules is a bit more difficult than selecting for RNA molecules that are able to perform a catalytic activity, as there is no length difference between active and inactive molecules. Native gel electrophoresis conditions that separate the RNA based upon shape must be identified (Basu and Strobel, 1999). Ideal separation conditions for a particular RNA system are often variable, but, in general, the gel solution should not contain denaturants such as urea, and should include polycations such as magnesium to promote folding of the RNA. The electrophoresis is typically performed at a lower voltage and the running buffer is cooled to promote retention of RNA structure within the gel.

Materials

Transcribed ribozyme RNA containing phosphorothioate-tagged analogs (see Support Protocol and Strategic Planning)
500 mM Tris/HEPES buffer, pH 7.5 (see recipe)
10 mM $MgCl_2$
50% glycerol
Native gel mix (see recipe)
Gel elution buffer (see recipe)
Ethanol (EtOH), cold

0.65-mL microcentrifuge tubes
70°C heating block
Autoradiography film
Platform rocker

Additional reagents and equipment for native gel electrophoresis (*UNIT 11.4*)

1. In a 0.65-mL microcentrifuge tube, mix:

    5 μL 1 mM RNA transcript
    2 μL 500mM Tris/HEPES buffer, pH 7.5
    4 μL 10 mM MgCl$_2$
    4 μL 50% glycerol
    5 μL ddH$_2$O.

2. Incubate the RNA 5 min at 70°C and cool slowly to room temperature.

    The folding conditions presented here are a suggestion, and should be optimized for the RNA molecule of interest.

3. Perform native gel electrophoresis as described in UNIT 11.4 using a gel mix optimized for the system of interest.

4. Expose the gel to autoradiography film and excise the RNA by physically cutting the products from the gel. Elute the RNA in 700 μL gel elution buffer by rocking for 2 to 4 hr at room temperature.

    Identification of the folded population can be made by comparing the substituted RNA reactions to those done with wild-type RNA.

5. Ethanol precipitate the RNA by adding 3 vol cold EtOH and incubating 15 min at −80°C. Centrifuge 15 min at 17,000 × g, 25°C. Save the pellet and allow the RNA to air dry for 5 to 10 min.

6. Determine the extent of interference at all positions (see Basic Protocol 4).

TRANSCRIPTION AND PURIFICATION OF PHOSPHOROTHIOATE-TAGGED RNA

This protocol describes run-off transcription for incorporating phosphorothioate-tagged nucleotides into the desired RNA. For each RNA incorporating a modified analog (δαS), a separate RNA incorporating the phosphorothioate-tagged parent nucleotide (NαS) will be needed as a NAIM control.

Materials

    10× transcription buffer (see recipe)
    1 M MgCl$_2$ (store at 4°C)
    20× NTPs: 20 mM each ATP, CTP, GTP, and UTP (see Table 6.9.1)
    10× NTPαS: modified (δαS) or parent (NαS) phosphorothioate-tagged nucleotide (see Table 6.9.1)
    1 μg/μL linearized template DNA, containing a T7 promoter upstream of the RNA gene to be transcribed
    0.1 mg/mL inorganic pyrophosphatase (store at −20°C)
    500 to 1000 U/μL T7 RNA polymerase (store at −20°C)
    2× formamide loading buffer (FLB; APPENDIX 2A)
    Gel elution buffer (see recipe)
    3 M sodium acetate (NaOAc), pH 5.2 (APPENDIX 2A)
    100% ethanol (EtOH), prechilled at −20°C
    T$_{10}$E$_{0.1}$ buffer, pH 7.5 (see recipe)

    1.5-mL microcentrifuge tubes
    37°C heating block
    Platform rocker

    Additional reagents and equipment for denaturing gel electrophoresis (APPENDIX 3B)

1. In a 1.5-mL microcentrifuge tube, combine:

   10 μL 10× transcription buffer
   2 μL 1 M MgCl$_2$
   5 μL 20× NTPs
   10 μL 10× NTPαS
   2 μL 1 μg/μL linearized DNA template
   1 μL 0.1 mg/mL inorganic pyrophosphatase
   5 μL 500 to 1000 U/μL T7 RNA polymerase
   65 μL ddH$_2$O.

   Incubate 1 to 2 hr at 37°C.

   The final 1× concentrations of tagged, modified nucleotides (δTPαS) are indicated in Table 6.9.1. For NAIM control transcripts, the 1× concentration of tagged parent nucleotide (AαS, CαS, GαS, or UαS) is 0.05 mM. For both modified and control transcripts, the 1× concentration of unmodified, untagged nucleotides (from the NTP mix) should be 1 mM in most cases, but see Table 6.9.1 for some exceptions.

   Any of the above concentrations can be altered to better suit the template being used in transcription, particularly the Mg^{2+} concentration. Larger volumes can be transcribed if more RNA is needed for the particular assay being run. The expected yield per transcription is ∼0.1 to 1 nmol/mL. If higher NTP concentrations are required for RNA transcription, proportional adjustments should be made to the analog and Mg^{2+} concentrations.

2. Add 100 μL of 2× formamide loading buffer to RNA transcripts and proceed with preparative denaturing gel electrophoresis (e.g., *APPENDIX 3B*).

3. Visualize the RNA using UV shadowing on a silica plate and physically remove the gel slice containing the RNA. Elute RNA in 700 μL gel elution buffer by rocking for 1 to 2 hr at room temperature.

4. Transfer the elution buffer to a clean 1.5-mL microcentrifuge tube and add 0.1 vol of 3 M NaOAc, pH 5.2, and 3 vol of 100% EtOH. Place 15 min at −80°C, then centrifuge 15 min at 17,000 × *g*, 25°C.

5. Carefully remove ethanol from tube, without disturbing the RNA pellet. Allow the pellet to dry for 5 to 10 min.

6. Resuspend the RNA pellet in 50 μL T$_{10}$E$_{0.1}$ buffer, pH 7.5.

7. Store RNA at −20°C.

IDENTIFYING POSITIONS OF INTERFERENCE AND CONTROLLING FOR THE LEVEL OF INCORPORATION

BASIC PROTOCOL 4

In this protocol, the selected RNA pools are treated with iodine to cleave at the site of the phosphorothioate tag. The cleaved RNAs are then separated by denaturing gel electrophoresis, allowing identification of the sites of analog interference.

The key to analysis of NAIM data is the inclusion of all relevant samples on a single gel. For every analog tested in a NAIM experiment, the relevant gel must also contain samples in which the parent nucleotide (i.e., unmodified NαS) has been tested in a parallel experiment. The gel will also have samples for both parent and analog nucleotides that have not been iodine treated. These serve as a control for nonspecific degradation of the RNA. Positions of strong degradation in this lane are uninformative in the NAIM assay, and cannot have any interference value determined.

Chemical and Enzymatic Probes for Nucleic Aicd Structure

6.9.11

An additional element in data analysis is the inclusion of RNA transcripts that have not been subjected to any selection step, but are end-labeled, iodine-treated, and analyzed on a sequencing gel. These unselected RNA molecules control for the overall level of analog incorporation and the specific level of incorporation at a particular residue in the sequence. If the incorporation level is too high, the bands at the bottom of the gel will be much stronger than those at the top of the gel, as multiple iodine-induced cleavages occur in each RNA; this should be avoided. The level of incorporation of a particular phosphorothioate nucleotide is closely correlated to the relative concentrations of the phosphorothioate and untagged nucleotide in the transcription mixture. Therefore, if over-incorporation is a problem, the concentration of the analog should be decreased or the untagged nucleotide increased in the transcription mixture. It should be noted that the cleavage of the phosphorothioate bond by iodine has an end point of ~15% reacted. Consequently, the majority of the RNA species will appear as full length on a sequencing gel unless there are too many substitutions in the RNA. The unselected RNA molecules also control for the level of incorporation at each position. There is sequence-specific variability for the extent of incorporation that is stronger for some analogs than for others. This includes sites of hyper- and non-incorporation, but information can still be gained from such positions because of the controls. These unselected controls do not have to be included on every gel, but must be performed every time an RNA molecule is transcribed and used in a NAIM reaction.

Any interferences observed with the tagged parent nucleotide can be attributed to a phosphorothioate effect. In the strongest cases, interferences at these sites are uninformative for subsequent analog studies, and an effort should be made to reduce them as much as possible. The inclusion of manganese in the reaction buffer sometimes serves to minimize phosphorothioate effects.

Quantitation of individual band intensities in the I_2 cleavage ladder is used to identify sites of interference resulting from analog substitution. The data are also normalized for loading differences and variation in the extent of the reaction. The following protocol calculates the extent of analog interference at each position within the RNA.

Materials

2× formamide loading buffer (FLB; *APPENDIX 2A*)
Selected RNA samples containing phosphorothioate-tagged analogs (see Basic Protocols 1, 2, and 3)
Parallel unselected RNA samples (control)
50 mM I_2/EtOH: 50 mM I_2 dissolved in ethanol

Phosphorimager and screen (Molecular Dynamics)
ImageQuant software

Additional reagents and equipment for denaturing gel electrophoresis (*APPENDIX 3B*)

1. Add 2 vol FLB to the selected RNA containing phosphorothioate-tagged analogs.

2. Split the selected RNA into two fractions. To one, add $^1/_{10}$ vol of 50 mM I_2/EtOH.

3. Heat the reaction 2 min at 90°C to denature the RNA.

4. Load all the reactions onto a denaturing polyacrylamide gel (see *APPENDIX 3B*). Run the gel for varying lengths of time, depending on the length of the RNA molecule and the section of the RNA that is being studied.

5. Dry the gel and expose to a phosphorimager screen.

6. Quantitate the peak intensities at each position of incorporation for both the parental nucleotide (NαS) and the nucleotide analog (δαS) observed in the ligation reaction by phosphorimager analysis using area integration. Within the ImageQuant software, draw an equivalent line through every lane on the gel and graph the radioactive intensity versus distance. Carefully define the baseline and make sure it is consistently defined between each of the lanes. Find the peaks by adjusting the noise and sensitivity variables. Compute the area under each peak and use these values in Equation 6.9.1 for each of the different variables and sites of incorporation.

7. Calculate the extent of interference at each position by substituting the individual band intensities into Equation 6.9.1.

$$\kappa = NF \times \frac{(N\alpha S\,experiment\,/\,\delta\alpha S\,experiment)}{(N\alpha S\,unselected\,/\,\delta\alpha S\,unselected)}$$

Equation 6.9.1

where at each site of incorporation, NαS experiment is the peak intensity for the parental selected RNA, δαS experiment is the peak intensity for the analog selected RNA, NαS unselected is the peak intensity for the parental unselected RNA, δαS unselected is the peak intensity for the analog unselected RNA, and NF is a normalization factor used to account for differences between loading and extent of reaction between lanes.

Values are initially calculated without the normalization factor, but then are normalized against each other to calculate κ. NF is determined by calculating the average interference value for all positions that are within two standard deviations of the mean, and dividing the value at each position by this average. NF is typically between 0.8 and 1.2.

Equation 6.9.1 provides a value defined as κ that indicates the extent of interference at each site of incorporation within the sequence. A κ value >2 defines a site of analog interference. A value <0.5 defines a site of enhancement, and a value of 1 defines a residue unaffected by the substitution. The noise in the κ measurement is such that values between 0.5 and 2 are positions of little or no interference, though this is a fairly conservative interpretation.

NUCLEOTIDE ANALOG INTERFERENCE SUPRESSION (NAIS)

Some of the interferences that are observed in a traditional NAIM experiment arise from the loss of a hydrogen bond and the consequent destabilization of the structure of the RNA. NAIM identifies the functional groups involved in hydrogen bonding, but NAIS can map tertiary contacts by identifying which functional groups are paired to each other in the active state of the complex (Boudvillain and Pyle, 1998; Strobel et al., 1998; Szewczak et al., 1998; Strobel and Ortoleva-Donnelly, 1999; Szewczak et al., 1999; Soukup et al., 2002).

NAIS can be used in conjunction with any of the previous types of selection. A mutation in the RNA or an alteration in a ligand can be used in combination with NAIM to identify the tertiary interactions within that complex. The approach utilizes a single mutation or functional group substitution at one important residue in the complex and phosphorothioate-tagged analogs throughout the RNA.

NAIS has been used extensively to identify hydrogen-bonded functional groups within ribozymes. Chimeric RNA molecules are constructed from a synthetic oligonucleotide containing the site-specific deletion of a functional group of interest and a transcribed RNA. These chimeric RNAs are ligated as outlined in *CPMB UNIT 3.15*. The RNA is transcribed with a phosphorothioate-tagged nucleotide analog that deletes the other functional group predicted to be involved in the tertiary interaction. NAIS utilizes the prediction that there will not be a further energetic penalty incurred upon the deletion of the second

Chemical and Enzymatic Probes for Nucleic Aicd Structure

6.9.13

member of a hydrogen-bonded pair and, as a result, the interference will be suppressed. This assay is dependent on the ability to reconstruct the selection assay such that there is activity despite the deletion of the functional group that initially gave rise to the interference. Positions of interference that arise from the hybrid molecules are identified and compared to those from RNAs that retain the specific functional group. Any interference that is suppressed upon the addition of the site-specific substitution is a strong candidate for involvement in a hydrogen bond with the modified functional group.

Additional controls must be included to show that the suppression of the interference is specific to both the nature of the substitution and the site of suppression. Ideally, this should include the following demonstrations of specificity. (1) The suppression must be specific to the site-specifically modified functional group. Mutations or site-specific substitutions at other residues should not lead to the same suppression pattern. (2) The suppression must be specific to a single residue among additional sites of interference. (3) The suppression must be specific to a single or closely related set of analogs.

For materials for selection, see Basic Protocol 1; for identification of sites of interference, see Basic Protocol 4; and for ligation of RNA molecules using T4 DNA ligase, see *CPMB UNIT 3.15*.

1. Use NAIM to identify key functional groups in the RNA.

2. Follow the protocol in *CPMB UNIT 3.15* for ligation of RNA molecules using T4 DNA ligase.

3. Use the hybrid RNA in a selection reaction as outlined previously.

    Because of the inherently lower activity of the hybrid RNA caused by the site-specific substitution, the selection reaction may need to be modified. Many variables, such as temperature, salt, and pH, can be altered to achieve a higher level of activity. This assay can only be performed if the substituted RNA retains some residual level of activity.

4. Identify the sites of interference as described. Compare κ values obtained using wild type ribozymes to those obtained using the hybrid RNA molecules.

    Sites that have high κ values in the wild-type experiment and lower values in the hybrid experiment are excellent candidates for tertiary contacts with the modified or mutated residue.

FUNCTIONAL GROUP IONIZATION DETERMINED BY NAIM

Several of the analog nucleotides used in NAIM experiments contain base modifications that, in addition to altering a functional group in the molecule, have altered pK_a values from the parent nucleotide (Oyelere and Strobel, 2000; Ryder et al., 2001; Oyelere et al., 2002; Jones and Strobel, 2003). The nucleotides used in this experiment are analogs of adenosine and cytosine. These analogs make it possible to screen for functionally important base ionization at every adenosine or cytosine residue in an RNA. The sites of ionization can play a role in either the structure of the RNA or, for those RNAs that are catalytic, in the chemistry of the molecule.

The adenosine analogs used to probe protonation at the N1 position include: adenosine ($pK_a = 3.5$), purine ($pK_a = 2.1$), 8-azaadenosine ($pK_a = 2.2$), formycin A ($pK_a = 4.4$), and 7-deazaadeonsine ($pK_a = 5.2$). The cytosine analogs used to identify protonation events at the N3 position include: cytidine ($pK_a = 4.2$), 6-azacytidine ($pK_a = 2.6$), 5-fluorocytidine ($pK_a = 2.3$), and pseudoisocytidine ($pK_a = 9.4$). These analogs are used in conjunction with each other to probe for the ionization of adenosine and cytosine nucleotides.

On the basis of previous studies with these analogs, the following interference pattern is expected at individual adenosine residues whose ionization is functionally important.

At elevated pH, interference is seen with PurαS and n^{8}AαS. The interference is at least partially rescued by lowering the pH of the buffer. Furthermore, enhanced activity may be seen with FormAαS or 7dAαS at sites of base ionization due to the increased basicity of these nucleosides, if base ionization is a limiting event in the reaction.

At those positions where the N3 ionization of cytidine within the RNA is important for activity, both n^{6}CαS and f^{5}CαS have been shown to cause interference at high pH due to the increased acidity of both analogs (Oyelere et al., 2002). But as is the case for the adenosine analogs, the interference can be rescued as the pH is reduced. Conversely, ΨiCαS will show enhanced activity at elevated pHs, but the effect will be diminished as the pH is reduced.

METAL BINDING FUNCTIONAL GROUPS IDENTIFIED BY NAIM

The use of sulfur-substituted analogs enables NAIM to be used in metal ion rescue experiments (Basu et al., 1998; Basu and Strobel, 2001). Substitution with sulfur disturbs metal ion binding sites, as sulfur is unable to form stable complexes with magnesium or potassium (Heitner et al., 1972; Douglas et al., 1990). Interferences seen with parental nucleotides may arise from disruption of metal ion binding to the pro-R$_P$ oxygen of the RNA backbone. Identification of guanosine-O^6-carbonyl ligands for metal ions is done with the analog 6-thioguanosine phosphorothioate (s^{6}GαS). Inclusion of thiophilic metals such as Mn^{2+}, Cd^{2+}, Co^{2+}, Ni^{2+}, Zn^{2+}, and Tl$^+$ in the selection reaction ameliorates the interference caused by the change from oxygen to sulfur.

Materials

RNA molecules transcribed in the presence of sulfur-substituted nucleotides
100 mM Mn(OAc)$_2$
100 mM TlOAc
Other mono- and divalent metals as required
2% 2-mercaptoethanol

Additional reagents and equipment for selection (see Basic Protocol 1) and for identification of sites of interference (see Basic Protocol 4)

1. Perform the selection activity at various concentrations of thiophilic metal ion concentration, from 0.1 to 5 mM or higher, if needed.

   Any buffers or salts with chloride as a counter ion cannot be used for thallium rescue experiments, as thallium chloride is insoluble.

   CAUTION: *Care should be taken when handling Tl$^+$ as it is toxic.*

2. Control for specificity of the metal ion rescue by using other mono- and divalent metal ions in the selection assay.

   To show that rescue is specific to the metal ion forming a stable complex with sulfur, and not merely overall structure stabilization by metal ions, other non-thiophilic metals such as K$^+$, NH$_4^+$, Li$^+$, Na$^+$, Ca^{2+}, and Sr^{2+} at high concentrations (up to 100 mM) should be included in the reaction buffers.

3. Prepare the samples for loading on a denaturing polyacrylamide gel. For RNA samples containing s^{6}GαS, add $^1/_{10}$ vol of 2% 2-mercaptoethanol directly before loading the samples on the gel in order to reduce streakiness in the gel.

4. Identify the sites of interference. Compare the κ values obtained for sites of interference at each metal ion concentration. Specific rescue of an interference by a single thiophilic metal suggests this functional group is a ligand for metal ion binding.

Chemical and Enzymatic Probes for Nucleic Aicd Structure

6.9.15

QUANTITATIVE NUCLEOTIDE ANALOG INTERFERENCE MAPPING (QNAIM)

QNAIM is a series of NAIM experiments performed at several ligand concentrations (Cochrane et al., 2003), which allows characterization of the energetics of binding. The RNA concentration is held substantially below the dissociation constant of the complex (K_d), and the ligand concentration is varied from one to several times the K_d concentration. Under these conditions, RNA-ligand binding is driven by the increase in the ligand concentration, which is already in vast excess. As a result, the RNAs do not compete with each other for binding; instead, the distribution of RNAs in the bound fraction results entirely from the intrinsic affinity of each RNA for the ligand. The population distribution in the bound fraction is revealed by the individual band intensities observed in the sequencing ladder that results from iodine treatment. By plotting the interference magnitude as a function of ligand concentration, it is possible to calculate the binding constant and free energy contribution ($\Delta\Delta G$) of each functional group substitution at each position in the RNA.

Because of slight differences in salt and ligand concentrations, the binding constant can vary between experiments and should be determined separately for each experiment. The binding constant for the bulk RNA population is determined using a dot blot method.

Materials

$1\times K_d$ RNA (labeled)
500 mM Tris/HEPES buffer, pH 7.5 (see recipe)
2 M KCl
1 M $MgCl_2$
0.1 M DTT
0.1% Igepal C-680
1 mg/mL tRNA
$200\times K_d$ stock of ligand
Ligand dilution buffer (see recipe)
Wash buffer (see recipe)

Whatman paper
Cling film
Phosphorimager and screen

Additional reagents and equipment for dot blots (see *CPMB UNIT 2.9B*), for selection of active molecules based on ligand binding (see Basic Protocol 1), and for determining sites of interference (see Basic Protocol 4)

Determine binding constant

1. Heat 120 μL of $1\times K_d$ RNA to 90°C for 1 min; cool on ice for at least 3 min.

2. To the RNA, add:

48 μL 500 mM Tris/HEPES, pH 7.5
120 μL 2 M KCl
12 μL 1 M $MgCl_2$
15 μL 0.1 M DTT
120 μL 0.1% Igepal C-680
120 μL 1 mg/mL tRNA
645 μL ddH$_2$O.

3. Make twelve 20-µL 1:2 serial dilutions from the $200\times K_d$ stock of ligand into ligand dilution buffer.

4. Add 80 µL of the RNA mix from step 2 into each of the twelve serial dilutions of ligand.

5. Pipet 40 µL from each ligand concentration into a separate dot blot hole and wash with 100 µL wash buffer.

6. Dismantle the dot blot apparatus and allow the filters to air dry for 1 to 2 min. Place the filters on Whatman paper, cover with cling film, and expose to a phosphorimager screen.

7. Quantitate the data by comparing the density of counts in the bound fraction to the unbound fraction at every ligand concentration. Determine the binding constant (K_d) by fitting the data to Equation 6.9.2:

$$\text{fraction bound} = \frac{(M \times [\text{protein}])}{(K_d \text{ bulk} + [\text{protein}])} + L$$

Equation 6.9.2

where K_d is the apparent equilibrium binding constant, M is the maximum fraction of RNA bound at the highest protein concentration (this varies from 0.7 to 0.9), and L is the minimum fraction of RNA bound at the lowest protein concentration (this is usually very close to 0).

Perform QNAIM

8. For each ligand concentration, follow the protocol outlined in Basic Protocol 1.

9. Determine the sites of interference by denaturing polyacrylamide gel electrophoresis as outlined in Basic Protocol 4.

10. Once the sites of interference have been found, determine $1/\kappa$ values for each site at every protein concentration.

11. Use these interference values to determine the relative binding constant for each RNA that has a substitution at a particular position:

$$1\!\!\Big/\!\kappa = \frac{(M \times [\text{protein}])}{(K_d \text{ interference} + [\text{protein}])} + L$$

Equation 6.9.3

where K_d is the apparent equilibrium binding constant, M is the maximum level of rescue of the particular interference (or the upper baseline), and L is the maximum extent of interference for the same position (or the lower baseline).

12. Determine the change in binding free energy ($\Delta\Delta G$) from Equation 6.9.4 using the measured bulk binding constant and the binding constant determined for the particular interference band:

$$\Delta\Delta G = -RT \times \ln (K_d \text{ bulk}/K_d \text{ interference})$$

Equation 6.9.4

where R is the gas constant and T is temperature.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps. For common stock solutions, see APPENDIX 2A; for suppliers, see SUPPLIERS APPENDIX.

Filter elution buffer

300 mM sodium acetate (NaOAc), pH 5.2 (*APPENDIX 2A*)
1% SDS (*APPENDIX 2A*)
$T_{10}E_{0.1}$ buffer (see recipe)
Store up to 3 months at 25°C

Gel elution buffer

2% SDS (*APPENDIX 2A*)
$T_{10}E_{0.1}$ buffer (see recipe)
Store up to 3 months at 25°C

Ligand dilution buffer

20 mM Tris/HEPES buffer, pH 7.5 (see recipe)
200 mM KCl
10 mM $MgCl_2$
0.1% Igepal C-680
Store up to 3 months at 4°C

Native gel mix

6% acrylamide
1× TBE electrophoresis buffer (*APPENDIX 2A*)
3 mM $MgCl_2$
Store up to 3 months at 25°C

$T_{10}E_{0.1}$ buffer, pH 7.5

10 mM Tris·Cl, pH 7.5 (*APPENDIX 2A*)
0.1 mM EDTA (*APPENDIX 2A*)

Transcription buffer, 10×

40 mM Tris·Cl, pH 7.5 (*APPENDIX 2A*)
1 mM spermidine
5 mM dithiothreitol
0.01% Triton X-100
Store up to 3 months at −20°C

Tris/HEPES buffer, 500 mM, pH 7.5

500 mM Tris
500 mM HEPES
Adjust pH to 7.5
Store up to 3 months at 25°C

Wash buffer

20 mM Tris/HEPES buffer, pH 7.5 (see recipe)
200 mM KCl
10 mM $MgCl_2$
Store up to 3 months at 4°C

COMMENTARY

Background Information

NAIM is a generalizable chemogenetic approach that identifies functional groups that are important for the activity of an RNA (Strobel and Shetty, 1997). This method uses a series of 5'-O-(1-thio) nucleoside analogs to probe the contribution specific functional groups make to RNA activity. NAIM utilizes a pool of randomly substituted RNAs generated by in vitro transcription, and the active RNAs are identified through a selection experiment. In this way, a particular RNA can be screened with multiple analogs in a time-efficient manner. NAIM is applicable to any RNA with a selectable function. It can be used to study RNA catalysis through cleavage or ligation (Strobel and Shetty, 1997; Boudvillain and Pyle, 1998; Kazantsev and Pace, 1998; Sood et al., 1998; Jones et al., 2001; Kaye et al., 2002), RNA interactions with protein and other ligands (Batey et al., 2001), as well as the steps of RNA folding (Basu and Strobel, 1999). It has even been used to look at RNP assembly in vivo (Szewczak et al., 2002). This approach should make it possible to identify the chemical groups important for a wide variety of RNA activities.

Critical Parameters

NAIM requires an assay that distinguishes active members of a substituted RNA population from less active members. Two types of methods have been used to achieve this separation. The first is to physically separate active from inactive RNAs in the course of the assay. The second approach, specific to ribozymes, is to make use of inherent RNA activity to selectively radiolabel active members in the population.

In general, the second type of assay is easier to perform, as one preparation of labeled substrate will allow interference mapping with many different analog-containing RNA transcripts. Experiments that require multiple gels are undesirable because they increase the time necessary for analysis. Additionally, longer experimental protocols could lead to increased levels of RNA degradation, which will interfere with quantitation of band intensity. It is important to design the RNA being studied by NAIM so that it is of sufficient length to contain several sites of analog substitution that do not show interference. These provide important internal controls in the interference experiments to account for differences in loading volume and extent of reaction between experiments.

These positions are used to calculate the normalization factor (NF) in Equation 6.9.1. NF becomes more accurate as the population of sites that do not show interference increases. These issues constrain the length of the RNA to no less than 40 to 50 nucleotides.

Once an appropriate assay has been designed, it is necessary to optimize the reaction in order to obtain the desired level of selectivity. This can be done by varying several aspects of the experimental conditions. For example, it has been observed that high concentrations of divalent metal ions and lower reaction temperatures tend to mask sites of weaker interference in the group I intron. This can be attributed to overall improvement in the stability of the RNA fold. In all cases, a spectrum of reaction conditions should be explored with a subset of analogs prior to completion of a full screen. Furthermore, sites that show phosphorothioate interference are uninformative in a NAIM assay, so some effort should be made to find conditions that minimize these effects, such as the inclusion of Mn^{2+} in the reaction buffer.

Anticipated Results

NAIM is a versatile method, used to quickly identify important functional groups within a structured RNA molecule. In conjunction with the other techniques outlined in this unit, a NAIM experiment can provide detailed structural and energetic information on any in vitro transcribed RNA molecule.

A typical NAIM experiment will yield information about the contribution to catalysis and/or stability of an individual functional group at every possible position within an RNA molecule. The magnitude of the interference (κ) will provide additional information about the relative importance of that modification when compared to other sites within the same molecule. By using the full range of nucleotide analogs available, the key functional groups in the RNA molecule can be specifically elucidated. Furthermore, information about hydrogen bonding partners (NAIS), metal ion ligands, pK_a-perturbed nucleotides, and the energetics of ligand binding (QNAIM) can be determined.

NAIM data analysis requires the comparison of the experimental data to several reference data sets. By carefully controlling for the incorporation level at a particular site and the effect of the phosphorothioate tag, a wealth of detailed data can be obtained from a single NAIM experiment.

Time Considerations

The NAIM experiment itself can be completed in a single day, once the necessary materials are assembled. Due to the inherent instability of RNA molecules, it is preferable to complete the experiment as quickly as possible. However, RNA molecules in which phosphorothioate-tagged nucleotides have been incorporated are no more prone to degradation than normal RNA molecules, and may be slightly less so.

Acknowledgments

Literature Cited

Basu, S. and Strobel, S.A. 1999. Thiophilic metal ion rescue of phosphorothioate interference within the *Tetrahymena* ribozyme P4–P6 domain. *RNA* 5:1399-1407.

Basu, S. and Strobel, S.A. 2001. Biochemical detection of monovalent metal ion binding sites within RNA. *Methods* 23:264-275.

Basu, S., Rambo, R.P., Strauss-Soukup, J., Cate, J.H., Ferre-D'Amare, A.R., Strobel, S.A., and Doudna, J.A. 1998. A specific monovalent metal ion integral to the AA platform of the RNA tetraloop receptor. *Nat. Struct. Biol.* 5:986-992.

Batey, R.T., Rambo, R.P., Lucast, L., Rha, B., and Doudna, J.A. 2000. Crystal structure of the ribonucleoprotein core of the signal recognition particle. *Science* 287:1232-1239.

Batey, R.T., Sagar, M.B., and Doudna, J.A. 2001. Structural and energetic analysis of RNA recognition by a universally conserved protein from the signal recognition particle. *J. Mol. Biol.* 307:229-246.

Boudvillain, M. and Pyle, A.M. 1998. Defining functional groups, core structural features and inter-domain tertiary contacts essential for group II intron self-splicing: A NAIM analysis. *EMBO J.* 17:7091-7104.

Cech, T.R. 1990. Self-splicing of group I introns. *Annu. Rev. Biochem.* 59:543-568.

Chamberlin, M., Kingston, R., Gilman, M., Wiggs, J., and deVera, A. 1983. Isolation of bacterial and bacteriophage RNA polymerases and their use in synthesis of RNA in vitro. *Methods Enzymol.* 101:540-568.

Cochrane, J.C., Batey, R.T., and Strobel, S.A. 2003. Quantitation of free energy profiles in RNA-ligand interactions by nucleotide analog interference mapping. *RNA* 9:1282-1289.

Conrad, F., Hanne, A., Gaur, R.K., and Krupp, G. 1995. Enzymatic synthesis of 2′-modified nucleic acids: Identification of important phosphate and ribose moieties in RNase P substrates. *Nucl. Acids Res.* 23:1845-1853.

Douglas, K.T., Bunni, M.A., and Baindur, S.R. 1990. Thallium in biochemistry. *Int. J. Biochem. Cell Biol.* 22:429-438.

Eckstein, F. 1985. Nucleoside phosphorothioates. *Annu. Rev. Biochem.* 54:367-402.

Fischer, B., Chulkin, A., Boyer, J.L., Harden, K.T., Gendron, F.P., Beaudoin, A.R., Chapal, J., Hillaire-Buys, D., and Petit, P. 1999. 2-Thioether 5′-O-(1-thiotriphosphate)adenosine derivatives as new insulin secretagogues acting through P2Y-Receptors. *J. Med. Chem.* 42:3636-3646.

Gaur, R.K. and Krupp, G. 1993. Enzymatic RNA synthesis with deoxynucleoside 5′-O-(1-thiotriphosphates). *FEBS Lett.* 315:56-60.

Gish, G. and Eckstein, F. 1988. DNA and RNA sequence determination based on phosphorothioate chemistry. *Science* 240:1520-1522.

Griffiths, A.D., Potter, B.V., and Eperon, I.C. 1987. Stereospecificity of nucleases towards phosphorothioate-substituted RNA: Stereochemistry of transcription by T7 RNA polymerase. *Nucl. Acids Res.* 15:4145-4162.

Hardt, W.D., Erdmann, V.A., and Hartmann, R.K. 1996. R_p-deoxy-phosphorothioate modification interference experiments identify 2′-OH groups in RNase P RNA that are crucial to tRNA binding. *RNA* 2:1189-1198.

Heitner, H.I., Sunshine, H.R., and Lippard, S.J. 1972. Metal binding by thionucleosides. *J. Am. Chem. Soc.* 94:8936-8937.

Jones, F.D. and Strobel, S.A. 2003. Ionization of a critical adenosine residue in the neurospora Varkud Satellite ribozyme active site. *Biochemistry* 42:4265-4276.

Jones, F.D., Ryder, S.P., and Strobel, S.A. 2001. An efficient ligation reaction promoted by a Varkud Satellite ribozyme with extended 5′- and 3′-termini. *Nucl. Acids Res.* 29:5115-5120.

Kaye, N.M., Christian, E.L., and Harris, M.E. 2002. NAIM and site-specific functional group modification analysis of RNase P RNA: Magnesium dependent structure within the conserved P1-P4 multihelix junction contributes to catalysis. *Biochemistry* 41:4533-4545.

Kazantsev, A.V. and Pace, N.R. 1998. Identification by modification-interference of purine N-7 and ribose 2′-OH groups critical for catalysis by bacterial ribonuclease P. *RNA* 4:937-947.

Ortoleva-Donnelly, L., Szewczak, A.A., Gutell, R.R., and Strobel, S.A. 1998. The chemical basis of adenosine conservation throughout the *Tetrahymena* ribozyme. *RNA* 4:498-519.

Oyelere, A.K. and Strobel, S.A. 2000. Biochemical detection of cytidine protonation within RNA. *J. Am. Chem. Soc.* 122:10259-10267.

Oyelere, A.K., Kardon, J.R., and Strobel, S.A. 2002. pK(a) perturbation in genomic Hepatitis Delta Virus ribozyme catalysis evidenced by nucleotide analogue interference mapping. *Biochemistry* 41:3667-3675.

Padilla, R. and Sousa, R. 1999. Efficient synthesis of nucleic acids heavily modified with noncanonical ribose 2′-groups using a mutant T7 RNA polymerase (RNAP). *Nucl. Acids Res.* 27:1561-1563.

Rife, J.P., Cheng, C.S., Moore, P.B., and Strobel, S.A. 1998. *N*-2-Methylguanosine is isoenergetic with guanosine in RNA duplexes and GNRA tetraloops. *Nucl. Acids Res.* 26:3640-3644.

Ryder, S.P. and Strobel, S.A. 1999a. Nucleotide analog interference mapping. *Methods* 18:38-50.

Ryder, S.P. and Strobel, S.A. 1999b. Nucleotide analog interference mapping of the hairpin ribozyme: Implications for secondary and tertiary structure formation. *J. Mol. Biol.* 291:295-311.

Ryder, S.P., Ortoleva-Donnelly, L., Kosek, A.B., and Strobel, S.A. 2000. Chemical probing of RNA by nucleotide analog interference mapping. *Methods Enzymol.* 317:92-109.

Ryder, S.P., Oyelere, A.K., Padilla, J.L., Klostermeier, D., Millar, D.P., and Strobel, S.A. 2001. Investigation of adenosine base ionization in the hairpin ribozyme by nucleotide analog interference mapping. *RNA* 7:1454-1463.

Schwans, J.P., Cortez, C.N., Olvera, J.M., and Piccirilli, J.A. 2003. 2′-Mercaptonucleotide interference reveals regions of close packing within folded RNA molecules. *J. Am. Chem. Soc.* 125:10012-10018.

Sood, V.D., Beattie, T.L., and Collins, R.A. 1998. Identification of phosphate groups involved in metal binding and tertiary interactions in the core of the *Neurospora VS* ribozyme. *J. Mol. Biol.* 282:741-750.

Soukup, J.K., Minakawa, N., Matsuda, A., and Strobel, S.A. 2002. Identification of A-minor tertiary interactions within a bacterial group I intron active site by 3-deazaadenosine interference mapping. *Biochemistry* 41:10426-10438.

Sousa, R. and Padilla, R. 1995. A mutant T7 RNA polymerase as a DNA polymerase. *EMBO J.* 14:4609-4621.

Strobel, S.A. and Ortoleva-Donnelly, L. 1999. A hydrogen-bonding triad stabilizes the chemical transition state of a group I ribozyme. *Chem. Biol.* 6:153-165.

Strobel, S.A. and Shetty, K. 1997. Defining the chemical groups essential for *Tetrahymena* group I intron function by nucleotide analog interference mapping. *Proc. Natl. Acad. Sci. U.S.A.* 94:2903-2908.

Strobel, S.A., Ortoleva-Donnelly, L., Ryder, S.P., Cate, J.H., and Moncoeur, E. 1998. Complementary sets of noncanonical base pairs mediate RNA helix packing in the group I intron active site. *Nat. Struct. Biol.* 5:60-66.

Szewczak, A.A., Ortoleva-Donnelly, L., Ryder, S.P., Moncoeur, E., and Strobel, S.A. 1998. A minor groove RNA triple helix within the catalytic core of a group I intron. *Nat. Struct. Biol.* 5:1037-1042.

Szewczak, A.A., Ortoleva-Donnelly, L., Zivarts, M.V., Oyelere, A.K., Kazantsev, A.V., and Strobel, S.A. 1999. An important base triple anchors the substrate helix recognition surface within the *Tetrahymena* ribozyme active site. *Proc. Natl. Acad. Sci. U.S.A.* 96:11183-11188.

Szewczak, L.B., DeGregorio, S.J., Strobel, S.A., and Steitz, J.A. 2002. Exclusive interaction of the 15.5 kD protein with the terminal box C/D motif of a methylation guide snoRNP. *Chem. Biol.* 9:1095-1107.

Wong, I. and Lohman, T.M. 1993. A double-filter method for nitrocellulose-filter binding: Application to protein-nucleic acid interactions. *Proc. Natl. Acad. Sci. U.S.A.* 90:5428-5432.

Contributed by Jesse C. Cochrane and
   Scott A. Strobel
Yale University
New Haven, Connecticut

Chemical and Enzymatic Probes for Nucleic Aicd Structure

6.9.21