ABSTRACT

JU,SHOW-YI. Verification of the Secondary Structure of RNA Diels – Alderases DA22/96. (Under the direction of Dr. Stefan Franzen).

The Diels-Alder reaction which generates two carbon-carbon bonds during a [4 + 2] cycloaddition is biochemically and pharmacologically significant. In order to expand and explore the utility and the application of this reaction to the assembly of complex natural products, an *in vitro* SELEX approach was used to obtain RNA biocatalysts of this transformation. A DA22/96 was prepared by the sequence truncation from the parent DA22. An essential requirement for catalytic activity was the incorporation of 5-(4- pyridylmethyl) carboxamide-uridine in place of uridine during the transcription of the DA22/96. My research was interested in finding the secondary structure of the DA22/96. This particular ribozyme exhibited catalytic activity similar to an enzyme. Mutual information and covariation analysis of 50 different catalytic sequences were used to identify the potential secondary structural elements. The comparisons between the MFold program derived secondary structures and mutual information analysis plots suggests that there were several conserved structural elements in all active DA22/96. RNase mapping method was used to validate the predicted secondary structure of DA22/96. The mapping data is largely consistent with the predicted secondary structure of DA22/96. Diels-Alderase activity assay experiment proved that DA22/96 had true enzyme activity. In addition, results showed that the modified DA22/96 had the better turn over rate when copper was presented.

VERIFICATION OF THE SECONDARY STRUCTURE OF RNA DIELS – ALDERASES DA22/96

by
SHOW-YI JU

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Science

CHEMISTRY

Raleigh, North Carolina

2006

APPROVED BY:

Dr. Steven A. Lommel	Dr. Stuart E Maxwell

Dr. Stefan Franzen
Chair of Advisory Committee

Biography

Born in Taiwan, I grew up in a family of seven boys and three girls. I am the youngest daughter in the family. My father had a special education degree in Japan and he always believed and stressed to us that achieving a higher education can help people enhance their lives.

After getting married, my husband and I made the decision to get a higher education in the United States, not only for us but also for our son. As a pharmacist by training and education, I then decided to put my education plans on hold to raise a family, but resumed my pursuit once my children were in high school. I took courses related to pharmacy from the universities and community colleges. After I received my pharmacy technician diploma, I felt more confident about myself.

In 2003 I had an opportunity to enter the master degree program to study chemistry at NCSU. During this time, I was able to learn more about all the technological advances that are pervasive in our everyday lives. The advanced scientific knowledge that I achieved carried over into my every day life so that my eyes were opened to new ideas about the world.

At NCSU, I genuinely enjoyed working with my peers and professors, despite my limited English vocabulary. Even now, as a grandmother, I believe that I can still give back from what I have gained from school so that others may have a better life.

I am grateful to NCSU which provided such a tremendous learning and working environment, friendly professors, advanced equipment, and many helpful people to make my dreams come true.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

1 Introduction

1.1 Motivation for studying an artificial ribozyme

The Diels-Alder reaction is one of the most important carbon-carbon bond-formation available to an organic chemist[1]. The reaction creates two carbon-carbon bonds and up to four new stereocenters[2]. RNA Diels-Alderase (DA22) ribozymes are RNA molecules that have catalytic activity. They fold into complex tertiary structures similar to a highly structured protein and catalyze a broad range of chemical reactions. RNA folding is a hierarchical process-first forming secondary structure element before folding into a three-dimensional complex that has binding pockets for ligands or substrates which provides catalytic centers for chemical reactions. Some of ribozymes occur naturally and can catalyze the cleavage or formation of phosphodiester bonds in the ribosome. The absence of enzymes capable of carrying out the Diels-Alderase reaction in the environment raises the compelling question of how a cycloaddition reaction might occur in a catalyst. The DA22/96 is a sequence truncated from DA22. The DA22/96 contains the same highly conserved consensus region from the parent DA22 but has a higher k_{cat}[3]. We know that the sequence in ribozyme DA22 after passing position 96 as nonessential for catalytic activity. It is clear that there is some form of limited interaction between the structure elements at 3'-end of these RNA catalysts and the highly conserved catalytic core. It is worthwhile to focus on DA22/96 secondary structure determination and establish the relationship between the structure and the functionality of this special ribozyme.

1.2 Diel-Alderases and Modified Uridine

The artificial ribozymes that have been selected from large pools of RNA sequences to catalyze reactions such as aminoacylation, N-alkylation, and Diels-Alder reactions have been published[4,5]. The artificial ribozyme DA22/96 can catalyze a Diels-Alder reaction. According to the published accounts, the ability of the pyridyl group to form a unique metal binding site in the RNA molecule affects the Diels Alderase enzyme activity. Although a number of metal ions were investigated in the selection the function of the DA22/96 ribozyme relies on the presence of divalent copper. DA22/96 contains 5-(4-pyridylmethyl) carboxamide-uridine modified nucleotides and its activity depends on the presence of cupric ion. The pyridyl modification provides a hydrophobic effect, a dipolar interaction, and additional hydrogen bonding. It also provides alternative σ-donor ligands for metal ion binding. The 5- position of pyridyl group can be very crucial to DA22/96 Diels-Alderase activity. When DA22/96 is binding to the substrates, it forms a pyridyl-cupric complex. Therefore the active site of DA22/96 has the potential to act as a Lewis acid catalyst[6].

1.3 How the in vitro Selection Works

The *in vitro* method selected was developed in the laboratories of Gold, Szostak and Joyce in the early 1990s[7]. It is based on the repeated selection of active species from nucleic acid libraries via the separation of un-reactive molecules and enzymatic amplification of the enriched library. The *in vitro* selection can be seen in figure 1.

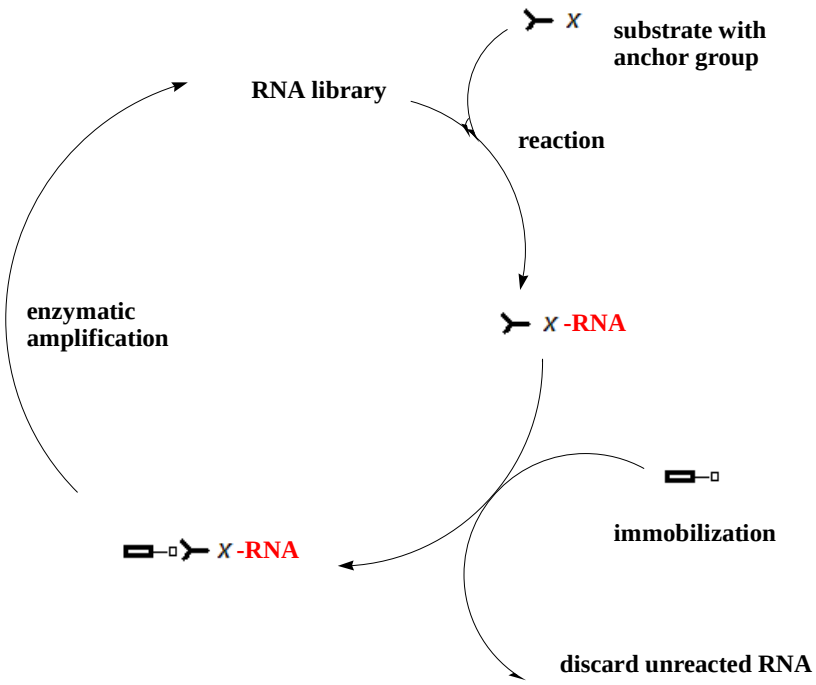


Figure 1 The principle of the SELEX method

The systematic **e**volution of **l**igands by **ex**ponential enrichment or SELEX method is based on two key steps. The first step is to select molecules with specific binding properties by immobilizing target molecules of interest in a solid matrix. The combinatorial nucleic acid library is then applied to the affinity matrix in a suitable buffer. Unbound nucleic acid molecules are washed away and the binding molecules are retained. The eluted molecules undergo enzymatic amplification, generating several thousand copies of each selected nucleic acid molecule. The end result is a library in which molecules with high affinity for the target are enriched. The process must be repeated several times until molecules with the desired properties dominate the enriched library. A sequencing process using catalytic function then identifies the target member of this library. The method with the highest rate of success for identifying nucleic acid catalysts is the

reaction with a substrate X, accelerated to isolate the desired member. The substrate X carries an anchor group so that the RNA molecules (which react with the substrate) acquire the anchor group and can subsequently be isolated by affinity chromatography on a suitably deviated matrix. RNA that are un-reacted do not bind to the matrix and can be removed through washing. This cycle is repeated until active molecules again dominate the library[8,9]. DA22 uses the SELEX method to select all the active molecules and is selected from 10^{14} RNA sequences that contain the modification 5-(4-pyridylmethyl) carboxamide-uridine[10,11].

RNA catalyst selection is achieved through a covalent attachment of diene and RNA via a polyethylene glycol (PEG) linker and a DNA 10-mer by bridge-medicated ligation of the dienophile on to the 5'-end of the RNA.

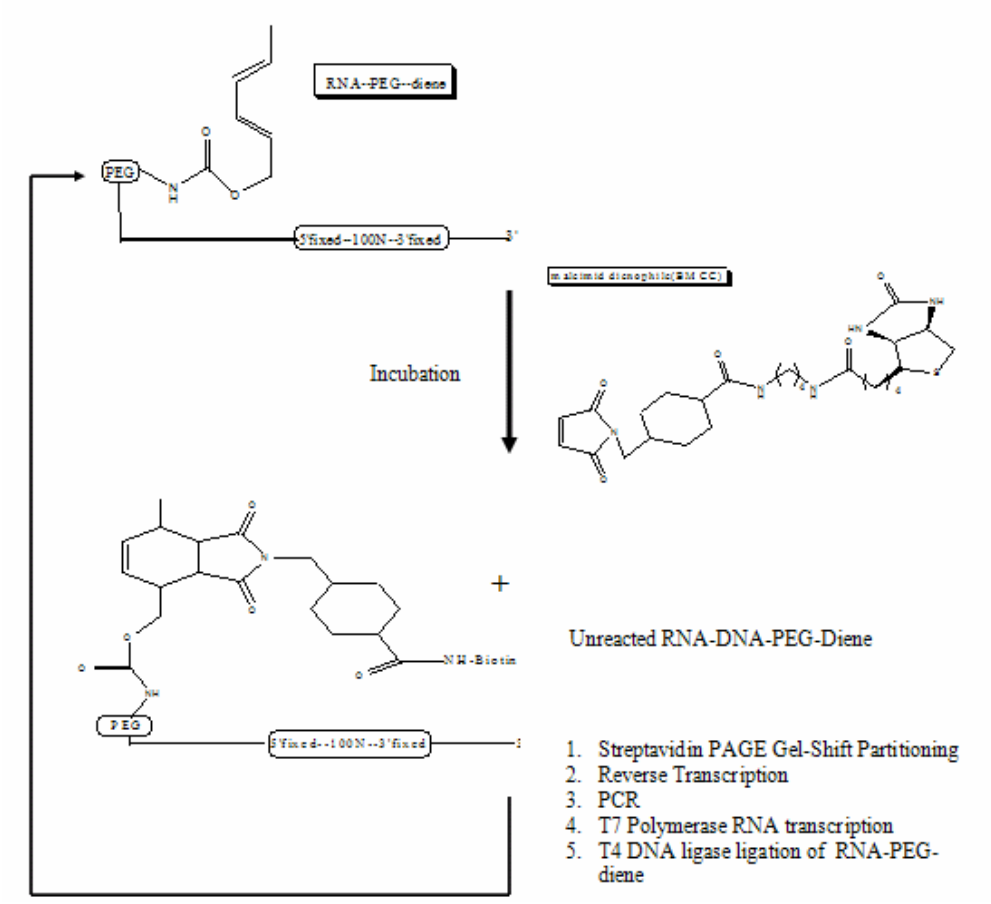


Figure 2 The Diels-Alder reaction between the acyclic diene conjugated to RNA through long PEG linker and maleimide dienophile.

1.4 Applications of RNA:

The expansion of *in vitro* selection techniques has broadened the scope of RNA catalysis for organic reactions and is a driving force behind the increase in the research involving the use of RNA for different purposes [12] . The understanding of the relationship among the structures of RNA, catalytic mechanism of ribozyme and functionality of RNA expands our understanding of biocatalysts[1]. In order to understand the structure-function relationship in RNA, a sensitive assay system is developed and specific ligands are isolated. The combination of high specificity in bimolecular catalysis coupled with the diversity

of organic and organometallic chemistry is another promising area for RNA applications. Synthetic hybrid catalysts that contain nucleic acid and an organometallic complex can generate various reactions[3]. Developing different *in vitro* evolution schemes can generate the most selective or efficient catalysts[13].

1.5 Goal

The main objective of my project is to determine whether the ribozyme DA22/96 has the enzymatic activity similar to a true enzyme while in a solution that can accelerate the reaction between the two free reactants under the copper present. The second goal is to identify the secondary structure of ribozyme that has a catalytic center of consisting of molecular RNA.

2 Methods and Experiments

DA22/96 is a sequence truncation of the full length DA22 construct. The DA22 molecules were constructed a contiguous 100-nucleotide randomized region, flanked by constant sequence segments to allow for amplification and other enzymatic process[14]. We developed a Diels-Alderases ribozyme activity assay for the DA22/96 from 49-mer oligoribonucleotides[1]. The crystal structure of unmodified 49-mers Diels-Alder ribozyme is reported[15]. By using the Mfold program, Bioedit software, and RNase mapping we find that the DA22/96 has a similar secondary structure of 49-mer ribozyme and its uridine contains the 4-pyridylmetnyl-carboxamide at 5 positions.

2.1 DA22/96 Diels-Alderases Ribozyme Activity Assay

The parent DA22 was selected by tethering PEG-linked RNA cognate created during the selection. The purpose of a PEG-linker is to allow the diene survey the entire surface of RNA[16]. It increased the efficiency of the reaction[17]. A synthetic 49-mer oligoribonucleotides contains a small common structure motif. Figure 3 shows that 49-mer RNA reports as a true catalyst without using the PEG-Linker for Diels-Alder reaction[18]. The DA22/96 contains the similar secondary structure motif. The activity assay for DA22/96 focuses on small organic substrates (9-Anthranenymethyl) hexaethylene and maleimidocaproic acid which reacts with DA22/96 ribozyme in the aqueous buffer.

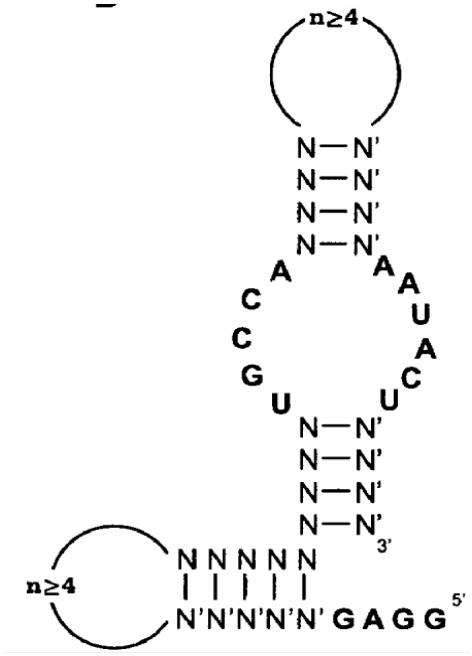


Figure 3 Minimum secondary structure motifs responsible for catalysis of Diels-Alder reaction.

Because it required large amount of RNA DA22/96 and small RNA J49 for enzyme assay experiment, we synthesized with small amount of RNA DA22/96 template and small RNA J49 template from IDT. We used PCR technique to amplify both templates and T7 transcription reaction to convert templates to transcripts.

2.1.1 Polymerase Chain Reaction and DNA template

2.1.1.1 DA22/96 PCR

Polymerase chain reaction (PCR) is a common method employed to create multiple copies of specific fragments of DNA. The PCR procedure rapidly amplifies a single DNA molecule into many billions of molecules. There are three major steps in PCR which are repeated for 30 or 40 cycles. The first step is Denaturation at temperature 94^0C. The second step is Annealing at 54^0C. The

third step is extension at 72^0C. By using those steps it could create millions of copies of DNA molecules in an hour, like a copy machine. The materials that are necessary for the PCR procedure are:

- DA22/96 T7 96mer Template:

  5'CCAGGCACGCGGGAGACAAGAATAAACGCTCAATTGA
  GTGTCTTGTAGAGCGGCGTGAAGAGCTTACGACTCTTCG
  TGTCCGTTTCTAACGCGTGC3'

- DA22/96 T7 down primer sequence:

  5'GCACGCGTTAGAAACG3'

- DA22/96 T7 up primer sequence:

  5'TAATACGACTCACTATACCAGGCACGCGGGAG3'

Before performing the PCR procedure, the three ingredients listed above are mixed with Taq.DNA polymerase in a 1.5 mL centrifuge tube.

Table 1 PCR Reaction for DA22/96 ribozyme

100 µL PCR reaction	**Final Concentration**
dH_2O	x µL
10X PCR buffer	10 µL
10mM dNTP's	1mM
Up-primer	2 µg/ µL
Down-primer	2 µg/ µL
Template	100-150 µg
Taq.DNA polymerase	5U/ µL

The following steps show the temperature and time during the PCR programming of DA22/96 ribozyme.

	Temperature	**Time**
1.	95^0C	2 min

2. 95^0C 1 min

3. 50^0C 1 min

4. 72^0C 30sec

5. Repeat step 2, 3, 4 for 31 times

6. 70^0C 10 min

7. 4^0C final hold

Following the completion of the PCR program, pure and small templates of J49 are used for the phenol/chloroform extraction and EtOH purification.

2.1.1.2 PCR Amplification of J49

The materials that are needed for PCR amplification of J49 are:

- Small 49mer T7 template:

  5'-GAGCTCGCTTCGGCGAGGCCGTGCCAGCTCTTCGGAGCAATACTCGGC-3'

- Small 49mer T7 up-primer

  5'-TAATACGACTCACTATAGGAGCTCGCTTCGGCGAG-3'

- Small 49mer T7 down-primer 5'-GCCGAGTATTGCTCCGAAGA-3'

Before the PCR procedure, the above three ingredients are mixed with Taq DNA polymerase in a 1.5 mL centrifuge tube.

Table 2 PCR reaction for small 49mer ribozyme

100 µL PCR reaction	**Final Concentration**
dH_2O	X µL
10X PCR buffer	10 µL
10mM dNTP's	1mM
Up-primer	2 µg/ µL
Down-primer	2 µg/ µL
Template	100-150 µg
Taq DNA polymerase	5U/ µL

The following steps show the temperature and time during the PCR program for small J49 ribozyme.

	Temperature	**Time**
1.	95^0C	1 min
2.	95^0C	1 min
3.	57^0C	1 min
4.	72^0C	30sec
5.	Repeat step 2, 3, 4 for 31 times	
6.	72^0C	10 min
7.	4^0C	final hold

Following the completion of the PCR program, pure and small templates of J49 are used for the phenol/chloroform extraction and EtOH purification.

Transcription Reaction for RNAs

The purpose of the transcription reaction is to create a RNA copy of the DNA template. It includes the purification processing after the transcription reaction. *E. Coli* T7 RNA polymerase was used to make RNA copies of the DNA template described above.

2.1.1.3 Unmodified DA22/96 Transcription

The following reaction components are required to perform the standard AmpliScribe T7 transcription reaction at room temperature.

Table 3 The DA22/96 ribozyme T7 transcription

Reaction Components	Final Concentration
RNase-Free water	x µL
1 µL linearize template DNA	50ng/µL
2 µL 10X Reaction Buffer	1X
1.8 µL 100mM ATP	9mM
1.8 µL 100mM CTP	9mM
1.8 µL 100mM GTP	9mM
1.8 µL 100mM UTP	9mM
2 µL 100mM DTT	10mM
µL AmpliScribe T7 –Flash Enzyme solution	

The total volume for the mixed reaction components is 20 µL. The combination of the above reaction components is incubated at 37^0C for 10 hours. GEL purification process is used to obtain pure RNA. The procedure is as follows:

1. The 6% PAGE gel is used to separate RNA. The 12mLof 20% acrylamid solution (in stock) is mixed with 1x TBE plus 7 M Urea solutions until the volume is 39.0mL. The solution is mixed thoroughly before being pulled into the gel machine along with 320 µL of APS (10%) and 40 µL of TMED.
2. Load 100 µL of RNA (containing the loading buffer) into the gel machine for 2 ½ hours under 40 mV heated condition. The gel is subsequently taken out of the machine,
3. Stain the gel with methylene blue water solution for 2 minutes.

4. Cut off the RNA band. Go through the 3mL .45 μm syringe and filter into a new 1.5mL tube.
5. Collect the filtrate (contains DA22/96) and perform the ETOH purification. Calculate the concentration of DA22/96.

2.1.1.4 Modified DA22/96 Transcription: (5-(4-pyridylmethyl) carboxamide-uridine DA22/96 Transcription)

All the materials and procedures are the same as above for unmodified DA22/96 Transcription. The only difference is that instead of using the previous mentioned uridine, 5-(4-pyridylmethyl) carboxamide-uridine is used for T7 transcription Reaction.

2.1.1.5 Small J49 RNA Transcription

Dr. Andres Jaschke group has proved that 49-mer ribozyme is Diels-Alderase [19]. This ribozyme is used as the positive control for the modified DA22/96 ribozyme enzyme activity assay. The following reaction components are combined at room temperature to implement the Standard AmpliScribe T7 Transcription Reaction process.

Table 4 Small 49mer (J49RNA) T7 transcription

Reaction Components	Final Concentration
RNase-Free water	x μL
1 μL linearize template DNA	50ng/μL
2 μL 10X Reaction Buffer	1X
1.8 μL 100mM ATP	9mM
1.8 μL 100mM CTP	9mM
1.8 μL 100mM GTP	9mM
1.8 μL 100mM UTP	9mM
2 μL 100mM DTT	10mM
μL AmpliScribe T7 –Flash Enzyme solution	

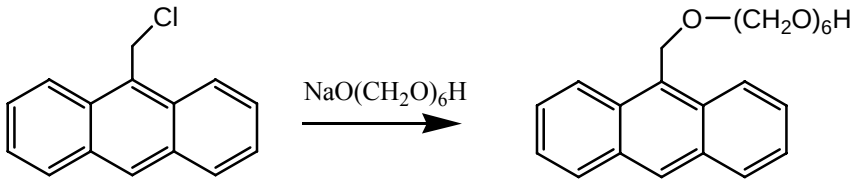
The total volume for the mixed reaction components is 20 μL. The combined reaction components listed above is incubated at 37^0C for 10 hours. Gel purification procedure is needed to obtain pure 49-mer RNA the. The procedure is identical to the preceding DA22/96 procedure above.

2.1.2 Catalysts of Diels-Alderases Ribozyme

The isolation of Diels-Alderase ribozyme (49-mer RNA) was achieved through the use of a combinatorial RNA library. The ribozyme accelerates up to 20-000-fold, carbon-carbon bond formation between anthracene, which is covalently tethered to the ribozyme, and a biotinylated maleimide[20]. The reaction condition for the modified DA22/96 ribozyme mimics the 49-mer RNA in order to prove that the modified DA22/96 is a Diels-Alderase in solution. The assay involves two organic substrate molecules, anthracene hexaethylene glycol and maleimidocaproic acid.

2.1.2.1 Material for Enzymatic Reaction

1. Synthesis substrate I: (9-Anthracenymethyl) hexaethylene glycol



According to published literature[21], 50ml Schlenk, 425mg (1.5mmol) hexaethylene glycol (HEG) was introduced in an Argon atmosphere and dissolved in 5ml anhydrous pyridine. The solution is subsequently evaporated to remove any trace of water. The procedure was then repeated once more. Next, 342mg of

(1.5mmol) 9-chloromethylanthracene and anhydrous DMF (3ml) were added. Finally, 40mg (1.6mmol) NaH was added and the mixture was stirred at room temperature overnight. Any volatile substance was evaporated in a vacuum and the residue treated with flash chromatography (SiO_2). The solution is then eluted with ethyl acetate, followed by ethyl acetate / methanol (10:1). The product obtained was amber oil. Yield: 630mg, 94% (see Figure 4 and 5) compared with published reports.

^{1}H NMR (δ, ppm, $CDCl_3$): 8.37 (d, 2H, J = 5 Hz,), 8.34 (s, 1H), 7.93 (d, 2H, J = 8Hz), 7.37-7.49 (m, 4H), 5.48 (s, 2H), 3.72 (t, J = 2Hz), 3.67-3.45 (m, 22H).

^{13}C NMR (δ, ppm, $CDCl_3$): 131.6, 131.2, 129.1, 128.5, 126.3, 125.1, 124.6, 72.8, 72.7, 70.7, 70.6, 70.4, 69.7, 65.5, 61.8.

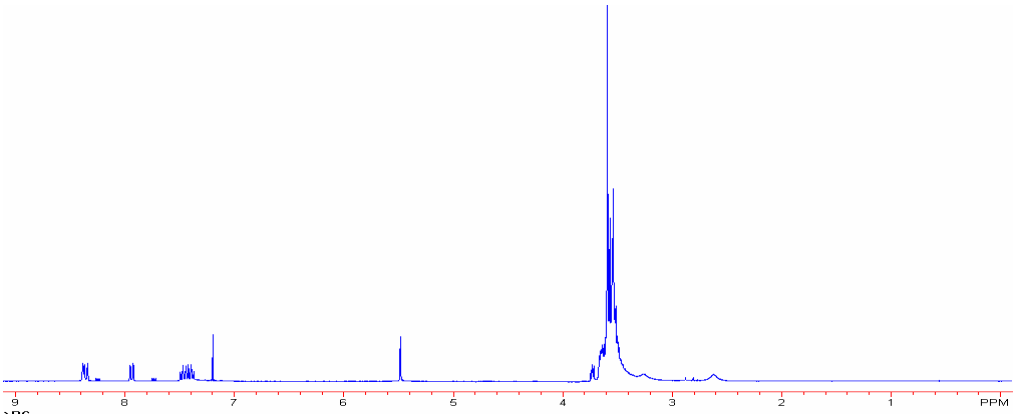


Figure 4 ^{1}H NMR of 9-Anthracenymethyl hexaethylene glycol

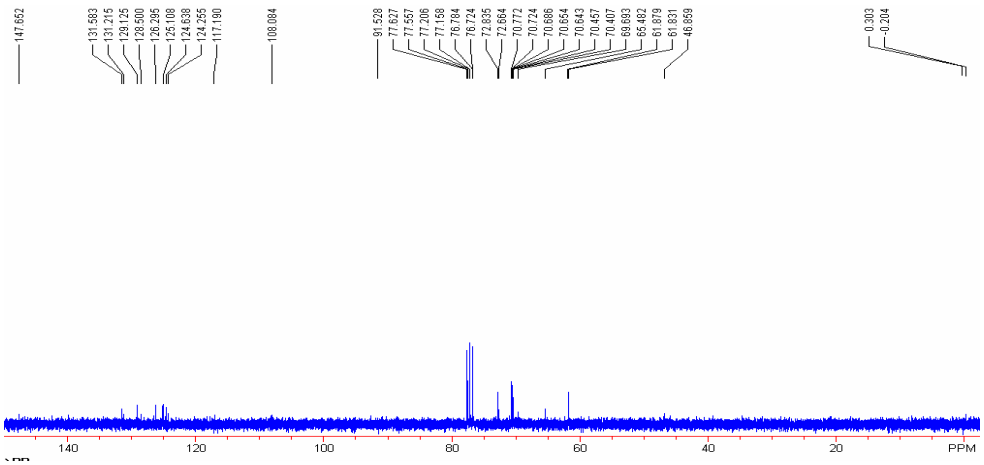


Figure 5 [13] CNMR of 9-Anthracenymethyl hexaethylene glycol

2. Substrate II:

   Maleimidocaproic acid is ordered from Sigma.

3. Buffer solution:

   The buffer solution contains 300mM NaCl, 80mM MgCl2, and 30mM Tris-HCl (pH7.46), 10% ethanol. [Tris equal tris (hydroxymethyl) aminomethane].

2.1.2.2 The modified DA22/96 enzymatic assay reaction

The reaction can be confirmed by using UV spectroscopy due to the fact that anthracene is absorbed at the 365 nm UV/VIS range. The following information describes the procedure for the assay reaction. The materials needed are: 100μM (9-anthracenylmethyl) hexaethylene glycol, 500μM maleimidocaproic acid, ribozyme (49-mer or 96mer) 70μM, and 7μL buffer solution. The total volume is 70μL reaction mixture. The reaction mixture is then transferred into a 70μL cuvette with a 1 cm light path (using a DU-64 spectrophotometer from BECKMAN). The concentration of the anthracene absorbance is record at 365nm with different time periods of 0, 1, 2, 5, 10, 20, 30,

60, 90,120 minutes. The same reaction procedure is repeated four more times on unmodified 96-mer ribozyme, modified 96-mer ribozyme, modified 96-mer with copper (The concentration of copper is 15μM in 70μL reaction volume if used), and small J49 ribozyme.

2.2 Secondary Structure of RA22/DA96 Ribozyme

Like DNA, most biologically active RNAs are extensively base paired to form double stranded helices. Structural analysis of these RNAs has revealed that they are not, "single-stranded" but rather highly structured. Unlike DNA, this structure is not just limited to long double-stranded helices but instead are collections of short helices packed together into structures akin to proteins. In this fashion, RNAs can achieve chemical catalysis, like enzymes.

2.2.1 Method

In RNA, secondary structures hold the key first step leading to the understanding of the functions of RNA molecules. RNA is generally single stranded with a double helix backbone, which is defined as the A-form. The factor determining the secondary structure of RNA is the stability of the base pair stacking. Since the computer is improving rapidly and many types of software are also available, two programs, Mfold and Bioedit are very good tools to predict the secondary structure of RNA. Laboratory work can then be developed to collect the physical data to confirm the predicted secondary structure of RNA. There are two approaches used for this purpose.

2.2.1.1 Thermodynamic Approaches

The most common method of RNA structure predictions is based on calculation the free energy estimates for different folds. The fold with the minimum free energy is the projected structure that is desired. These free energy values are the combination of energy values that calculate for each pair of adjacent base pairs, plus loop or bulge energies. The energy values in turn are derived from melting point determination with synthetic oligoribonucleotides. Thermodynamic approaches can find a single RNA sequence that has an optimal secondary structure with global minimal free energy. This approach has proven quite successful for many short sequences of RNA[22].

2.2.1.2 Comparative Sequence Analysis Approaches

A second approach for secondary structure prediction is searching for conserved stem regions in related sequences. Conserved stem regions mean that the primary sequence may vary but stem structure is retained by way of covariant mutations. Comparative approaches are more reliable than thermodynamic approaches when we have access to a number of aligned homologous sequences[23].

Each approach has its own advantages. A higher level of accuracy can be obtained if both approaches are used together.

2.2.1.3 Final structure of predicted secondary structure

The final predicated secondary structure should always be verified using an experimental probe of the RNA structure. These experimental methods

include RNA mapping technique, NMR, and X-ray crystallography. In order to identify DA22/96 ribozyme secondary structure, we used chemical enzymatic mapping experiments. Those experiments used RNase to cut RNA so that DA22/96 fragments can be identified by using 5'-[^{32}P] labeled RNA during the polyacrylamide gel electrophoresis separation.

2.2.2 Experiment

2.2.2.1 RNA Folding by MFOLD algorithm

MFold is a program used to predict RNA secondary structure[24]. The MFold program normally will generate several different possible secondary structures from a single input RNA sequence[25]. The following three predicted secondary structures have the lowest free energy from the Mfold output for the input DA22/96 primary sequence. These secondary structure predictions form the basis for testing by RNase mapping methods.

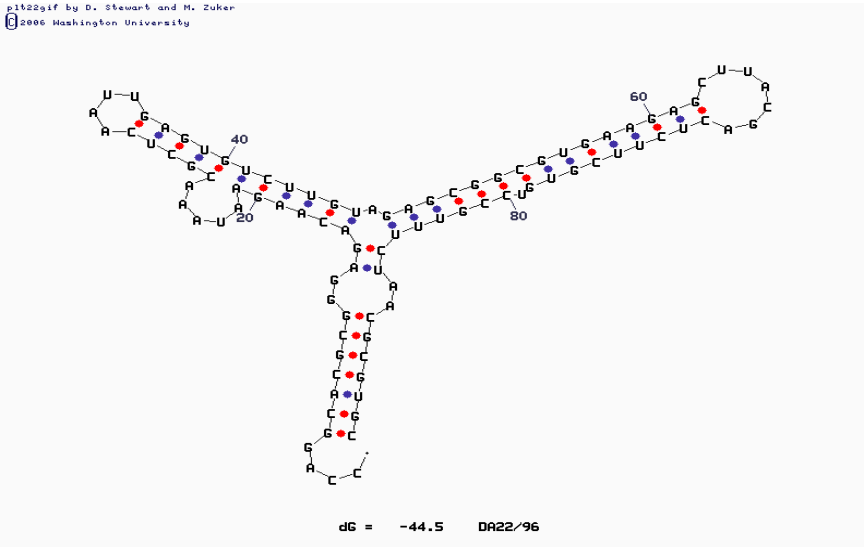


Figure 6 Mfold -44.5 Da22/96

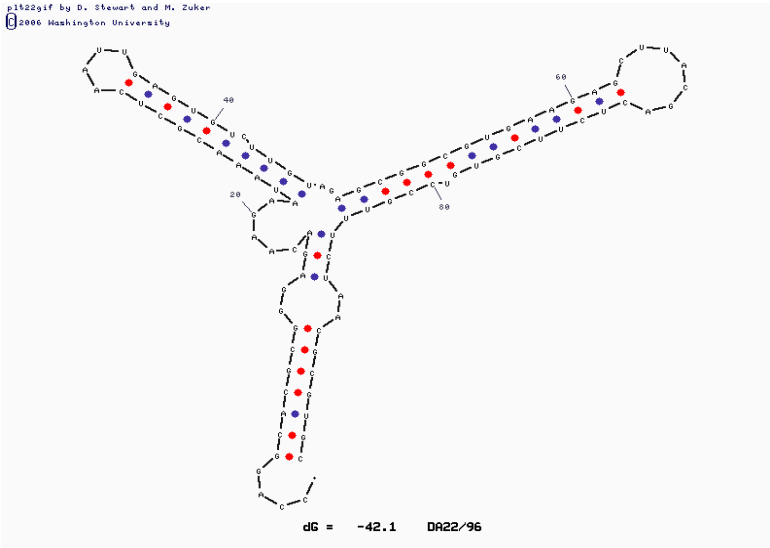


Figure 7 Mfold -42.1 DA22/96

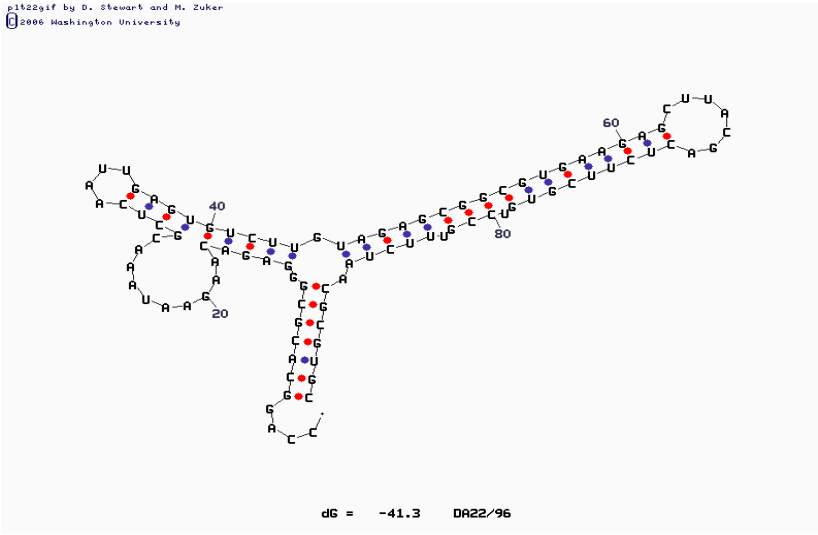


Figure 8 Mfold -41.3 DA22/96

Under a natural situation, the correct secondary structure of RNA should have lower free energy. One of these three structures could be the right secondary structure of DA22/96 ribozyme.

2.2.2.2 RNA Folding by Comparative Sequence analysis

In general molecular structure is often correlated with function such that there is a tendency to conserve protein and RNA secondary structure even when the primary amino or nucleic acie sequence is altered. RNA is composed hairpins, pseudoknots and other characteristic secondary structure elements. Within these the stem (A-form double stranded helical) regions are base-paired. When one base of a pair changes, we usually find that its partner also changes so as to conserve that base pair. This phenomenon is called a *compensatory base change*.

Suppose that we have an alignment of a group of homologous RNAs and that we are certain that the alignment DA22/96 RNA and its nucleotides are the same length. Formally, we are given a multiple alignment of **m** RNA sequences:

$$\begin{aligned} R_1 &= r_1(1), r_1(2), r_1(3), \ldots, r_1(n), \\ R_2 &= r_2(1), r_2(2), r_2(3), \ldots, r_2(n), \\ R_3 &= r_3(1), r_3(2), r_3(3), \ldots, r_3(n), \\ &\vdots \quad\quad \vdots \\ R_m &= r_m(1), r_m(2), r_m(3), \ldots, r_m(n). \end{aligned}$$

They are all the same length because they have already been aligned and some of the characters may be gaps (``-''). Suppose that base pairs $r_2(i) - r_2(j)$ can form in sequence **k** for all **m** sequences - meaning that there are no gaps in columns **i** or **j**. This is not necessarily evidence for a conserved base pair. What is needed is extra evidence in the form of compensatory base changes. Suppose that there are n_1 G-C, n_2 C-G, n_3 A-U and n_4 U-A base pairs between the 2 columns. Thus the minimum number of compensatory changes that must have occurred

during evolution is determined by 1 less than the number of the n_js that are not **0**. The evidence becomes more convincing when all the consecutive base pairs in a helix are conserved and have 1 or 2 compensatory base changes. If the evolutionary past can be reconstructed, then the conclusion can be made that compensatory base changes have occurred over and over again during evolution.

Let us map A, C, G and U to the numbers 1, 2, 3 and 4, respectively. We can let e_i be the sample mean of the **ith** column. Then assuming independence of the columns, we would expect the covariance between distinct columns, **i** and **j**, to be 0. The covariance is given by:

$$cov(i,j) = \sum_{k=1}^{m} (r_k(i) - e_i)(r_k(j) - e_j).$$

The method called the *mutual information content* of 2 columns. Since RNA secondary structure is often conserved in evolution, mutual information measure for identifying covarying sites in an alignment can be used to identify structural elements[25].

The 50 sequences that have similar and different characters of DA22/96 ribozyme are obtained by truncating up to position 96 from original DA22 sequences[26]. It is shown in following table.

Table 5 The 50 alignment of DA22/96 ribozyme sequences

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|.
                    5         15         25         35         45         55         65         75         85         95
DA-22           GGGAGACAAG AAUAAACGCU CAAUUGAGUG U~CUUGUAGA GCGGCGUGAA GA~GCUUACG ACU~CUUCGU GUCCGUUUCU AACGCGUGCC CAUGAA
1-1             GGGAGACAAG AAUAAACGCU CAAUUGAGAG U CUUGUGGA GUGGGGUCAA GA GCUUACG ACU CUUGGA GUCCGCUUCU AACGCGUGCC CAUGCA
1-17            GGGAGACAAG AAUAAACGCU CAAUUGAGAG U CUUGUGGA GUGGGGUCAA GA GC  ACG ACU CUUGGA GUCCGCUUCU AACGCGUGCC CAUGC
1-2             GGGAGACAAG AAUAAACGCU CAAUGGAGUG U CUUGUGGA CCAGCGUGUA GA GCUCA   ACU CUGCAU GUAUGGUUCU AACGCGUGAC CAUGUA
1-10            GGGAGACAAG AAUAAACGCU CAAUGGAGUG U CUUGUGGA GCGGCGUGAA GA GUUUACG ACU CUUCGU GACCGUUUCU AACGCGUGGG CUUGAA
1-7             GGGAGACAAG AAUAAACGCU CAAUUGAGUG U CUUGUGGA GUGGAAUGGA GU GCAUACA UCG CUCCGG GUCCGCUUCU AACGAGUGCC GGCCAG
1-37            GGGAGACAAG AAUAAACGCU CAAUGGAGUG U CUUGUAGA GCGGCACAAG GU CCUUAAG ACU UUGCGG AGCCUUUUCU AACGCGUCGG GUUGAU
1-12            GGGAGACAAG AAUAAACGCU CAAUUGAGAG U CUUGUGGA GUGGGGUCAA GA GCUUACG ACU CUUGGA GUCCGCUUCU AACGCGUGCC CAUGCA
1-30            GGGAGACAAG AAUAAACGCU CAAUUGAGAG U CUUGUGGA GUGGGGUCAA GA GUUUACG ACU CUUGGA GUCCGCUUCU AACGCGUGCC UAUGCA
1-26            GGGAGACAAG AAUAAACGCU CAAUUGAGAG U CUUGUGGA GUGGGGUCAA GA GCAUACG ACU CUUGGA GUCCGCUUCU AACGCGUGCC CAUGCA
1-3             GGGAGACAAG AAUAAACGCU CAAUGGAGUG U CUUGUGGA GCGGCGUGAA GA GUUGACG ACU CUUCGU GACCGUUUCU AACGCGUGCC CGUGUC
1-32            GGGAGACAAG AAUAAACGCU CAACUGAGUG U CUUGUGGA GCGGUCUGCA GA GUCUUCG ACC CUGCGU GACCGUUUCU AACGCGUGCC CGUGAA
1-14            GGGAGACAAG AAUAAACGCU CAAUCGAGUG U CUUGUGGA GUUGGAUGAA GA GCUUAGG CCU CUUCGG GUCAGCUUCU AACGCGUGCU UCUGAG
1-47            GGGAGACAAG AAUAAACGCU CAAGUGAGUG U CUUGUGGA GCGGCGUGAA GA GCUUUGG ACU CUUCGU GACCGUUUCU AACGCGUGCG AAUGCA
1-33            GGGAGACAAG AAUAAACGCU CAAACGAGUG U CUUGUGGA GCGGUCUGAU GA ACAUAAG AAU CUUCGU GACCGUUUCU AACGCGUGAC  AUGAG
1-41            GGGAGACAAG AAUAAACGCU CAAUUGAGUG U CUUGUGGA GCGGCGGGAA GA GUUUAAG ACU ACUUGC GGCCGUUUCU AACGCGUGCG CAUGAA
1-21            GGGAGACAAG AAUAAACGCU CAAUUGAGUG U CUUGUGGA GCGGCUUGAA CA CCCUACG AGG CUUCGU GGCCGUUUCU AACGCGCGCC AAUGAA
1-44            GGGAGACAAG AAUAAACGCU CAAUUGAGUG U CUUGUGGA GCGGCGUUAU GC GCAUACG GCG CUGCGU GGCCGUUUCU AACGCGUGCU AAGGCA
1-46            GGGAGACAAG AAUAAACGCU CAAUUGAGAG U CUUGUGGA GCGGAGUGAA GA GUUUAAG ACU CUUCGU GUCCGUUUCU AACGCGUGCC CUUGGC
1-25            GGGAGACAAG AAUAAACGCU CAAUCGAGUG U CUUGUGGA GCUGGUUGAG GA GCGUACG ACU CUUCGU GGCCGUUUCU AACGCGUGGG CAUGAG
1-43            GGGAGACAAG AAUAAACGCU CAAUCGAGUG U CUUGUGGA GCGGCGUGAA GA GCUUAAG ACU CUUCGU GUCCGUUUCU AACGCGUGCC CUAGGG
1-15            GGGAGACAAG AAUAAACGCU CAAUAGAGAG U CUUGUGGA GCGGCUCGAA GA GCCUAC  ACU CUUCGU GUCCGUUUCU AACGCGUGGG CCUGAA
1-29            GGGAGACAAG AAUAAACGCU CAAUAGAGUG U CUUGUGGA GCUGCGUGAA GA GUUUGCG ACU CUUCGU GGCCGUUUCU AACGCGUGGC AGUGAA
1-11            GGGAGACAAG AAUAAACGCU CAAGUGAGUG U CUUGUGGA GCGGCAAGAA GA GCUUACG GCU CUUCGG GACCGUUUCU AACGCGUGCC CGUGUG
1-31            GGGAGACAAG AAUAAACGCU CAAAUGAGUG U CUUGUGGA GCGGGGUGAA GA AGAGACG ACU CUUCGG GCCCGUUUCU AACGCGUGCG AAUGAA
1-19            GGGAGACAAG AAUAAACGCU CAAUUGGGGG U CUUGUGGA GCGGCGUGAA GA CUUUACG AUU CUUCAU GGCCGUUUCU AACGCGUGCG CAUGAA
1-16            GGGAGACAAG AAUAAACGCU CAAUGGAGUG U CUUGUGGA CCAGCUUGAC GA ACUGAUG AUU CGUCGU GUAUGGUUCU AACGCGUGCC CAUGAA
1-28            GGGAGACAAG AAUAAACGCU CAAUGGAGUG U CUUGUGGA CCAGCGUGUA GA GCU  CG ACU CUGCAU GUAUGGUUCU AACGCGUGCC AAUGAG
1-9             GGGAGACAAG AAUAAACGCU CAAUUGAGUG U CUUGUGGA GCAGCGCAAG GA GAUGAUG UCU CCUCGU GUAUGCUUCU AACGCGUGAC CAUGAG
1-40            GGGAGACAAG AAUAAACGCU CAAUUGAGUG U CUUGUGGA GUGGGGUCAA GA GGUUACG ACU CAUGGA GUCCGCUUCU AACGCGUGGC CAGGAG
1-24            GGGAGACAAG AAUAAACGCU CAAUUGAGUG U CUUGUGGA CCAGCAUGAA GA GACUACG AUU CUUCGU GUAUGGUUCU AACGCGUGCU CAUGAA
1-5             GGGAGACAAG AAUAAACGCU CAAUUGAGAG U CUUGUGGA GUGGGGUCAA GA GCUUACG ACU CUUGGA GUCCGCUUCU AACGCGUGCC CAUGCA
1-35            GGGAGACAAG AAUAAACGCU CAAUUGAGAG U CUUGUGGA GUGGAAUGAA GU GCUUAGC GCA CUUCGG GUCCGCUUCU AACGCGUGCC GUGGCU
1-49            GGGAGACAAG AAUAAACGCU CAAUUGAGUG U CUUGUGGA GCGGCGGGAC GA GCUUAUG ACG CUACGU GGCCGUUUCU AACGCGUGUC CAUGGC
1-50            GGGAGACAAG AAUAAACGCU CAAUUGAGUG G CUUGUGGA GCGGCGUGGA UA GCUUAAG ACU GUUCAU GACCGUUUCU AACGCGUGCC CAAGAG
1-27            GGGAGACAAG AAUAAACGCU CAAUUGAGUG U CUUGUGGA GCUGCUUGAA GU GCCUACG ACA CAUCGU GGCAGUUUCU AACGCGUGGC AGUGAA
1-4             GGGAGACAAG AAUAAACGCU CAAUCGAGUG U CUUGUGGA GUGGGGUCAA GA GAGUAGG ACU CUUGGA GUCCGCUUCU AACGCGUGCG CUUGAA
1-6             GGGAGACAAG AAUAAACGCU CAAUCGUGCG U CUUGUGGA GCAGCGUUAA GA GAU ACG CCU CUUCGU GUAUGCUUCU AACGCGUGCC CAUGCA
1-18            GGGAGACAAG AAUAAACGCU CAAUCGAGUG U CUUGUGGA GCAGCGUGAA GA GCUUACG ACU CUUCGU GUAUGCUUCU AACGCGUGCC CCGACC
1-34            GGGAGACAAG AAUAAACGCU CAAUCGAGUG U CUUGUGGA GUGGGGUUAA GA GC  ACG ACU CUUCGA GUCCGCUUCU AACGCGUGCU GAUGUA
1-13            GGGAGACAAG AAUAAACGCU CAAUCGAGUG U CUUGUGGA GCUGCGUGGA CA G UUGCG AAU CUGCCU UGCGGUUUCU AACGCGUGCC CUUGUA
1-20            GGGAGACAAG AAUAAACGCU CAAAUGCGUG A CUUGUGGA GCGGAGUGAG GA GCUUACG ACU CUUCGG GUCCGUUUCU AACGCGUGCG CCGGAC
1-36            GGGAGACAAG AAUAAACGCU CAAAUGAGUG U CUUGUGGA CCAGCGUGAC GG    UACG AAU CGUCGC GUAUGGUUCU AACGCGGGCC CUUCGC
1-38            GGGAGACAAG AAUAAACGCU CAAUGGAGUG U CUUGUCGA CCAGCUUGGC GA ACUGAUG AUU CGUCGU GUAUGGUUCU AACGCGUGCC CAUGAA
1-42            GGGAGACAAG AAUAAACGCU CAAUGGAGAG U CUUGUGGA CCAGCUUGAC GA ACUGAUG ACU CGUCGU GUAUGGUUCU AACGCGUGCC CAUGAA
1-42            GGGAGACAAG AAUAAACGCU CAAUGGAGUU A CUUGUGGA GUGGCGUGAA GA GGCUGCA CUU CUUCGG GACCAUUUCU AACGCGUGAC CAUGAA
1-45            GGGAGACAAG AAUAAACGCU CAAUGGAGUG U CUUGUGGA CCAGCUUGAU GA ACUGAUG AUU CGUCGU GUAUGGUUCU AACGCGUGCC CAUGAA
1-39            GGGAGACAAG AAUAAACGCU CAACUGAGCG C CUUGUGGA GCGGCCCGGA GA GGUUCCG ACC UCUCAA GACCGUUCCU AACGCGUGCG CAUGCC
1-22            GGGAGACAAG AAUAAACGCU CAAGUGAGUG U CUUGUGGA GCGGUGCGAA GAAGAGUACG ACU CUUCGU GGCCGUUUCU AACGCGUGCC CCUGUA
1-23            GGGAGACAAG AAUAAACGCU CAAAUGAGUG UUCUUGUGGA GCGGCGUGAC GA GCUCGAG ACU CUUCGU GACCGUUUCU AACGCGUGCA CAGGGA
1-8             GGGAGACAAG AAUAAACGCU CAAUGGAGUG U CUUGUGGA CCAGCUUGAC GA ACUGAUG AUUUCGUCGU GUAUGGUUCU AACGCGUGCC CAUGAA
german enz      ggagcucgcu ucggcgaggc cgug~~~~~~ ~~~~~~~~~~ ~~~~~ccagc ucuucggagc aauacucggc ~......... .......... ......
top 7 90%       GGGAGACAAG AAUAAACGCU CAAU*GAG*G U~CUUGU*GA ***G****** G*~****A** *C*~*U**** ******UUCU AACG*GUG** ******
Consensus       GGGAGACAAG AAUAAACGCU CAA**GAG*G UUCUUGUGGA ***G****** G*A******G ***UC***** G***G*UUCU AACGCGUG** ***G**
```

The mutual information plot for DA22/96 ribozyme can be obtained by using BioEdit software. Below, Figure 9 shows the matrix plot of mutual information data from the 50 alignment DA22/96 sequences.

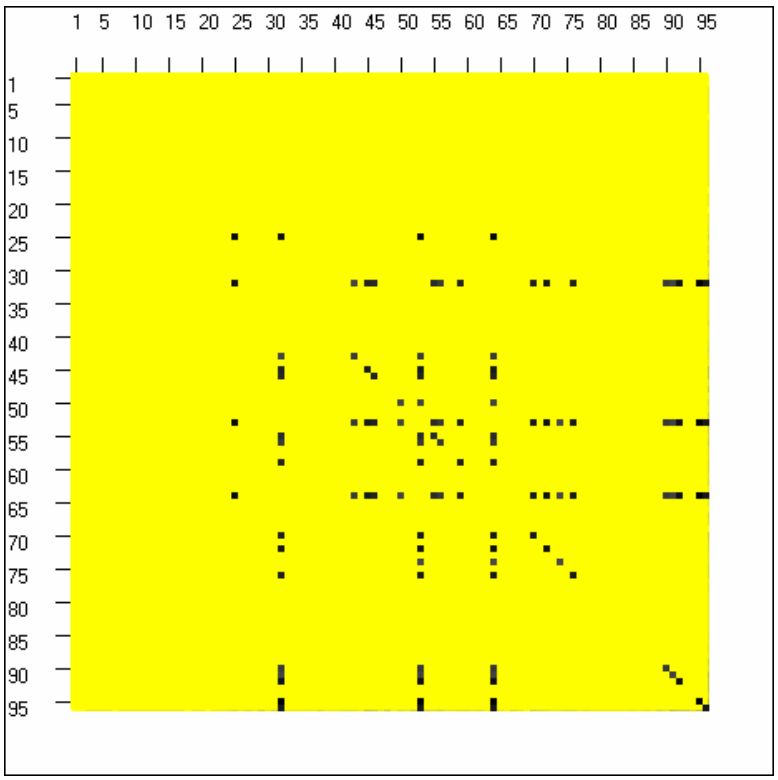


Figure 9 The mutual information plot of DA22/96 ribozyme

High information pairs are represented in the plot by the regions running perpendicular to the diagonal. Those regions represent pairs of positions with highly correlated identities (they appear to influence each other). These regions suggest that they are base-pair. Diagonal runs of high information also strongly suggest the presence of base-paired helices.

2.3 Refinement by RNase Mapping of Modified DA22/96 ribozyme.

Modified DA22/96 was transcribed *in vitro* using Ampliscribe Flash T7 Transcription Kits (Epicentre) and gel-purified by denaturing polyacrylamide gel electrophoresis (PAGE). Purified Modified DA22/96 ribozyme was radiolabeled at the 5' terminus by first digesting with calf intestine alkaline Phosphatase (CIAP,

Promega) then radio labeling using T4 polynucleotide kinase (PNK, Promega) and [γ -^{32} P] ATP (MP Biomedical). Radiolabeled Modified DA22/96 ribozyme was re-purified by denaturing PAGE. Modified DA22/96 template oligonucleotides are chemically synthesized by Integrated DNA Technologies. The following materials were used:

a. DA22/96 template sequences:

   5'CCAGGCACGCGGGAGACAAGAATAAACGCTCAATTGAGT
   GTCTTGTAGAGCGGCGTGAAGAGCTTACGACTCTTCGTGTCC
   GTTTCTAACGCGTGC3'

b. Down primer sequence:

   5'GCACGCGTTAGAAACG3'

c. DA22/96 T7 up primer sequence:

   5'TAATACGACTCACTATACCAGGCACGCGGGAG3'

5-(4-pyridylmethyl) carboxamide-uridine (replace the uridine) was used during T7 transcription Reaction. The Modified DA22/96 ribozyme can then be generated.

2.3.1 5'-end label and Gel-purify RNA

A rapid and general method for mapping RNA structure in the resolution is using structure-specific enzyme as probe[27]. The 5' terminal phosphate of RNA was removed with alkaline phosphates and replaced with [^{32}P] phosphate. This procedure is described at following figure from Ambion Company.

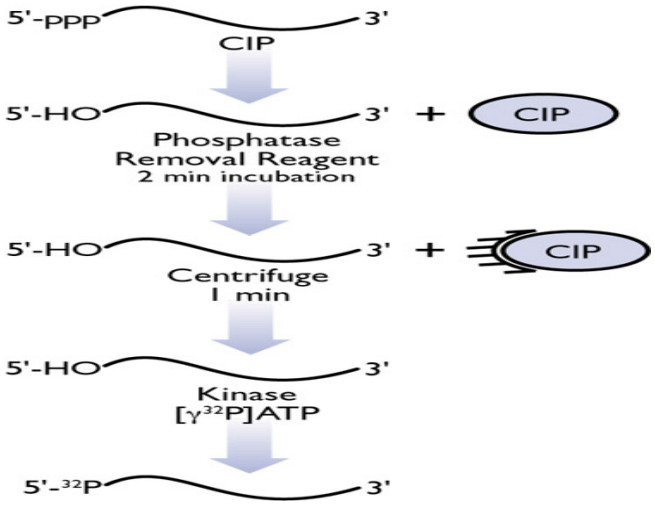


Figure 10 5'-End-Labeling Procedure

The resolution of the gel electrophoresis is very important in RNA sequencing. The gel should be able to resolve the oligonucleotide length with single-base accuracy.

2.3.1.1 Dephosphorylate

Due to the fact that DA22/96 is *in vitro* when transcribed it needs to be dephosphorylated first at the 5'-end. The dephosphorylation procedure is the following:

1. Mix 10 µg with 10 µL10x CIP Buffer (contain 0.5 M Tris-HCl pH 8.5 and 1 mM EDTA, pH 8.5)
2. Add 5 µL CIAP (Calf Intestinal Alkaline Phosphatase) (about 10U) and add ddH2O to 100 µL.
3. Incubate at 37º C for 45-60 minutes.
4. Perform Phenol/chloroform extract and RNA EtOH precipitation.

5. Resuspend in 30µL ddH2O and check the concentration with the spec photometer.

2.3.1.2 Label 5'-end with T4 Polynucleotide kinase (PNK)

The procedure for this step is the following:

1. Mix 50-100 pmol CIP-treated RNA (~1-2 µg), 5 µL 10X PNK buffer, 15µL $[\gamma]^{32}$- P ATP , 2µL PNK (20U) together and add ddH20 to 50µL volume.
2. Incubate the mixture at 37º C for 1.5 hours
3. Do Phenol/chloroform extraction for DA22/96.

2.3.1.3 Prepare and run the gel

1. Resuspend dry CIP treated RNA in 50µL GP Buffer. Boil for 3-5 minutes.
2. Pour 6% urea-PAGE gel. Gel will be whatever % acrylamide (19:1), 1X TBE, and 7M urea.
3. Pre-run gel to heat it up (about15-20 minutes). Load sample and in a separate lane load 10-20µL GP Dye. Run gel at about 40-45mA. It takes about 1.5 to 2 hours.
4. Crush up gel slice into a fine paste in 1.5mL tube. Add 400µL RNA Elution Buffer and vortex to suspend. Rock in cold room (4ºC) overnight (16-20 hours) or at room temperature for 30 minutes.
5. Recover RNA by spinning in microfuge at the highest speed for 2 minutes. Before that need to wet the filter for 100 µL. Move supernatant

to top of 3mL 0.45μm syringe and filter into a new 1.5mL tube. Add 300 μL more RNA Elution Buffer to gel and resuspend. Rock at room temperature for 15 minutes. Spin as before and filter again through same syringe to pool with previous supernatant. Push an additional 100μL fresh RNA Elution Buffer through filter to recover any remaining RNA. EtOH precipitate by adding 100% EtOH to a final of 70% for about 1 hour at -20ºC.

6. Spin down at top speed in microfuge for 30 minutes. Wash pelleted RNA with 500μL ice cold 70% EtOH by inverting. Incubate at -20º for more than 30 minutes then centrifuge again 20 minutes.
7. Pour off EtOH again and dry pellet by laying tube on its side on bench top or lyophilizing (speedvac). Dissolve pellet in 50μL H2O and check counts in scintillation counter. Label tube with date and CPMs and freeze at -20ºC for next reaction. The concentration of modified DA22/96 count 331487/ μL, unmodified DA22/96 count 1271051/ μL.

2.3.2 Nuclease and Pb^{2+}, Cu^{2+} Mapping of DA22/96 ribozyme

For the purpose of mapping the ribozyme, a total volume of 10 μL for all reactions was used. Radiolabeled DA22/96 RNA ($\geq 1 \times 10^5$ cpms) is assembled in 1x MAB (Mapping Assembly buffer) buffer (20mM HEPES, pH 7.0, 100 mM NaCl, 3 mM $MgCl_2$, 0.2 mM EDTA) containing 10 mg/mL tRNA, substrate I (anthracene) 100 μM, substrate II(maleimide acid) 500 μM. Mung bean nuclease (Epicentre) was added to a final of 5.125 units/μL and reactions incubated at 25°C

for 12 minutes. Then lead acetate was added to a final concentration 10mM for the repeated experiment under the same conditions.

For the actual mapping of the DA22/96 RNA assembly, T1 digestion ladders are generated by incubating 5'- radiolabeled DA22/96RNA with 10 mg/mL tRNA and 0.025 units/ μL RNase T1 (Ambion) in assembly buffer at 25°C. Alkaline hydrolysis ladders are generated by incubating 5'- radiolabeled sRNA with 10 mg/mL tRNA in 10 mM $NaHCO_3$, pH 10.0, 1 mM EDTA at 95°C. Null reaction (NR) samples were identically handled with the exception of the absence of added nuclease or Pb^{2+}. All reactions were halted by the addition of 10 volumes of 2% $LiClO_4$ in acetone. Pellets are collected by centrifugation, washed with acetone, and then air-dried. RNA samples were resuspended in loading buffer (1 X Tris-borate-EDTA, 4M urea, 0.25 mg/mL bromophenol blue, 0.25 mg/mL xylene cyanol, 12% (w/v sucrose), boiled for 5 minutes, and then resolved on 12% polyacrylamide sequencing gels containing Tris-borate-EDTA buffer, pH 8.3, and 7 M urea. Vacuum-dried gels were visually inspected using a phosphorimager.

3 Results and Discussions

The above experiments produced two interesting results. One involves the modified DA22/96 Diels-Alderase ribozyme activity assay and the other, a prediction of the secondary structure of modified DA22/96 RNA.

3.1 Modified DA22/96 Diels-Alderase Ribozyme Activity

From the assay experiment of DA22/96 Diels-Alderase, the modified ribozyme accelerates the Diels-alder reaction when copper is present in the solution

3.1.1 Results

The unmodified DA22/96 is the negative control while J49 is the positive control. J49 exhibited Diels Alderase activity in solution and further backed by the Jaschke group[28]. The experiment shows that J49 has a higher level of activity than the unmodified DA22/96. The modified DA22/96 preliminary data shown in Figure 10 clearly has the higher activity than unmodified DA22/96 and J49. We also observed that the uridine modified DA22/96 with copper ion has the highest activity among those ribozymes. The modified DA22/96 is the first 96 nucleotides of the DA22 ribosome[29]. Based on published accounts of the activity the DA22/96 construct has the same activity as full-length DA22.

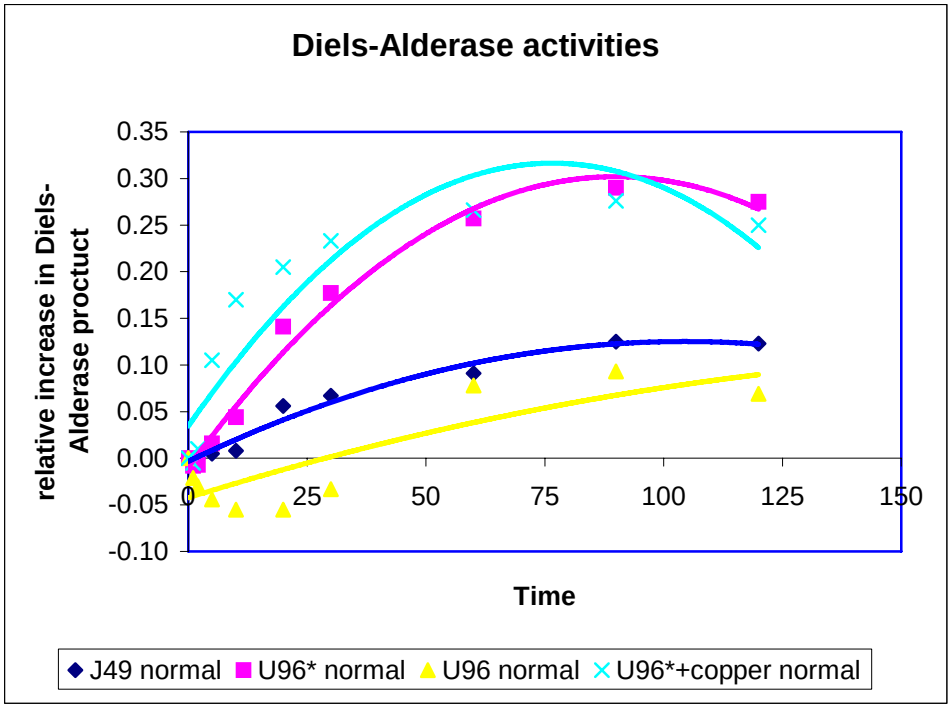


Figure 11 Diels-Alderase activity relative increase products in three ribozymes

The activity of isolate parent's DA22 was found to depend on the presence of copper ion in solution. No Diels-Alderase activity was observed in the absence of copper ion[30]. The figure 11 from the experiments showed that the DA22/96 ribozyme has same product amount with or without copper ion at time about 100 minutes. At 30 minutes, both modified DA22/96 with and without copper ion hit the maximum rate (V_{max}). In the first 10 minutes, the DA22/96 cognate with copper ion has a faster initial velocity (as shown in the earliest part the curve) than the DA22/96 without copper. In other words with or without copper didn't make significant difference.

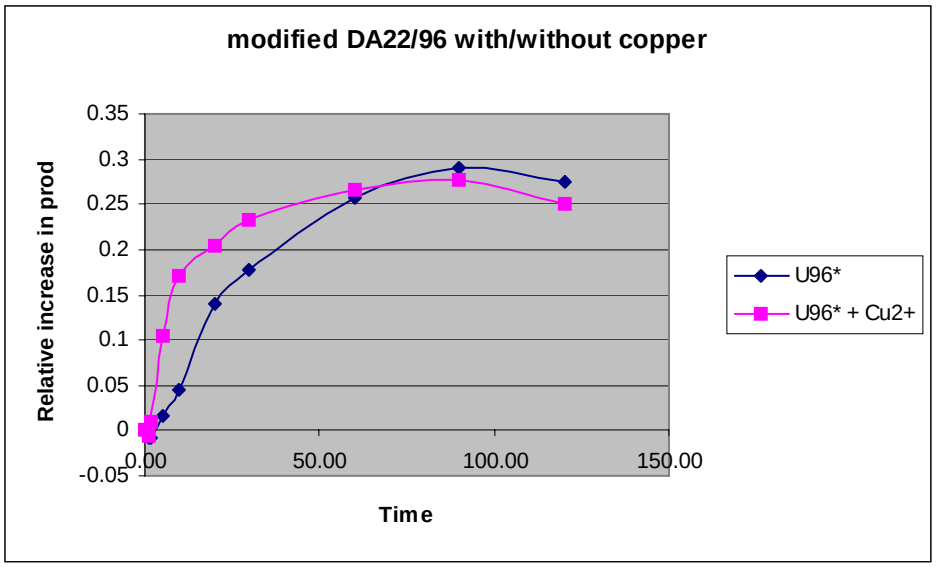


Figure 12 Comparison between amount of product of Diels-Alder reaction with copper and without copper of DA22/96 ribozyme

3.1.2 Discussion

When the SELEX method is used to isolate RNA cognates, it becomes a self-modifier and can only perform the reaction once. The assay result proves that the DA22/96 is the true catalyst just like parent's DA22[31]. All pentose in DA22/96 have ribose (β-D-ribofuranose) but only 10 pentose in parent's DA22 are deoxyribose (2-Deoxy β-D-ribofuranose).

Ribozymes are molecules that have catalytic properties. From this result now we have a simple in-solution activity assay for Diels-Alderase activity that can be easily monitored by a decrease in the absorbance of Anthracene substrate at 365 nm in any standard spectrophotometer, as reported by the Jaschke group[1]. Also we can assure that modified DA22/96 has the capacity to catalyze the Diels-Alder reaction in the solution. Before this assay we could not make such claims with a high degree of certainty. The enzyme reaction is also at a higher rate than

J49 in the absence of copper. Unmodified DA22/96 lacks functional group diversity thus limiting the type of catalytic mechanism RNA when compared to modified enzyme. It is the acid -base chemical reaction that causes the Diels-alder reaction to happen. The modified uridine for DA22/96 is the main reason for the higher rate of enzyme reaction. Graph 11 clearly shows the modified DA22/96 catalyzing the same amount of product with or without copper (see figure 11); however, with copper, its initial velocity (the earliest part of the curve) is much faster. The behavior shows that copper ion is not essential for the reaction but definitively enhances the rate of the reaction.

3.2 Prediction of Modified DA22/96 RNA Secondary Structure

The modified DA22/96 RNA that contains consensus sequence is folded using the Zucker algorithm-based program Mfold to have lowest energy[32]. The result shows the lowest energy folding model is the correct secondary final fold structure of DA22/96 ribozyme.

3.2.1 Results

3.2.1.1 Predicted Secondary structure of Modified DA22/96

Comparative sequence analysis is based on the biological paradigm that macromolecules are the product of their evolution. A higher-order structure can be derived from different primary structures. Comparative structure analysis requires an alignment of those sequences that have similar or differences in sequence collection. The better the sequence alignment is, the more meaningful information can be extracted. Mutual information is a way to measure the degree

of covariance between two columns in multiple alignments. From the mutual information plot, we know all the positions which secondary structure base pairings are involved and which correlates the best with their cognate pairing partner.

Close examination of the mutual information plot shows that the secondary structure in figure 12 is the closest to the correct secondary structure of DA22/96 ribozyme.

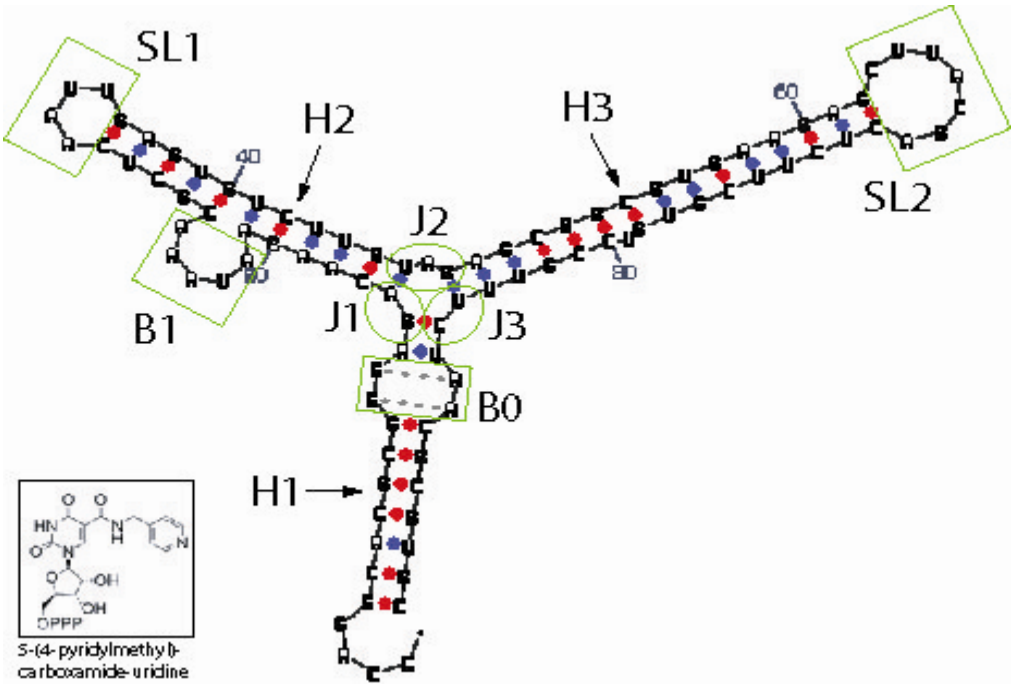


Figure 13 The predicted secondary structures of DA22/96.J mean junction. B mean bulge, SL means stem-loop, H mean helix. SL1, B1, J1, J2, J3, SL2, B0 are the enzyme access sites.

The results from the mapping experiment also proves the above secondary structure for DA22/96 is the correct one.

3.2.1.2 Mapping of Modified DA22/96 and Unmodified DA22/96

In order to test the validity of the structures proposed by MFOLD, RNase mapping experiments were conducted. The analysis of these experiments led to experimental constraints on the overall conformation of wild type DA22/96. DA22/96 is probed *in vitro* with a single strand-sensitive mung bean nuclease. We have produced consistent probing data with predicted secondary structures DA22/96. We observe that the asymmetric B1 and SL1&2 are accessible to cleavage on both the modified and unmodified DA22/96 ribozyme. Symmetric B0 region actually forms the helix continuous with H1 or G-A base pairs. The junction regions surrounding predicted catalytic center are not accessible to cleavage. The inaccessible regions are conserved areas. The SL1 region in DA22/96* does not have the signal in gel because the modified uridines will resist the MB cleavage. Three junction areas, j1, j2, and j3 are inaccessible in the gel, thus showing the junction region as the predicted catalytic center.

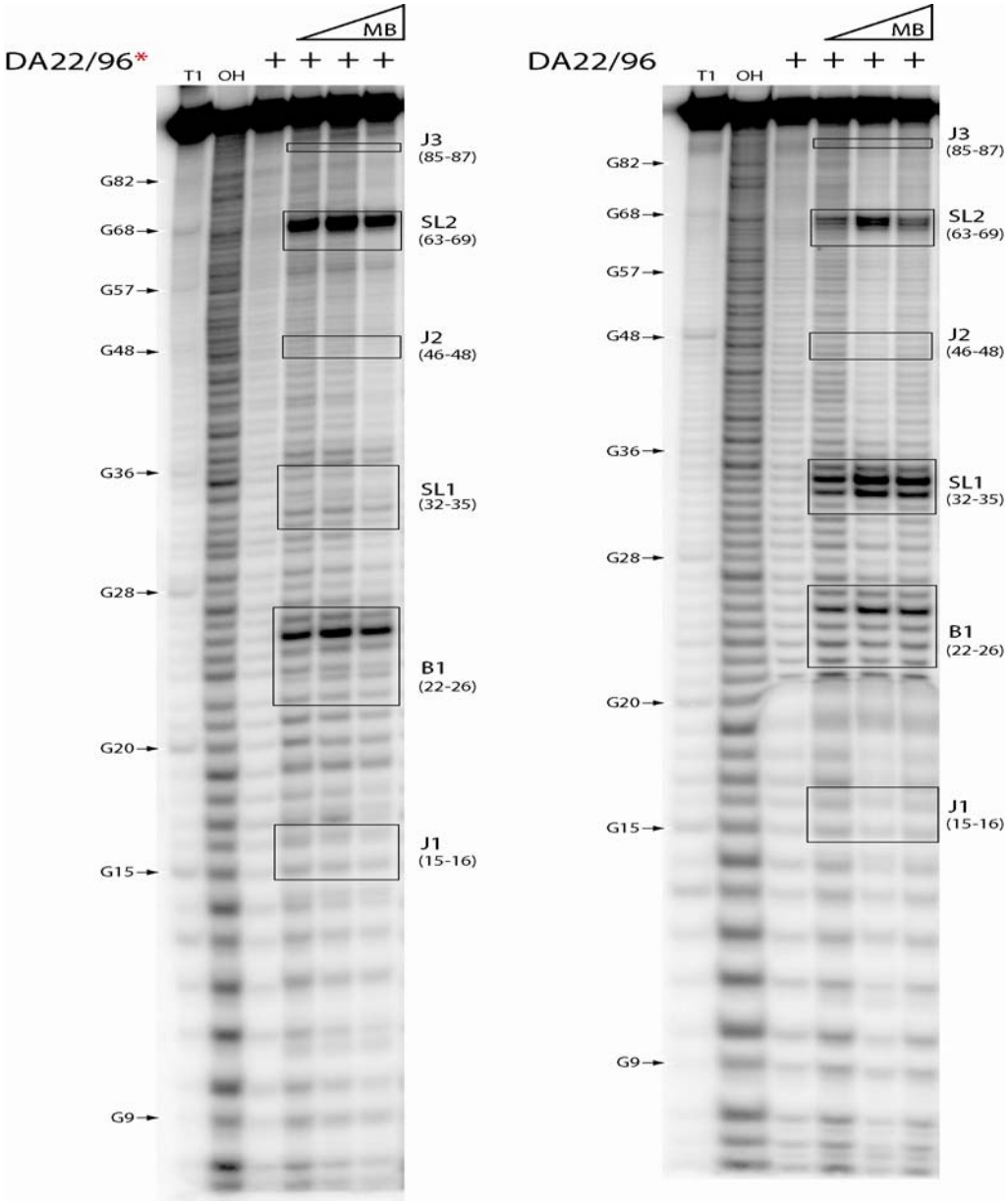


Figure 14 The result of mapping modified DA22/96 and unmodified DA22/96.

* =contain modified uridine

3.2.1.3 Comparing cleavage of modified and unmodified DA22/96 Ribozyme

Lead (II) acetate (Pb^{2+} acetate) is a useful and versatile probe to extract information on RNA structure. The ion primarily not only induces highly specific cleavages at positions of tight metal ion binding sites, but also cleaves

RNA within single-stranded regions, loops and bulges. The secondary structure for both modified and unmodified DA22/96 ribozyme appears to be identical except for the reduced cleavage and upwardly shifted band where modified uridines occur. Red dots are placed to the right of each uridine band in the MB digested modified and unmodified lanes (3 and 4) to demonstrate the loss of signal and upward shift. This could either be the result of the modified uridine size or chemistry of modification 5-(4-pyridylmetnyl) carboxamide-uridine.

The following mapping data suggested that modified uridine did not significantly affect the secondary structure of DA22/96 but may instead have its greatest impact on catalytic site chemistry reaction.

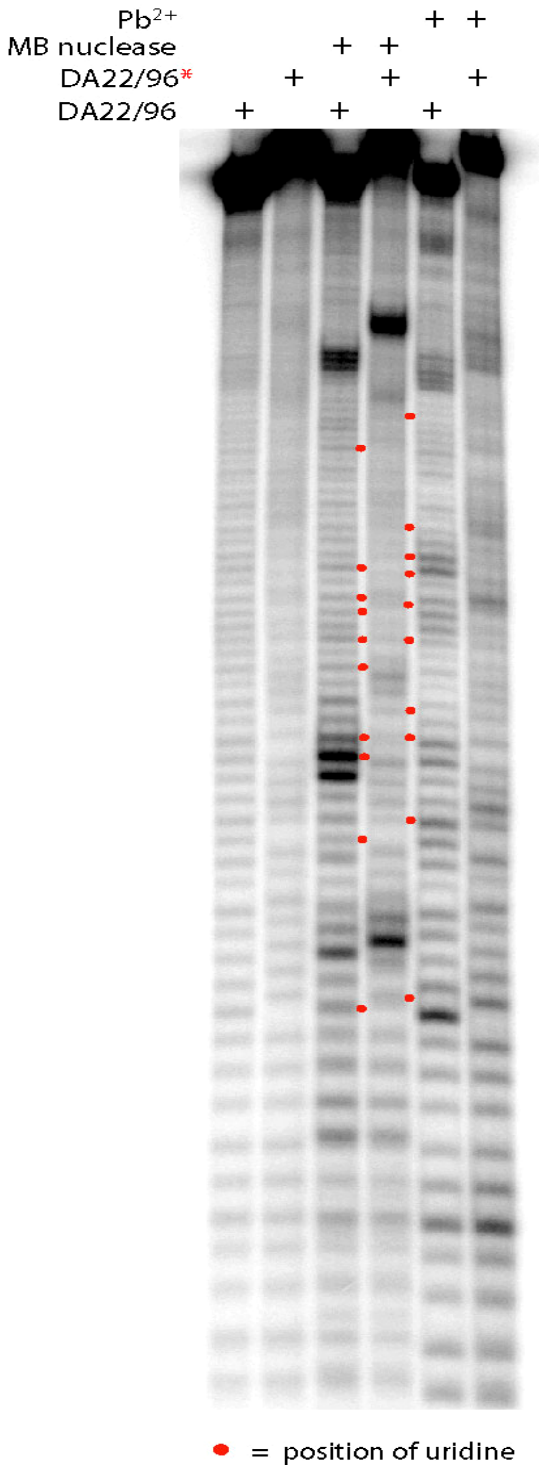


Figure 15 The result of cleavage of the modified DA22/96 and unmodified DA22/96

3.2.1.4 Mapping of substrate binding to DA22/96

Enzymes and chemical agents that cleave or modified RNA with the specificity can be used to analyze the structure of RNA. In this experiment, MB nuclease and Pb^2 were used for the probing of single stranded and dynamic region of the RNA. The reason for selecting MB (with single strand-sensitive mung bean) nuclease to map the DA22/96 secondary structure is that MB achieves better results. We tried to use other nucleases to cut the RNA, however, the mapping results were not as clear as when we used the MB. *When adding the* copper or different substrates in the reaction, no significant structure change was detected in the mapping gels, meaning either the copper or different substrates significant alterations in their structure. This may reflect the relatively slow turnover rate of the reaction and indicating that either a small change in the structure during catalysis occurred or that only a small population of RNA bound at any given time.

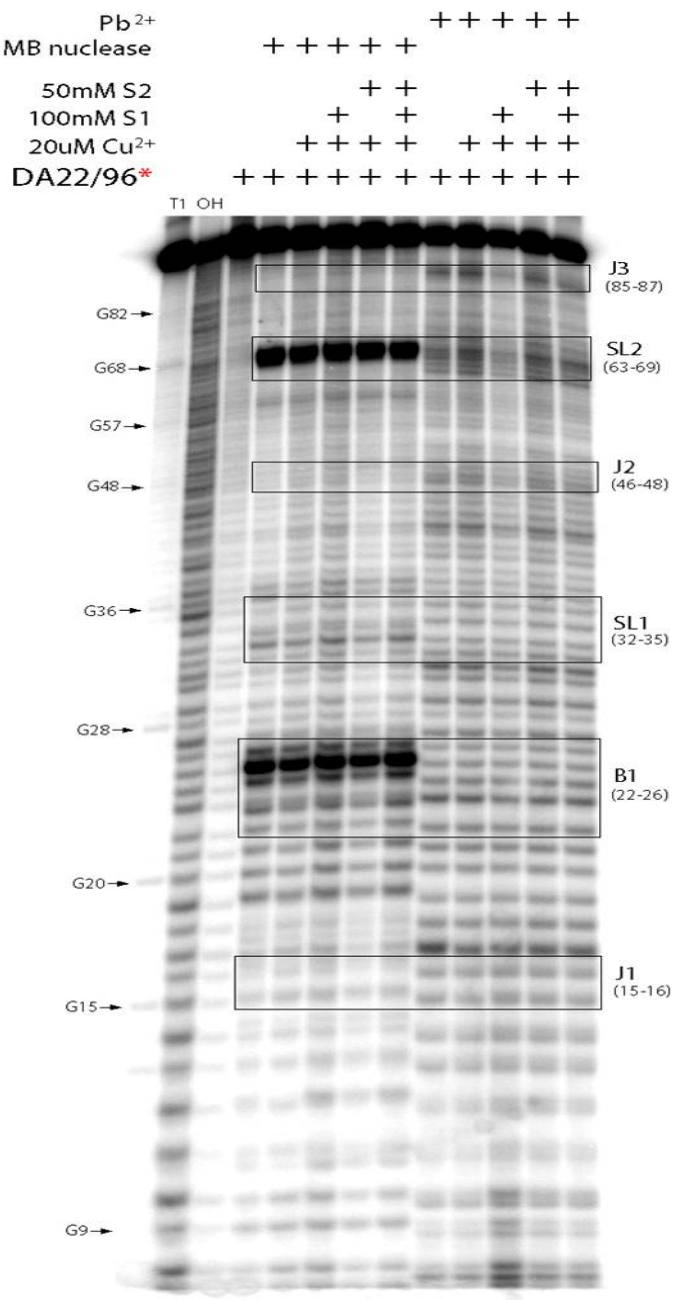


Figure 16 The result of the substrate binding site mapping.

3.2.2 Discussion

A RNA Diels-Alderase (DA22) containing modified uridine base has been reported to have catalytic [4+2] cycloaddition the same as a protein that functions like a catalyst[33]. The question is how the DA22/96 RNA primary sequence and concomitant secondary structure relates to RNA Diels-Alderase activity for this particular diene and dienophile pair. The results from the mapping experiments, we could conclude that the junction region is the catalytic activity center for DA22/96. The three helix stabilized the RNA molecular when RNA brought substrates together. The DA22/96 has the similar junction structure as the J49 ribozyme. This implies that the full-length DA22Diels-Alderase has the similar junction structure.

In figure 15, the structures of the two RNA are identical with the exception of two things: the strength of cleavage and the position of the cleavages. The cleavage of unmodified DA22/96 shows a stronger signal, and therefore creates darker bands at the sites of cleavage on the gel. The loss of signal in the modified DA22/96 is observed by the lighter bands in gel lane 4 containing DA22/96*. The positions of the cleavages are also slightly different. The cleavage sites in modified DA22/96* appear to be shifted upwards when in comparison with the unmodified DA22/96. Although the bands are all shifted in lane 4 of the gel, the spacing between each band in that lane has identical spacing as the band spacing in lane 3, where the unmodified DA22/96 ran. The similar band spacing suggests that the modified uridine does not significantly affect the secondary structure of DA22/96, leading to the conclusion that the greatest impact of the

modified uridine may lay in catalytic site chemistry. The upward shift of the modified uridine is caused by the difference in molecular weights between the modified and unmodified RNA. The modified DA22/96* is heavier, due to the modified uridine, increasing the molecular weight of the RNA. The increased weight of the RNA causes the RNA to run slower in the gel, and thus causes an upward shift in the bands of uridine digested.

RNase mapping method is used to determine the secondary structure because it only takes a small amount of RNA (pmol) to run the mapping analysis. From an economic perspective this method is the better choice than other methods like MS and NMR. Some residues may not be accessible to chemical or enzymatic probe because an RNA molecule is folded in a specific manner. In this case, changed digestion pattern at different temperatures may provide additional information. Also there are other nucleases like S1 nuclease, RNase V1 or chemicals like DMS, DEPC and CMCT that can be chosen to confirm the mapping results.

REFERENCES

[1] Seelig,B.; Keiper, S.; Stuhlmann, F.; Jaschke, A.; Angew. *Chem. Int.Ed.2000, 39, No.24, 4576-4579*

[2] Selectivity and Novel RNA-Catalysed Activity: A. Jaschke, F. stuhlmann, D. Bebenroth, S. Keiper and R. Wombacher. *2002 Biochemistry Society*

[3] Tarasow, T.; Kellogg, E.; Holley, B.; Nieuwlandr, D.; Tarasow, S.; Eaton. B.; *J. AM. Chem. Soc. 2004, 126. 11843-11851*

[4] Theodore, T.; Tarasow, M.; Tarasow, S.; Tu, C.; Kellogg, E.;Eaton, B.;. *J. AM. Chem. Soc., 121 (15), 3614-3617, 1999*

[5] Stuhlmann, F.; Jaschke, A.; *J. AM. Chem. Soc. 124(13) 3238-44*

[6] Tarasow, T.; Tarasow, S.; Eaton, B.; *J. AM. Chem. Soc. 2000, 122, 1015-1021*

[7] Highlights in Bioorganic Chemistry: Carsten Schmuck, Helma Wesnnemers, *433-435*

[8] Frauendorf, C.; Jaschke, A.; Angew. *Chem. Int. Ed. 1998, 37(10), 1378-1381*

[9] Seelig, B.; Jaschke, A.; *Chemistry & Biology March 1999, 6:167-176*

[10] Tarasow, T.; Eaton, B.; *CMLS, Cell. Mol. Life Sci. 5 (1999) 1463-1472*

[11] RNA- catalysed carbon-carbon bond formation: Theodore M. tarasow, Sandra L. Tarasow and Bruce Eaton. *Nature September 1997(4), vol 389, 54-57*

[12] Dewey, T.; Eaton, B.; *J. AM. Chem. Soc. 2004, 126, 11231-11237*

[13] Schlatterer, J.; Dipl.-Chem, Stuhlmann, F.; Jaschke, A.; *ChemBioChem. Volume 4 Issue 10, Pages 1089-1092*

[14] T. Tarasow; S. Tafasow; B. Eaton; *Nature, 1997(9), vol 389, 54-57*

[15] Serganov, A.; Keiper, S.; Malinina, L.; Tereshko, V.; Skripkin, E.; Hobartner, C.; Polonskaia, A.; Tuan Phan, A.; Wombacher, R.; Micura, R.; Dauter, Z.; Jaschke A.; Patel.D.; *Nature Structural and Molecular biology 2005(8), vol 3, 218-224*

[16] Tarasow, T.; Tarasow, S.; Eaton, B.; *J. AM. Chem. Soc. 2000, 122, 1015-1021*

[17] RNA-Catalyzed Carbon-Carbon Bond Formation: Andres Jaschke. *Biol. Chem., Vol 382,1321-1325*

[18] Selectivity and Novel RNA-Catalysed Activity: A. Jaschke, F. stuhlmann, D. Bebenroth, S. Keiper and R. Wombacher. *2002 Biochemistry Society*

[19] Keeper, S.; Bebenroth, D.; Sellig, C.; Westhof, E.; Jaschke, A.; *Chemistry & Biology Vol 11, Issue 9, September 2004, pages 1217-1227*

[20] Seelig,B.; Keiper, S.; Stuhlmann, F.; Jaschke, A.; Angew. *Chem. Int.Ed.2000, 39, No.24, 4576-4579*

[21] Stuhlmann, F.; Jaschke, A.; *J. AM. Chem. Soc. 2002 124, 3238-3244*

[22] Tinoco, I.; Uhlenbeck, O.; Levine, M.; Estimation of Secondary RNA Structure in Ribonucleic Acids, *Nature Vol 230, April 9 1971, 262-267*

[23] Gutell, R.; Power, A.; Hertz, G.; Putz, E.; Stormo, G.; Nucleic Acids Research, *Vol 20, No. 21, 5785-5795*

[24] Zuker, M.; Mfold web server for nucleic acid folding and hybridization prediction; *Nucleic Acids Research Vol. 31, No.13 , 3406-3415*

[25] RNA Folding by Comparative Sequence Analysis: Michael Zuker, Rensselaer, *August 18 1998*

[26] Tarasow, T.; Kellogg, E.; Holley, B.; Nieuwlandr, D.; Tarasow, S.; Eaton, B.; *J. AM. Chem. Soc. 2004, 126. 11843-11851*

[27] Regina, M.; John, N.; Maxam, A.; *J. AM. Chem. Soc. 1978, 4493-4499*

[28] Helm, M.; Petemeier, M.; Ge, B.; Fiammengo, R.; Jaschke, A.; *J. AM. Chem. Soc. 2005, 127, 10492 – 10493*

[29] Tarasow, T.; Kellogg, E.; Holley, B.; Nieuwlandr, D.; Tarasow, S.; Eaton, B.; *J. AM. Chem. Soc. 2004, 126. 11843-11851*

[30] Tarasow, T.; Tarasow, S.; Eaton, B.; *Nature Vol 389 4 September 1997*

[31] Tarasow, T.; Kellogg, E.; Holley, B.; Nieuwlandr, D.; Tarasow, S.; Eaton, B.; *J. AM. Chem. Soc. 2004, 126. 11843-11851*

[32] Zuker, M.; *Nucleic Acids Researchs 2003, Vol 31, No 13, 3406-3415*

[33] Tarasow, T.; Eaton, B.; CMLS, Cell. Mol. *Life Sci. 55 (1999) 1463-1472*