ABSTRACT

ZHAO, JIE. Improved RNA Diels-Alderase by Mutation. (Under the direction of Bruce Eaton.)

DA-22 is a well-studied Diels-Alderase with good catalysis. DA-2 and DA-22 share a highly conserved sequence. The RNA selection would have converged on a local catalytic maximum and that further mutation might lead to better catalysis. My research was to mutagenize DA-2 by mutagenic PCR to diversify the number of sequences and to determine if better Diels-Alderases could be found. RNA Diels-Alderase DA-2 was selected through *in vitro* evolution cycles. The strategy is to transcribe mutagenized DNA to RNA with T7 RNA polymerase, conjugate RNA with a diene via a PEG linker in a DNA bridge-mediated ligation, incubate RNA-diene conjugate with dienophile, then use streptavidin to separate the reacted and unreacted RNA. After the separation, RNA can be reverse transcribed to DNA and then the cycles can be repeated.

Improved RNA Diels-Alderase by Mutation

by

Jie Zhao

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

2004

APPROVED BY

Chair of Advisory Committee

DEDICATION

I would like to dedicate this work to my parents in China and my husband. They gave me unconditional love, support and encouragement during the course of my studies in U.S. I also like to dedicate it to my dearest extended families, Mr. Blevins' in Statesville. It was them who helped me to know American families and this culture. They comforted an anxious Chinese girl's heart with their loves and gave her, who first came to U.S., strength.

BIOGRAPHY

Twenty-eight years ago, a girl was born in a beautiful city of Northwest China, named Urumqi and her name was Jie Zhao, the author of this work.

She loved science and sports. She had been a track star since middle school. In 1994 she went to Nankai University to pursue her B.S. degree in chemistry. Immediately after she graduated in 1998, she was admitted by the chemistry graduate school with scholarship and had a rare exemption from the required exams. In 2001 she got her M.S. degree in analytical chemistry from Nankai University. Then she had a good opportunity to continue to work in this area in NCSU in U.S. Living abroad was really exciting and it seemed like a total new life to her. She was learning new stuff every day. Working on DNA and RNA in Dr. Eaton's research group broadened her knowledge to biochemistry area and benefited her a lot.

ACKNOWLEDGEMENTS

I would like to thank a lot of people who helped me on my academic studies in chemistry department of NCSU. First of all, I want to thank my advisor, Dr. Eaton, who gave me the chance to join his research group, led me to a DNA, RNA world, a joint of biology and chemistry, and let me gain good experiences in a new area to me. I also want to thank Dr. Bowden, who sets himself a good example to everyone. While I knew little about U.S. three years ago, it was him who guided me towards the right direction at my first stop in U.S. Thanks to Dr. Franzen, who ignited my interest in nucleic acids and proteins through his class. A special thanks to Dr. Maxwell in biochemistry department, who gave me a huge academic help in the area of biochemistry gene expression. He helped me to build up a theory foundation for my research. Finally I want to thank everyone in Eaton's group, Ben, Lina, Dana, Tong and Jon, who helped me on my research and made me feel at home.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 Catalytic RNA

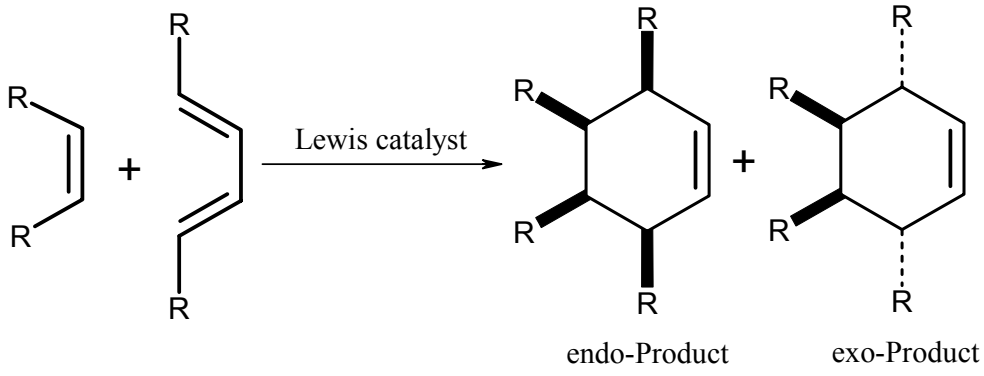
With the discovery of self-splicing[1,2] and self-cleaving[3-7] catalytic RNA (ribozymes, RNA molecules with catalytic properties), naturally occurring ribozymes have been shown to efficiently catalyze the formation and cleavage of nucleic-acid phosphodiester bonds.[8] These ribozyme-substrate interactions are governed by nucleobase-pairing interaction.

Oligonucleotides are very unique among macromolecular catalytic platforms.[9] RNA was once thought to only be able to interact as catalysts with other nucleic acid substrates and not to possess the required functionality to catalyze the synthesis of small organic structures. Can RNA be made to catalyze reactions that are fundamentally different than those catalyzed by natural ribozymes? In contrast to proteins with diverse functional groups as side chains that can provide intricate interactions with substrates or cofactors,[9] nucleic acids are lacking of functional group diversity. Incorporation of modified nucleotides with a variety of functional groups, which extend even beyond the diversity present in proteins has overcome this limitation.[10,11] Recent advances in the chemical and enzymatic modification of RNA are being used to increase the structural and functional diversity in nucleic acids and provide the opportunity to tune the functionality of the macromolecular platform for the desired catalytic outcome.[9-24] In addition, RNA catalysts have demonstrated substrate specificity[25,26] and sometimes stereospecific product formation.[27] Combining these properties makes catalytic oligonucleotides a useful

platform for creating new catalysts and studying reaction mechanisms. The evidence is that the narrow range of RNA-catalyzed reactions has now been subsequently expanded by *in vitro* selection methods to include acyl transfers,[28-38] peptidyl transfers,[39] ester and amide bond formation,[40,41] cofactor synthesis,[42] Michael additions,[43] nucleophilic substitutions,[44] porphyrin metallations[45,46] and Diels-Alder cycloaddtions.[47-50]

1.2 Diels-Alder reaction

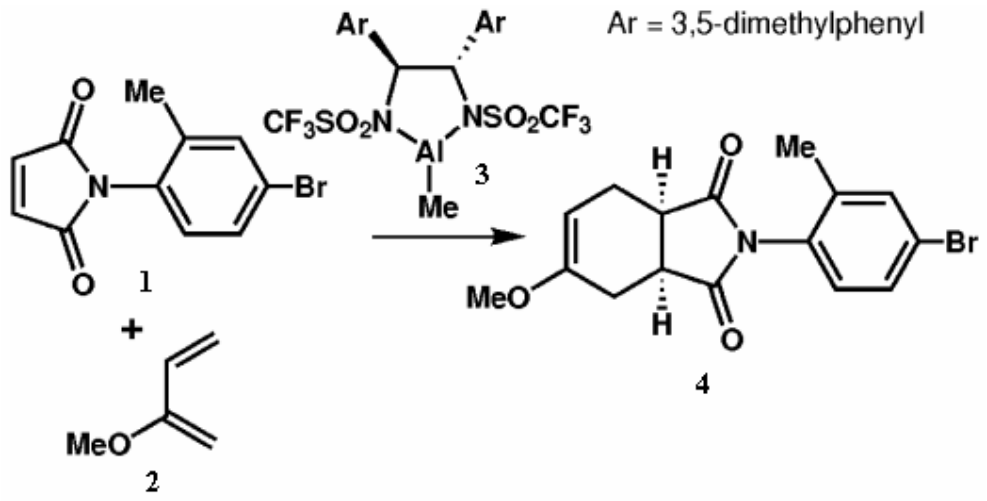
Carbon-carbon bond formation and the creation of asymmetric centers are both very bioactively important. Diels-Alder cycloaddition is historically among the most powerful and widely used of all synthetic organic transformations.[51] It creates two carbon-carbon bonds and up to four new stereocenters and can typically be accelerated by Lewis acid catalysts. There are exo-diasteroselectivity and endo-diasteroselectivity as shown in equation 1.1, but few known catalysts can give the exo-diasteroselectivity for an acyclic 1,3-butadiene in a normal demand Diels-Alder reaction.

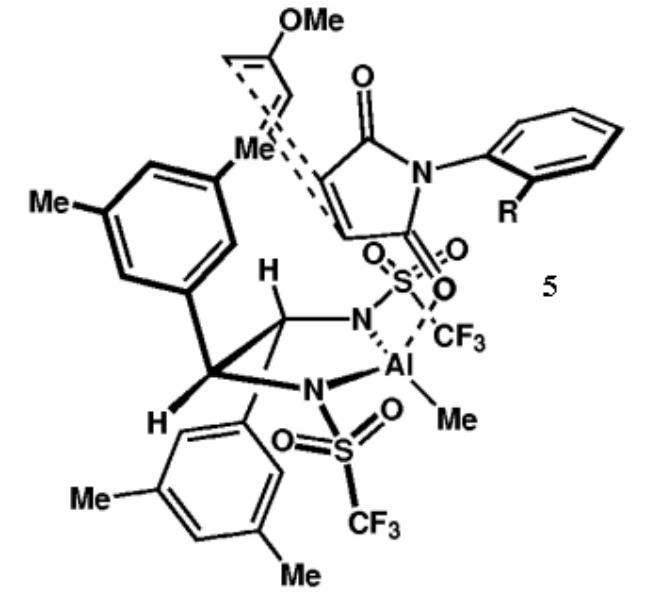


Equation 1.1. Diels-Alder cycloaddition

π-electron-rich aromatic groups in [4+2] catalysts could be used both to stabilize specific transition-state geometries and to provide stereoselectivity for the development of enantioselective catalysts for Diels-Alder reactions. In figure 1.1, the dienophile (1) was activated by Al-coordination and the structure (5) allowed optimal π-π interaction between the Al-coordinated carbonyl group and C2 of the nearby 3,5-dimethylphenyl group in (3) to stabilize the specific transition-state geometries. The phenyl group in (3) can block the [4+2] cycloaddition of (2) to the front face of N-arylmaleimide (1) to favor the formation of (5) by forcing the diene to add to the rear face, rendering this endo-product stereoselectivity. The stereoselectivity of the reaction was high and formation of (6) was disfavored.[52,53]

RNA catalysis to assemble organic molecules could provide major advancements in understanding the mechanisms of these organic reactions. It may also reveal new possibilities for low reactivity cycloaddition substrates and diastereoselectivity.


Ar = 3,5-dimethylphenyl
Ar
Ar
CF3SO2N
NSO2CF3
Al
Me
3
O
Me
N
Br
O
1
+
MeO
2
H
O
Me
N
Br
MeO
H
O
4


OMe
Me
O
N
R
Me
O
S
O
H
N
CF3
Al
N
H
Me
O
S
O
Me
CF3
Me
5

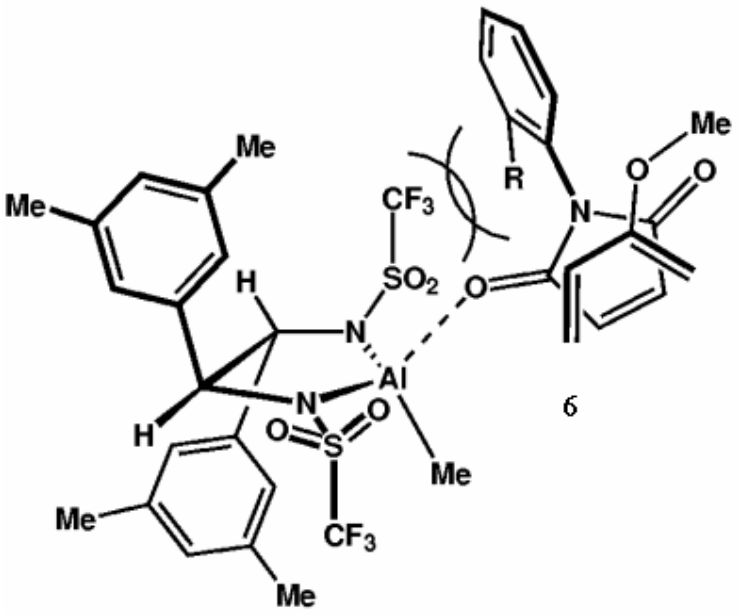


Figure 1.1. Enantioselective synthesis of 4 through the reaction of 2-methoxybutadiene 2 with the N-arylmaleimide 1 in the presence of catalyst 3. 5 was favored form and 6 was disfavored form.

2. LITERATURE REVIEW

2.1 RNA Diels-Alderase binding interactions

Most RNA structures are not catalysts, but when RNA is chemically modified with functional groups like imidazoles and pyridines used in an *in vitro* selection, its catalytic capabilities are expanded. A Diels-Alder reaction catalyst was reported consisting of modified RNA with additional functionalities attached to the nucleobases.[47] The pyridyl modification expands the functional diversity of the RNA by providing additional hydrogen bonding, hydrophobic and dipolar interactions as well as an alternative

sigma-donor ligand for metal ion binding. Novel metal binding sites created by the precise positioning of pyridyl groups may be crucial to the Diels-Alderase activity. Most of the RNA Diels-Alderases have demonstrated a high degree of substrate specificity and a dependence on the presence of cupric ion (Cu^{2+}). Modified RNA bound with crucial metals folds specifically to recognize the two reactants. Hydrophobic interactions play an important role between RNA Diels-Alderases and substrates. To date, no catalysis of carbon-carbon bond formation by unmodified RNA has been described.[34] Either the RNA or the substrates has to be modified. In the work of Jaschke on the Diels-Alder [4+2] reaction, no modified bases were used in the RNA, but modified substrate, anthracene (modified diene), served as a diene in Diels-Alder reactions and was also a good intercalator for binding nucleic acids.[27,50]

RNA can fold up in three-dimensional complex shapes with intricate binding pockets. Some RNA can form catalytic pockets and some may only bind small molecules and proteins. Proposed secondary structures based on the Zucker folding algorithm are shown in figure 2.1. It is evident from the structures that the RNA Diels-Alderases have the potential to generate extended helical regions resulting in stable secondary structures.[48] It is clear that large portions of the secondary structures involve the constant regions, which are required for the various enzymatic steps of *in vitro* selection cycles. The lowest energy structures generated by the Zucker algorithm show the locations and secondary structures of the 10 base conserved regions, which are largely centralized in the primary sequence and are involved in a variety of secondary interactions. The 10 base

conserved regions have been expanded in figure 2.1 to show that they may share similar secondary structures consisting of a bulge and they can fold with the bulges in a special manner. It is likely that these conserved regions play similar roles in creating catalytically active folded topologies. The mechanisms by which RNA accelerates Diels-Alder cycloaddition have received much attention in recent years and there is currently much interest in developing methods for improving its rate and selectivity.

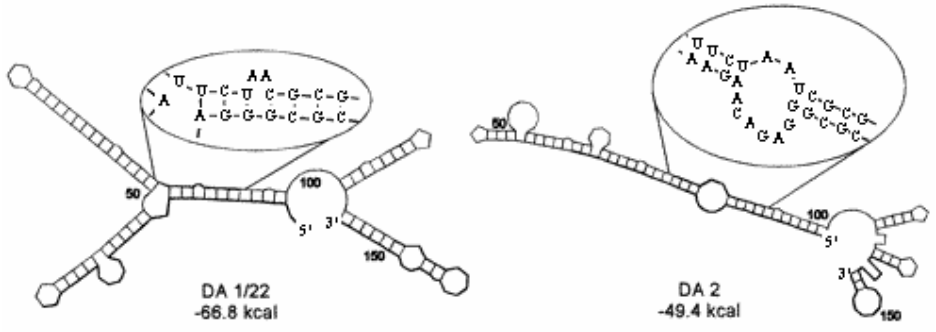


Figure 2.1. Proposed secondary structures of the two RNA Diels-Alderase isolates. The 10 base conserved sequences have been enlarged to show secondary structural details.

2.2 RNA Diels-Alderase *in vitro* selection

A variety of catalytic molecules can be present in a diverse pool of nucleic acids with random sequence. How to identify or isolate those rare molecules that have the desired catalytic properties is a challenging problem. The use of combinatorial methods is playing a central role in expanding RNA Diels-Alderase functions. An *in vitro* selection scheme to isolate RNA molecules with Diels-Alderase activity was carried out as

described previously.[41,47] The strategy is to generate a randomized pool of RNA that is transcribed from DNA. Many different modified ribonucleotides can be efficiently incorporated into the transcripts with high fidelity by using T7 RNA polymerase. Compared to natural nucleotide catalysts, the scope of catalysis can be greatly increased by adding functional groups to the nucleotide bases. Conformational flexibility, nucleophilicity, electrophilicity, metal coordination sites and acid or base characteristics are then added.[41,47] T7 RNA polymerase can well tolerate these modifications with regard to sequence diversity and overall fidelity. RNA conjugates can be transcribed with a diene attached to a long flexible polyethylene glycol (PEG) tether and a DNA 10-mer by bridge-mediated ligation of the 10-mer onto the 5'-end of the RNA. The long flexible PEG linker (average MW 2,000) allows the diene to survey the entire surface of the RNA. This relatively inert linear polymer is soluble in aqueous buffers so that it is used for many RNA catalyzed reactions and may improve the solubility of substrates in some cases. Cycles of *in vitro* selection are carried out by incubating the RNA-diene conjugate with the dienophile. The resulting pool of product bears a biotin substituent contributed by the dienophile for selection of RNA molecules that catalyze carbon-carbon bond formation between the two reactants. Selective capturing of the active nucleic acid sequences is a crucial factor in determining the success of a RNA catalyst *in vitro* selection. A convenient technique for partitioning active from inactive RNA sequences involves capture of biotin by streptavidin. Biotin is a widely used affinity tag because of its strong binding to streptavidin. The newly formed biotinylated Diels-Alder product

would then be covalently attached to the RNA through the PEG linker allowing for separation of reacted and unreacted RNA molecules by virtue of streptavidin binding and concomitant electrophoretic mobility shift. The streptavidin dependent gel shifted product can be quantitated by phosphorimaging or scintillation counting of excised gel bands.[47] After removal of all nonbiotinylated members of the pool, active RNA molecules can be reverse transcribed and the resulting PCR-amplified DNA used in the next cycle is the template for transcription. Multiple cycles of *in vitro* selection produce a population of RNA molecules capable of catalyzing the desired Diels-Alder reaction. Mixtures of transition metals are used in the incubation for Diels-Alder cycloaddition. They could form unique structural or Lewis acidic sites with the modified RNA. Divalent metal ions are essential components present in successful nucleic acid catalyzed reactions, such as Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Al^{3+}, and Ga^{2+}.[35,41] A buffer without functional groups that could react is also required. HEPES (2-[4-(2-hydroxyethyl)-1-piperazine] ethanesulfonic acid) is a commonly used buffer. These essential components maintain RNA tertiary structure and prevent its aggregation.[43,47,54] Unlike any other molecule type, nucleic acids can be enzymatically replicated and this provides the opportunity to directly select for a desired catalytic function by *in vitro* techniques.

2.3 Mutagenic PCR

If a significant number of mutations were introduced during PCR amplification, an *in vitro* evolution would diversify the pool of catalytic RNAs. Most practitioners of PCR prefer to carry out DNA amplification in an accurate manner, introducing as few base

substitutions as possible. This is especially critical when one is studying clonal isolates and must distinguish natural variation from artifactual variation that is introduced by polymerase error. Thermostable DNA polymerases are available that operate with high fidelity because of an intrinsic 3' to 5' exonuclease activity. Manipulation of PCR conditions can lead to further improvement of copying accuracy.[55] The other side of the fidelity issue can also be considered, those instances where promiscuity is a virtue. In probing the structure or function of a protein or nucleic acid, it maybe desirable to generate a library of mutants and apply a screening method to isolate sequences that exhibit a particular property.[56-59] For mutations over a short sequence within a cloned gene, it is appropriate to replace a portion of the gene with a synthetic DNA fragment that contains random or partially randomized nucleotides. One convenient way to introduce random mutations is through inaccurate copying by a DNA polymerase, especially if the polymerase is a thermastable enzyme that can operate in the context of the PCR. Each pass of the polymerase during the PCR allows for the possibility of mutation, so that the cumulative error rate can become substantial. The error rate of Taq polymerase is the highest of the known thermostable DNA polymerase. Over the course of 20-25 cycles of the mutagenic PCR, the polymerase can make a mutation rate from one to a few mutations per base. 20-25 cycles of the mutagenic PCR are carried out to ensure that mismatched termini have ample opportunity to become extended to produce complete copies and there are very little chance of any two molecules isolated from the final population carrying the same mutation as a consequence of deriving from a common

ancestor. Therefore, after mutation, more sequences are created.

Through mutagenesis a cloned catalyst can be re-randomized and examined for changed activity. An evolving pool can periodically be mutagenized and re-selected. After subpopulations are selected, they can then be used directly as input for a second evolution cycle of mutagenic PCR. Repeating the cycle of selection and mutagenic amplification allows one to carry out *in vitro* evolution of nucleic acids, including those that have catalytic function.[60-62] Theoretically this can result in selection convergence on sequences around the most active catalytic motifs. The type of mutagenic PCR employed can control the degree of mutation introduced at each position. Chemical mutagenesis uses incorporation of nucleotide analogs to cause base pair mismatching during PCR amplification.

Important advantages of mutagenic PCR are to introduce the various types of mutations in an unbiased fashion rather than to achieve a high overall level of amplification and to allow repeated randomization of populations of nucleic acids without isolating clones and obtaining sequence information.

3. *IN VITRO* SELECTION OF RNA DIELS-ALDERASE IN MY RESEARCH

3.1 Objectives

My research is to utilize mutagenic PCR and isolate mutated RNA molecules that catalyze the formation of carbon-carbon bonds between a tethered diene and a

biotinylated maleimide dienophile by [4+2] cycloaddition (Diels-Alder reaction).

3.2 Materials and methods

The process begins with dsDNA DA-2. The dsDNA templates contain fixed regions required for the enzymes used in the amplification. The fixed regions were not mutated, as these are essential for PCR amplification and the other enzymatic steps of the selection process. The center of the templates contains a variable region of 100 nucleotides long that is practically the maximum length in DNA synthesis of a random sequence, as shown in figure 3.1.[47] Mutation is only in the 100 nucleotide random region. Mutation in this variable region can diversify the number of sequences after cycles of selection to help explore the details of the evolving catalytic region. It seems reasonable that further mutation might lead to better catalysis.

Mutagenic PCR is derived from "standard" PCR conditions: 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris (pH 8.3 at 25°C), 0.2 mM each dNTP, 0.3 μM each primer and 2.5 units of Taq polymerase in a 100-μl volume, and the reaction mixtures were incubated for 30 cycles of 95°C for 1 min, 55°C for 1 min and 72°C for 1 min in a conventional thermocycler. Some changes are made to enhance the mutation rate and introduce the mutagenic PCR. In my research, a RNA sequence (DA-2) was mutated using PCR in the variable and highly conserved sequence regions (Scheme 3.1) while maintaining the fixed regions by increasing the concentrations of Mn^{2+} and Mg^{2+}.

Family 1

DA-01

GUGAGUGUCUUGUAGAGCGGCGUGAAGAGCUUACGACUCUUCGUGUCCGU**UUCUAACGCG**UGUCCAUGAACCCGUCCCUAGUGGGGGGGGGCGCCUG

DA-14

CUGAGUGUCUUGUAGAGCGGUGUGAAGAGCUUACGACUCUUCGUGUCCGU**UUCUAACGCG**UGCCCAUGAACCCUCCCUAGUUGGGGGGGCGCCUG

DA-22

UUGAGUGUCUUGUAGAGCGGCGUGAAGAGCUUACGACUCUUCGUGUCCGU**UUCUAACGCG**UGCCCAUGAACCCAUCCCUAGUGGGUGGGGGCGCCUG

DA-47

UUGAGUGUCUUGUAGAGUGGCGUGAAGAGCUUACGACUCUUCGUGUCCGU**UUCUAACGCG**UGCCCAUGAACCCAUCCCUAGUGGGUUGCGGCGCCUG

Family 2

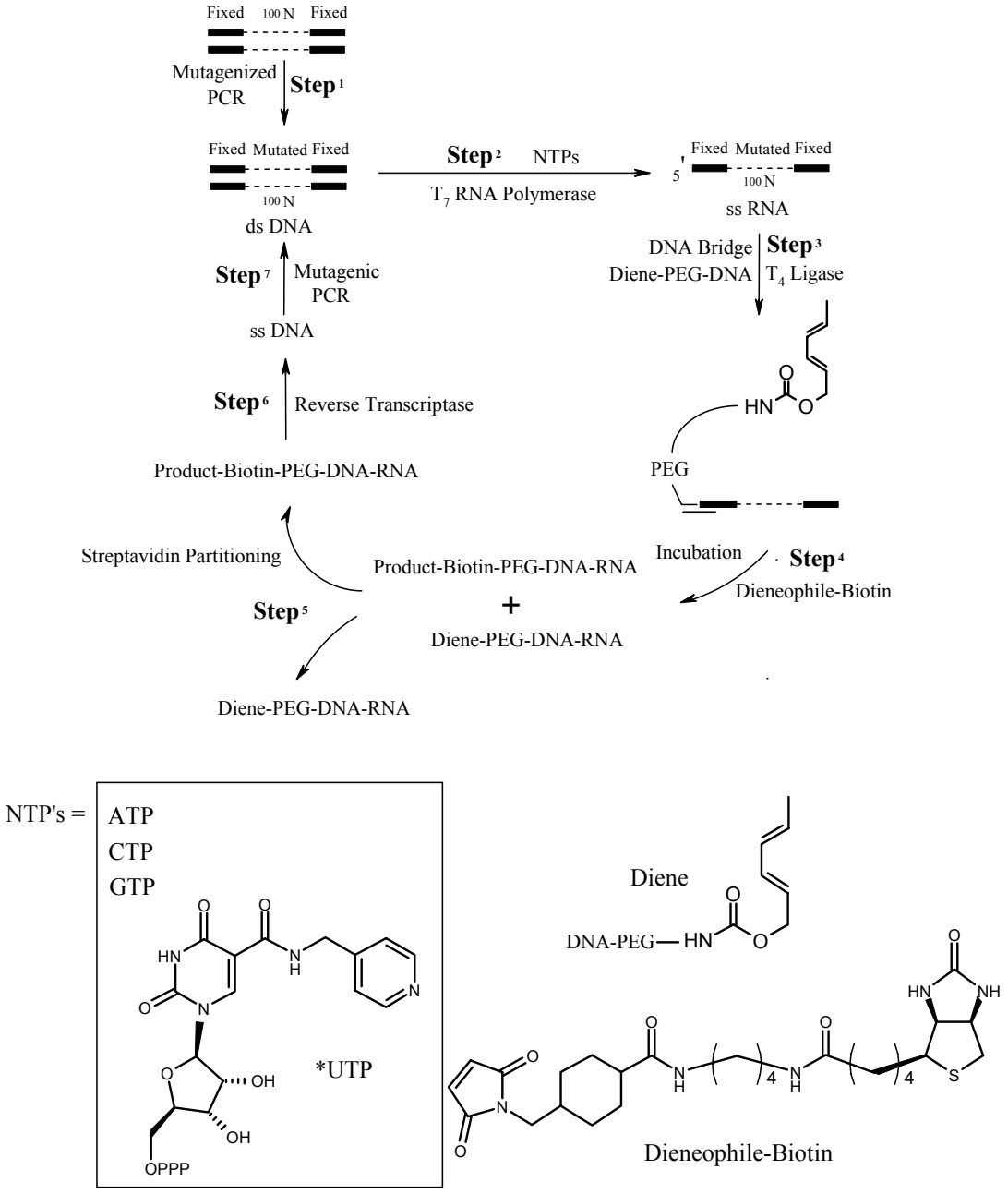
DA-2

GGCUUUCCAGUAUGGAUACCCUUGGGGCUCUUUGUAGACAGCGUAGUUG**UUCUAAUCGCG**UAGGCCUGGUGAGGUACGUGUUAGAUGCGGUUAUCGAGGG

Figure 3.1. Variable region (~100 N) and highly conserved region underlined, sequences of 5 isolates obtained from the Diels-Alderases *in vitro* selection. The number to the left of the sequences identifies isolates. Members of clonal families are labeled. The computer-identified consensus sequence is in bold and underlined for each of the isolates from families.

In step 1 (Scheme 1), 0.5 mM Mn^{2+} was added to diminish the template specificity of the polymerase;[63,64] higher Mg^{2+} concentration (7 mM) was used to stabilize noncomplementary pairs;[65,66] The Concentrations of dCTP and dTTP were increased five fold over that of dGTP and dATP to promote misincorporation;[64,67] Doubling the units of Taq polymerase to 5 units was done to promote chain extension beyond positions of base mismatches.[68] Reaction mixtures were incubated for 30 cycles of 95°C for 1 min, 55°C for 1 min, 72°C for 1 min and the reaction products were purified with a silica purification kit (Qiagen). Then a small portion of the purified products was run on a 6% analytical polyacrylamide (19:1) gel stained with ethidium bromide to confirm a

satisfactory yield of full-length material.



Scheme 3.1. Steps of the RNA catalyst selection cycle

In step 2 (Scheme 1), mutagenized dsDNA library was then transcribed to produce a new RNA library with T7 RNA polymerase. During transcription step it is easy to incorporate additional functional groups into the RNA by using modified UTP. The pyridine modified UTP shown in Scheme 1 (*UTP) has been shown to be essential in a selection for RNA Diels-Alderases. 20 mM GMP, 0.5 mM each of ATP, GTP, CTP and 1 mM modified UTP (in 100 ul total reaction volume) were incorporated. High concentration of GMP was used to make sure that the 5' end of transcribed RNA was GMP for Diene-PEG-DNA ligation. Reactants were incubated at 37°C for 6 hours.

Once the transcribed RNA was purified, in step 3 (Scheme 1) it was then attached to a diene via a long flexible polyethylene glycol (PEG) and a DNA 10-mer by bridge-mediated ligation of the 10-mer onto the GMP 5'-end of the RNA by T4 DNA ligase. The PEG linker is long enough (44 units ± 6) that on average the diene has sufficient freedom to find an active site along the length of the catalyst. The ratio of modified transcript of RNA to 10-mer-PEG-substrate to 20-mer DNA bridge oligonucleotide was 1:2:3. These components were mixed thoroughly and incubated at 72°C for 3 min and then allowed to cool to room temperature. Ligation buffer with T4 DNA ligase and RNasin were added and the mixture was incubated in thermocycler at 37°C for 3 hours, at 22°C for 3 hours, 17°C for 3 hours and held at 4°C overnight.

In step 4 (Scheme 1), biotinylated maleimide (final concentration 200 μM) as the Diels-Alder dienophile was added to the ligation product. The reaction mixture contained HEPES pH 7.0, 200 mM NaCl, 200 mM KCl, 1mM $CaCl_2$, 1 mM $MgCl_2$, 10 uM each

$Al(lactate)_3$, $GaSO_4$, $MnCl_2$, $FeCl_2$, $CoCl_2$, $NiCl_2$, $CuCl_2$ and $ZnCl_2$, 10% ethanol (EtOH) and 2% dimethyl sulfoxide (DMSO). The mixture was incubated at 25°C for 3 hours. One major drawback was the low solubility of the biotinylated substrates.

Streptavidin bound biotin in the biotin-product-PEG-DNA-RNA ligation product to produce molecules with large molecular weight. Through running electrophoretic gel, streptavidin-biotin-product-PEG-DNA-RNA with large molecular weight moved slowest in the gel to produce a gel shift. The gel shift was the reacted diene-PEG-DNA-RNA sequences in step 5 (Scheme 1). After removal of all nonbiotinylated members of the pool, in step 6 (Scheme 1) active mutagenized RNA molecules were then reverse transcribed in a mixture of buffer components, dNTP's, 3' primer, RNasin and superscript II reverse transcriptase by incubating at 45°C for 45 min. to give a cDNA copy of the "winning" sequences. These sequences were amplified by performing multiple cycles of mutagenic PCR for further evolution and then used in iterative selection cycles.

Each mutagenic evolution cycle expanded the pool of RNA sequences. It was unclear if the mutated RNA library was enriched in sequences that could catalyze the Diels-Alder reaction. A large population of sequence variants indicates that there are many unique sequences with comparable ability to catalyze this single Diels-Alder reaction. The mutation of a RNA Diels-Alderase can generate numerous related catalysts for a given [4+2] cycloaddition. The majority of the selected mutagenized ssRNA sequences contains a special folded three dimensional structure domain bound with crucial metals and produces intricate binding pockets that specifically recognize diene

and maleimide and catalyze the carbon-carbon bond formation. The number and potential structures possible to bind small molecules or catalyze reactions are unknown. An important goal of my research was to determine if including mutagenesis as part of Diels-Alderase selection would allow better catalysts to be found.

4. RESULTS AND DISCUSSIONS

4.1 DA-2 mutation library generation and *in vitro* selection

The beginning mutated DNA sequence library (~10^{14}) was prepared by mutagenic PCR of clone DNA DA-2 with 157 base pairs. A small portion of the purified products on a 6% analytical polyacrylamide (19:1) gel was stained with ethidium bromide to confirm a satisfactory yield of full-length material, 157 base pair mutated DA-2 as shown in figure 4.1.

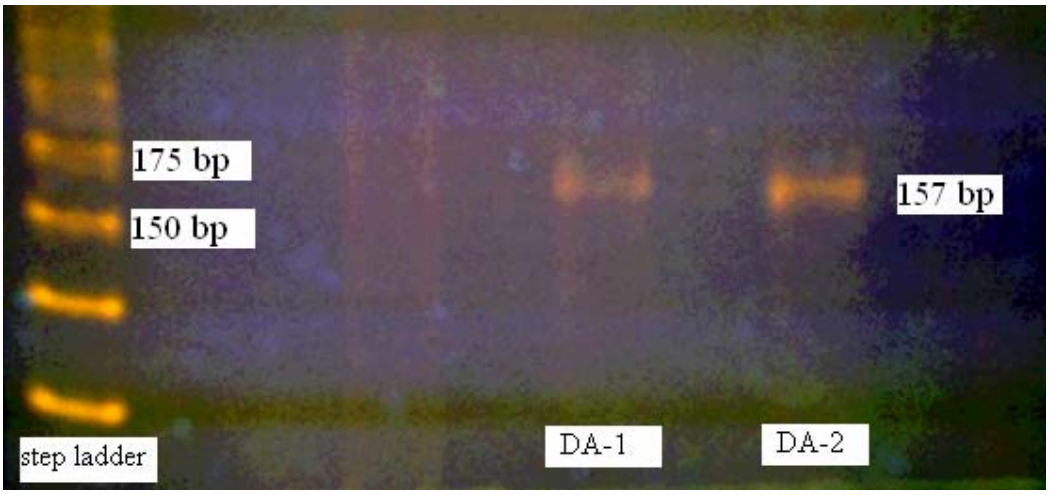


Figure 4.1. 6% analytical polyacrylamide gel for 30 cycles of mutagenic PCR products. 157 base pairs are the right full length of mutagenized DNA products.

The mutated RNA sequence library was then prepared by transcription of the cDNA library generated by mutagenic PCR. A diene was attached to the transcribed RNA through a long flexible polyethylene glycol (PEG) and a DNA 10-mer by bridge-mediated ligation of the 10-mer onto the GMP 5'-end of the RNA by T4 DNA ligase to form the ligation product. Electrophoresis was used to separate ligation products. Voltage was applied to charged molecules, inducing them to migrate through the gel, causing the bigger molecules to move more slowly. The ligation products have a larger molecular weight to charge ratio, so they migrate slower, as seen as the upper bands shown in figure 4.2.

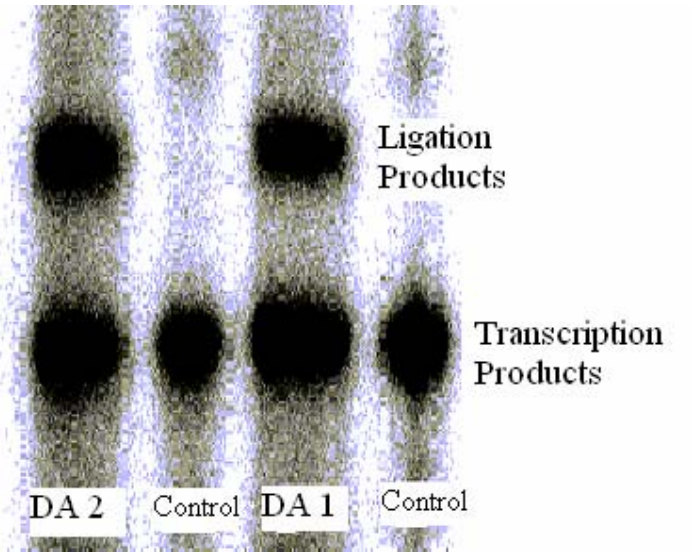


Figure 4.2. 6% denaturing polyacrylamide, 8.3 M Urea gel for ligation. The lower mobility bands correspond to the ligation products, which were excised for Diels-Alder reaction and the higher mobility bands were unligated transcription products.

Catalyzed Diels-Alder cycloaddition products were isolated from the gel shifts, which were streptavidin-bound with biotin in the dienophile, as shown in figure 4.3.

Cycles of *in vitro* selection were continued as described in Scheme 3.1.

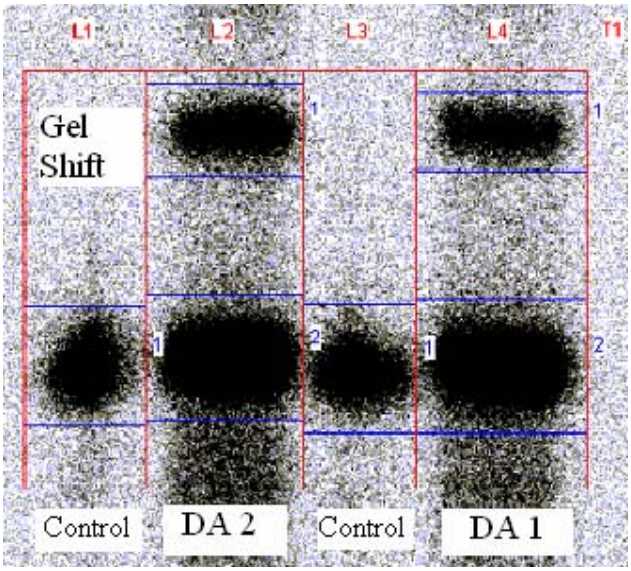


Figure 4.3. 6% denaturing polyacrylamide, 8.3 M Urea gel for Diels-Alder cycloaddion product analysis. The lower mobility bands, gel shifts correspond to the Diels-Alder products, which were excised and brought forward to next cycle. The higher mobility bands were ligation products, which were also the unreacted Diels-Alder products.

4.2 Mutated Diels-Alderase catalytic characterization

Since the sequence of DA-2 was diversified by mutagenic PCR and selected by each evolution cycle, new mutated sequences were introduced in each cycle and the concentration of these different sequences was relatively low in Diels-Alder reaction.

Also, one mutated RNA sequence might have much better or even worse catalytic possibility. Therefore the percentage of reacted Diels-Alder product was getting lower after each cycle of selection shown in table 4.1, showing that sequence diversity increased more than catalytic activity. However, after giving up mutagenic PCR and doing several standard PCR evolution selections, those selected catalytic RNA sequences would converge and we might be able to isolate better Diels-Alderases.

Table 4.1. Progress of the 5 rounds of selection for mutagenized RNA Diels-Alderase from DA-2

Round	Reaction time (min.)	Percentage of streptavidin-bound product per input pmol of RNA (%/pmol RNA)
1	180	1.53
2	180	0.35
3	180	0.78
4	180	0.23
5	60	0.036

4.3 Metal Dependence

The original *in vitro* selection that produced the parent DA-2 included a mixture of metal ions (Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Al^{3+}, Ga^{2+}). This provides the pyridyl modified RNA the opportunity to utilize any metal or metal combination in structural or functional roles. It was found that the position of the nitrogen in the pyridine modification

(*UTP in Scheme 1) was essential for catalytic activity and it suggested that the organization of certain pyridine groups to form a specific copper ion-binding site contributes to the observed rate enhancement through Lewis acid catalysis.

The 5-position pyridyl modified uridine in RNA is crucial to Diels-Alderases because it expands the functional diversity of the RNA by providing additional hydrogen bonding, hydrophobic and dipolar interactions as well as an alternative sigma-donor ligand for metal ion binding.

4.4 Significant effect of RNA catalysis

The recent explosion of genomic and proteomic information has created an abundance of protein targets that have potential as therapeutics. Expanding our understanding of RNA catalyzed cycloadditions will provide the ability to rapidly create and screen a large library of small molecules against protein targets. With a broader understanding of RNA catalysis and the ability to create chemical diversity through cycloaddition reactions it could be possible to assemble libraries of compounds to be targeted to protein active sites that might lead to new structures for pharmaceuticals and agrochemicals.

5. RECOMMENDATIONS FOR FUTURE RESEARCH

5.1 RNA Diels-Alderase sequence analysis

At the end of a selection experiment, the individual “winning” sequences can be spatially isolated from the evolved pool by standard cloning procedures and the

concentrations of the diversified new catalytic RNA Diels-Alderase could be brought up. To isolate the individual sequences of the PCR product from step 6 in Scheme 1, the last cycle will be used in standard plasmid cloning procedures that allows for the isolation of pure bacteria colonies containing individual catalyst sequences. The amplified DNA pool from the cycle of selection can be treated as a "pure" solution of insert and cloned in a variety of ways most typically by blunt-ended ligation into any variety of plasmid vectors. Bacteria can be transformed, plated, isolated and plasmid-purified using standard protocols or high-throughout facilities. Sequencing will be performed on the isolated DNA from the colonies. This can be done by dideoxy cycle sequencing with thermosequenase and a ^{32}P end labeled primer. Sequencing of pools recovered as PCR product from cycles of selection can be compared to the starting DNA pool. This would provide information about the convergence of sequence information in the evolving pool. The sequencing would provide an indication of how the selection preceded.

5.2 Mutagenic evolution of DA-22

The best-studied RNA Diels-Alderase DA-22 will be used under the same condition and compared to the catalytic ability of mutated DA-2.

5.3 Further modifications of RNA

There remain many further modifications of RNA that could be tested to determine what functionality will impart the greatest catalytic activity. For cycloadditions, we currently use 4-pyridyl modification on the U-bases to help form Lewis acid metal coordination sites. We have the potential to modify U-bases or C-bases with new

chemical groups or modify both the U-bases and C-bases in RNA simultaneously. Further modification could be done to render RNA with better catalysis.

5.4 Substrate Specificity

Previously reported Diels-Alderases had been shown to posses a high degree of substrate specificity even when presented with dienophile substrates that were structurally and electronically very similar to the biotinylated maleimide dienophile used in its selection.[49] Substrate specificity could be performed on mutated RNA sequences to detect whether or not the mutation would sacrifice substrate molecular recognition.

5.5 New metal dependence

The mutated Diels-Alderases were incubated with the same mixture of metal ions that was used in the original selection to allow for new metal dependent activity to emerge. All the generations of Diels-Alderases showed a dependence on cupric ion, but with the enrichment of RNA Diels-Alderases, we can predict that new transition metals would be required because the expanded sequences would fold into a different binding site, which fits some other new transition metals. The unique metal dependence of Diels-Alderases would demonstrate that RNA sequences with new structural and/or mechanistic features could be generated through mutation of existing RNA catalysts.

6. LIST OF REFERENCES

1. Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. *Cell* **1982**, 31, 147-157.
2. Peebles, C. L.; Perlman, P. S.; Mecklenburg, K. L.; Pertrillo, M. L.; Tabor, J. H.; Jarrell, K. A.; Cheng, H. L. *Cell* **1986**, 44, 213-223.
3. Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. *Cell* **1983**, 35, 849-857.
4. Prody, G. A.; Bakow, J. T.; Buzayan, J. M.; Schneider, I. R.; Bruening, G. *Science* **1986**, 231, 1577-1580.
5. Buzayan, J. M.; Gerlach, W. L.; Bruening, G. *Nature* **1986**, 323, 349-353.
6. Sharmeen, L.; Kuo, M. Y. P.; Dinter-Gottieb, G.; Taylor, J. J. *Virol.* **1988**, 62, 2674-2679.
7. Saville, B. J.; Collins, R. A. *Cell* **1990**, 61, 685-696.
8. Breaker R. R. *Chem. Rev.* **1997**, 97, 371-390.
9. Tarasow, T. M.; Eaton, B. E. *Biopolymers* **1998**, 48, 29-37.
10. Dewey, T. M. *J. Am. Chem. Soc*. **1995**, 117, 8474-75.
11. Eaton, B. E. *Current Opinion in Chem. Bio*. **1997**, 1, 10-16.
12. Vaish, N. K. *Nucleic Acids Research* **2000**, 28, 3316-22.
13. Lutz, S.; Burgstaller, P.; Benner, S. A. *Nucleic Acids Research* **1999**, 27, 2792-98.
14. Jurczyk, S. C. *Helvetica Chimica Acta* **1999**, 82, 1005-15.
15. Benner, S. A. *Abs. Papers Am. Chem. Soc*. **1998**, 216, U264.
16. Dewey, T. M.; Zyzniewski, M. C.; Eaton, B. E. *Nucleosides & Nucleotides* **1996**, 15, 1611-17.
17. Eaton, B. E.; Pieken, W. A. *Annual Rev. Biochem*. **1995**, 64, 837-63.
18. Gourlain, T. *Nucleic Acids Research* **2001**, 29, 1898-905.
19. Aurup, H.; Williams, D. M.; Eckstein, F. *Biochem*. **1992**, 31, 9636-41.
20. Perrin, D. M.; Garestier, T.; Helene, C. *Nucleosides & Nucleotides* **1999**, 18, 377-91.
21. Thum, O.; Jager, S.; Famulok, M. *Angewandte Chemie-Intl. Ed*. **2001**, 40, 3990.
22. Teramoto, N.; Imanishi, Y. I. Y. *Bioconjugate Chem*. **2000**, 11, 744-48.
23. Matulic-Adamic, J.; Daniher, A. T.; Karpeisky, A. *Bioorg. Med. Chem. Letters* **2000**, 10, 1299-302.
24. Sakthivel, K.; Barbas, C. F. *Angewandte Chemie-Intl. Ed*. **1998**, 37, 2872-73.
25. Seelig, B. *Angewandte Chemie-Intl. Ed*. **2000**, 39, 4576.
26. Sengle, G. *Bioorg. Med. Chem*. **2000**, 8, 1317-29.
27. Stuhlmann, F.; Jaschke, A. *J. Am. Chem. Soc*. **2002**, 124, 3238-44.
28. Suga, H., P. A. Lohse; Szostak, J. W. *J. Am. Chem. Soc*. **1998**, 120, 1151-56.
29. Lohse, P. A.; Szostak, J. W. *Nature* **1996**, 381, 442-44.
30. Jadhav, V. R.; Yarus, M. *Biochem*. **2002**, 41, 723-29.
31. Illangasekare, M.; Yakus, M. *Science* **1995**, 267, 643-47.

32. Illangasekare, M.; Yakus, M. *J. Mol. Bio*. **1997**, 274, 519-29.
33. Illangasekare, M.; Yakus, M. *J. Mol. Bio*. **1997**, 268, 631-39.
34. Illangasekare, M.; Yakus, M. *Rna-A Publication of the Rna Soc*. **1999**, 5, 1482-89.
35. Illangasekare, M.; Yakus, M. *Proceedings of the National Academy of Sciences of U.S.A.* **1999**, 96, 5470-75.
36. Lee, N. *Nature Structure Biology* **2000**, 7, 28-33.
37. Lee, S. E. *Nucleic Acids Research* **2001**, 29, 1565-73.
38. Kumar, R. K.; Yakus, M. *Bio.* **2001**, 40, 6998-7004.
39. Zhang, B. L.; Cech, T. R. *Chem. Bio*. **1998**, 5, 539-53.
40. Zhang, B. L.; Cech, T. R. *Nature* **1997**, 390, 96-100.
41. Wiegand, T. W.; Janssen, R. C.; Eaton, B. E. *Chem. Bio*. **1997**, 4, 675-83.
42. Huang, F. Q.; Bugg, C. W.; Yarus, M. *Biochemistry* **2000**, 39, 15548-55.
43. Sengle, G. *Chem. Bio*. **2001**, 8, 459-73.
44. Wecker, M.; Smith, D.; Gold, L. *Rna-A Publication of the Rna Soc*. **1996**, 2, 982-94.
45. Li, Y. F.; Sen, D. *Nature Structural Biology* **1996**, 3, 743-47.
46. Li, Y. F.; Sen, D. *Biochem*. **1997**, 36, 5589-99.
47. Tarasow, T. M.; Eaton, B. E. *Nature* **1997**, 389, 54-57.
48. Tarasow, T. M.; Eaton, B. E. *J. Am. Chem. Soc*. **2000**, 122, 1015-21.
49. Tarasow, T. M.; Eaton, B. E. *J. Am. Chem. Soc*. **1999**, 121, 3614-17.
50. Seelig, B.; Jaschke, A. *Chem. Bio*. **1999**, 6, 167-76.
51. Corey, E. J.; Cheng, X. M. *Logic Chem. Syn. Wiley: NY*, **1989**, 18-23.
52. Corey, E. J. *Angewandte Chemie-Intl. Ed*. **2002**, 41, 1650-1667.
53. Nicolaou, K. C.; Snyder, S. A.; Montagnon, T. *Angewandte Chemie-Intl. Ed*. **2002**, 41, 1668-1698.
54. Nieuwlandt, D.; West, M. Chem. *Biochem*. **2003**, 4, 649-662.
55. Cha, R. S.; Thilly, W. G. *PCR Methods Applic*. **1993**, 3, S18-S29.
56. Matteucci, M. D.; Heyncker, H. L. *Nucleic Acids Res*. **1983**, 11, 3113-3121.
57. Wells, J. A.; Vasser, M.; Powers, D. B. *Gene* **1985**, 34, 315-323.
58. Oliphant, A. R.; Nussbaum, A. L; Struhl, K. *Gene* **1986**, 44, 177-183.
59. Ner, S. S.; Goodin, D. B.; Smith, M. *DNA* **1988**, 7, 127-134.
60. Beaudry, A. A.; Joyce, G. F. *Science* **1992**, 257, 635-641
61. Lehman, N.; Joyce, G. F. *Nature* **1993** 361, 182-185.
62. Bartel, D. P.; Szostak, J. W. *Science* **1993**, 261, 1411-1417.
63. Beckman, R. A.; Mildvan, A. S. Loeb, I. A. *Biochem*. **1985**, 24, 5810-5817.
64. Leung, D. W.; Chen, E.; Goeddel, D. V. *Technique* **1989**, 1, 11-15.
65. Eckert, K. A.; Kunkel, T. A. *PCR Methods Applic.* **1991**, 1, 17-24.
66. Ling, L. L.; Keohavong, P.; Dias, C. *PCR Methods Applic.* **1991**, 1, 63-69.
67. Cadwell, C.; Joyce, G. F. *PCR Methods Applic.* **1992**, 2, 28-33.
68. Gelfand, D. H.; White, T. J. *PCR Prot*. **1990**, 129-141.