RNA Diels−Alderases: Relationships between Unique Sequences and Catalytic Function

Theodore M. Tarasow, Sandra L. Tarasow, and Bruce E. Eaton*

Contribution from the Chemistry Division, NeXstar Pharmaceuticals, Inc., 2860 Wilderness Place, Boulder, Colorado 80301

Received July 12, 1999

Abstract: A group of eight unique RNA sequences isolated from a single in vitro selection experiment were characterized based on their Diels−Alderase properties. Five of the eight isolates contain a 10-base conserved region, while the remaining three do not share any sequence similarities with any of the others. This series of RNA Diels−Alderases displayed k_{cat} and K_m values that spanned 74 and 80-fold ranges, respectively. These values appeared to compensate for one another, as the value of k_{cat}/K_m varied by only 5-fold among the isolates studied. Product inhibition experiments revealed a general trend between affinity for the cycloaddition product and catalytic ability as measured by k_{cat}. The activities of each of the RNA Diels−Alderases were completely dependent on the presence of cupric ion. In addition, all of the RNA catalysts demonstrated a high degree of substrate selectivity, even for structurally and electronically similar dienophiles. Overall, the data demonstrate how a variety of RNA sequences can solve a specific chemical problem. The information gathered provides a foundation from which to build our understanding of how sequence, structure, and function are interrelated. Such knowledge could be applied to create new catalysts, both large and small.

Introduction

A new era is emerging for the biocatalysis of small molecules.[1] RNA catalysts are now demonstrating new potential for valued synthetic transformations. The ability to systematically modify vast RNA libraries with diverse functionality known to be important in catalysis has expanded the functional capabilities of RNA.[2] Unlike any other molecule type, RNA can be enzymatically replicated, and this provides the opportunity to directly select for a desired catalytic function via in vitro techniques. To fully realize the potential of RNA as a catalytic platform[3] and understand its place relative to other biocatalysts, more must be learned about the details of how small molecule substrates are brought together within RNA active sites. Herein we describe for the first time how different RNA Diels−Alderases catalyze carbon−carbon bond formation to assemble cyclohexene carbocycles.

Given the importance of carbocycles as pharmaceuticals, we have focused on the Diels−Alder cycloaddition reaction. Historically, the Diels−Alder reaction has been one of the most valuable and widely used transformations available to organic chemists.[4] This [4 + 2] cycloaddition between a 1,3 diene and an electron-deficient alkene has been proven to be a powerful transformation for the assembly of complex natural products and drug substances.[5] The reaction creates two carbon−carbon bonds and up to four new stereocenters and can typically be accelerated by an aqueous environment or by Lewis acid catalysis. This latter fact makes catalysis by RNA a daunting proposition since the rate enhancements in going from an organic to an aqueous medium are already significant ($>10^2$).[6] Nevertheless, we recently described preliminary results of an in vitro selection for RNA molecules capable of catalyzing a Diels−Alder cycloaddition.[7] These RNA Diels−Alderases are the first examples of RNA facilitating carbon−carbon bond formation.

The in vitro selection strategy shown in Scheme 1, panel A was used to isolate Diels−Alderase RNA catalysts from a starting pool of more than 10^{14} unique sequences.[7a] The RNA molecules were constructed of a randomized region 100 nucleotides in length flanked by regions of constant sequence to allow for amplification and other enzyme-mediated steps in the selection process. Previous studies suggested that unmodified RNA had no inherent Diels−Alderase activity,[8] although a recent publication using a very hydrophobic diene capable of intercalation demonstrated that native RNA could be shown to catalyze a [4 + 2] cycloaddition with a maleimide dienophile.[9] Increasing the functional diversity and conformational flexibility through nucleotide modifications can create RNA that has the potential to be a more versatile and efficient catalytic platform. Nature has chosen to modify almost all but the most benign forms of RNA.[10] Indeed, in contemporary biochemistry the function of RNA is typically dependent on the type and position

(1) (a) Joyce, G. F. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5845−5847 (b) Frauendorf, C.; Jaschke, A. *Angew. Chem., Int. Ed.* **1998**, *37*, 1378−1381.

(2) Tarasow, T. M.; Eaton, B. E. *Biopolymers* **1998**, *48*, 29−37.

(3) Tarasow, T. M.; Eaton, B. E. *CMLS* **1999**, *55*, 1463−1472.

(4) Corey E. J.; Cheng X- M. *The Logic of Chemical Synthesis*; Wiley: New York, 1989; pp 18−23.

(5) For some current examples, see: (a) Roush, W. R.; Sciotti, R. J. *J. Am. Chem. Soc.* **1998**, *120*, 7411−7419. (b) Corey, E. J.; Letavic, M. A. *J. Am. Chem. Soc.* **1995**, *117*, 9616−9617. (c) Corey, E. J.; Guzman-Perez, A.; Loh, T- P. *J. Am. Chem. Soc.* **1994**, *116*, 3611−3612. (d) Kozmin, S. A.; Rawal, V. H. *J. Am. Chem. Soc.* **1997**, *119*, 7165−7166. (e) Fallis, A. G. *Acc. Chem. Res.* **1999**, *32*, 464−474.
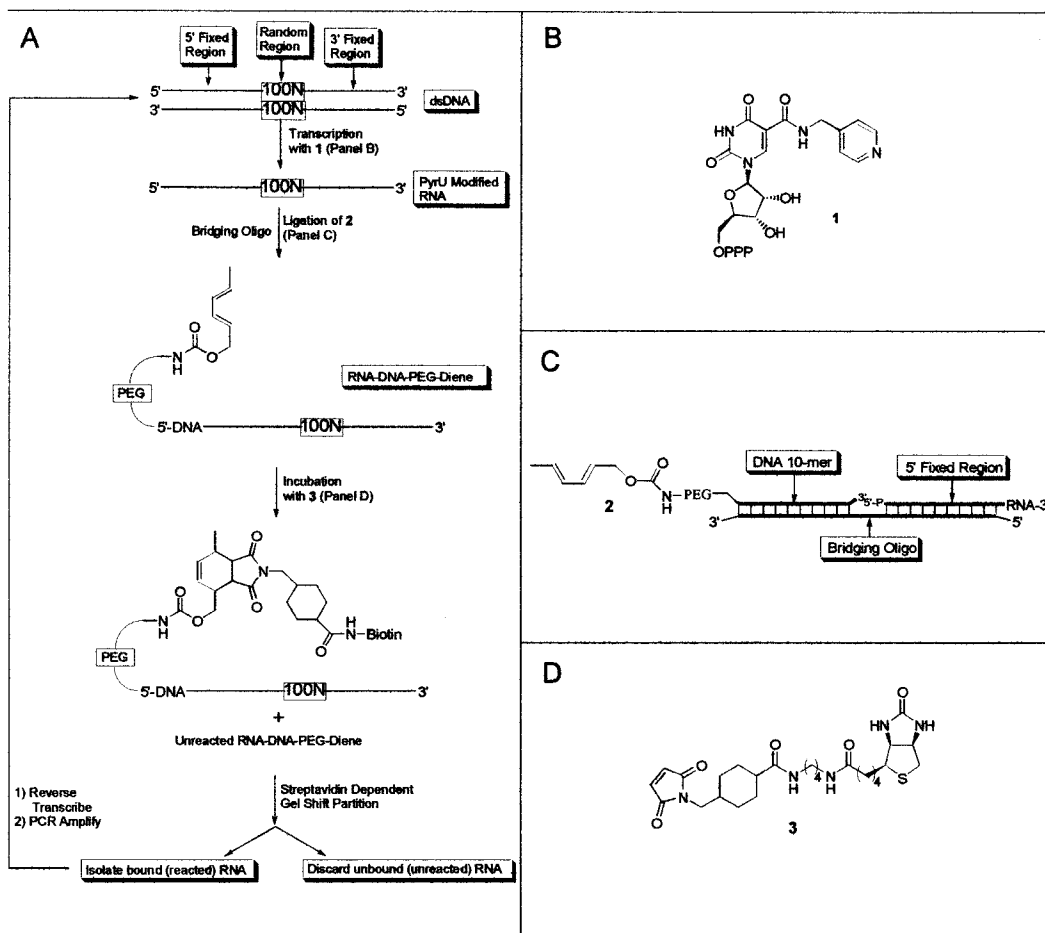
(6) Garner, P. P. In *Organic Synthesis in Water, Diels−Alder Reactions in Aqueous Media*; Grieco, P. A., Ed.; Thomson Science: New York, 1998; pp 1−42.

(7) (a) Tarasow, T. M.; Tarasow, S. L.; Eaton, B. E. *Nature* **1997**, *389*, 54−57. (b) Tarasow, T. M.; Tarasow, S. L.; Tu, C.; Kellogg, E.; Eaton, B. E. *J. Am. Chem. Soc.* **1999**, *121*, 3614−3617.

(8) Morris, K. N. et al. *Proc. Natl. Acad. Sci. U.S.A.* **1994** *91*, 14028−13032.

(9) Seelig, B.; Jaschke A. *Chem. Biol.* **1999**, *6*, 167−176.

Scheme 1

of the modification.[11] Many of these modifications occur at the 5-position of pyrimidine bases. Similarly, the RNA utilized in this Diels−Alderase selection contained 5-pyridylmethylcarboxy amide-U (**1**) instead of uridine, which is just one of a wide array of modified uridines[2,12] and deoxyuridines[13] known to be compatible with in vitro selection methods. The pyridyl-modified uridine **1** expands the functional diversity of the RNA by providing additional hydrogen bonding, hydrophobic, and dipolar interactions. Furthermore, the pyridyl functionality of **1** provides the opportunity for the formation of unique metal binding sites which are likely important for generating novel two- and three-dimensional structures as well as having the potential to act as ligands to coordinate metals that can act as a Lewis acid.

RNA catalyst selection was accomplished by covalent attachment of the diene and RNA via a poly(ethylene glycol) (PEG) linker and a DNA 10-mer by bridge-mediated ligation of the 10-mer (**2**) onto the 5′-end of the RNA (Scheme 1, panel C). Diels−Alder cycloaddition between the RNA−diene con-

jugate and dienophile **3** results in biotinylation of the RNA. The long flexible PEG linker was employed to allow the diene to survey the entire surface of the RNA. Cycles of in vitro selection were carried out by incubating the RNA−diene conjugate with the dienophile **3** shown in Scheme 1, panel D. A mixture of transition metals were included in the incubation which could form unique structural or Lewis acidic sites within the modified RNA. The biotinylated maleimide dienophile allowed for separation of reacted and unreacted RNA molecules by virtue of streptavidin binding and concomitant electrophoresis mobility shift. Partitioned RNA molecules were subsequently reverse transcribed while still bound to streptavidin, the resulting cDNA was PCR-amplified, and the next cycle of RNA produced by T7 RNA polymerase mediated transcription. Multiple cycles of in vitro selection produced a population of RNA molecules capable of catalyzing the desired Diels−Alder reaction. Multiple families of nonclonally derived catalysts were obtained, and their analysis is the subject of this publication.

Results and Discussion

In vitro RNA Diels−Alderase Selection. The results from 12 cycles of selection for Diels−Alder catalysts are shown in Table 1. RNA and dienophile concentrations were held constant throughout the selection. Consistent values for what was essentially the uncatalyzed reaction for cycles 1−7 demonstrate the reproducibility of the gel-partitioning technique and provided a baseline for detection of the slightest increase in percent bound, which became apparent in round 8. As a means of more stringent selection, incubation times were steadily decreased from 8 h in round 8 to 3 min in round 12. The pool's increased ability to

(10) (a) Limbach, P. A.; Crain, P. F.; McCloskey, J. A. *Nucleic Acids Res.* **1994**, *22*, 2183−2196. (b) Rozenski, J. Crain, P. F.; McCloskey, J. A. *Nucleic Acids Res.* **1999**, *27*, 196−197.

(11) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984; pp 178−189.

(12) (a) Dewey, T. M.; Mundt, A. A.; Crouch, G. J.; Zyzniewski, M. C.; Eaton, B. E. *J. Am. Chem. Soc.* **1995**, *117*, 8474−8475. (b) Kerr, C. E. Ph.D. Thesis, Washington State University, December, 1997. (c) Dewey, T. M.; Zyzniewski, C.; Eaton, B. E. *Nucleosides Nucleotides* **1996**, *15*, 1611−1617. (d) Latham, J. A.; Johnson, R.; Toole, J. J. *Nucleic Acids Res.* **1994**, *22*, 2817−2822.

(13) Sakthivel, K.; Barbas, C. F., III. *Angew. Chem., Int. Ed.* **1998**, *37*, 2872−2875.

Table 1. Progress of the 12 Rounds of Selection for RNA Diels−Alderases

round	input RNA (pmol)	reaction time (min)	streptavidin bound (%)	fraction reacted/time $(\text{min}^{-1}) \times 10^3$
1	2000	480	2.4	0.05
2	750	480	2.6	0.05
3	200	480	2.6	0.05
4	100	480	2.6	0.05
5	100	480	2.6	0.05
6	100	480	2.6	0.05
7	100	480	2.7	0.06
8	100	480	3.6	0.08
9	100	120	6.7	0.56
10	100	30	8.6	2.9
11	100	10	4.9	4.9
12	100	3	3.3	11

facilitate the cycloaddition can be illustrated by comparing the fraction of RNA reacted as a function of time over the course of the selection (Table 1). The data indicate that the Diels−Alder reaction rate was approximately 200-fold faster in the presence of the round 12 pool.

Sequence Analysis. Following round 12, standard cloning and sequencing procedures produced 46 sequences from which eight nonclonally derived families were identified.[14] A curious feature of the largest clonal family was a stretch of G residues up to 10 consecutive nucleotides in length located at the 3′-end of the random region. Clonal family base composition analysis revealed no clear bias for or against incorporation of the modified U (26.5%), consistent with studies of modified base incorporation and recognition by the transcription enzymes utilized.[12b] A 10 base conserved sequence was found in five of the eight unique sequences.

Proposed secondary structures for representative clones based on the Zucker folding algorithm are shown in Figure 1.[15] It should be noted that these structures are speculative since the algorithm does not take into account the effect of the 5-position uridine modifications. Nevertheless, it is evident from the structures that these RNA Diels−Alderases have the potential to generate extended helical regions resulting in stable secondary structures. Many suboptimal folds can be generated with energies within 90% of the optimal fold, which may explain why less than 100% of a given sequence is active. Not surprisingly, large portions of the secondary structures involve the constant regions, which were required for the various enzymatic steps of the in vitro selection cycle. It is interesting to speculate that these catalysts were forced to use the constant regions and as a result are less than optimal. Perhaps more efficient catalysts could be selected from RNA libraries constructed with smaller or even deleted constant regions.
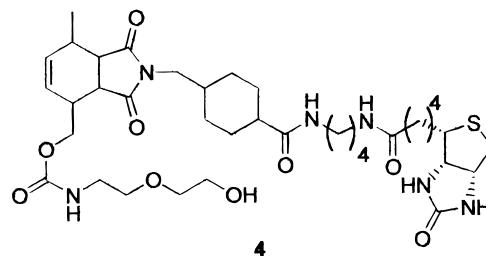
Additional analysis of the lowest-energy structures generated by the Zucker algorithm showed the locations and secondary structures of the 10 base conserved region found in five of the eight sequence families. These regions, largely centralized in the primary sequence, are involved in a variety of secondary interactions, some of which are entirely within the constant region (isolate 2) or entirely within the random region (isolates 17 and 10) or with both constant and random regions (isolates 22, 1, and 7). The 10 base conserved regions have been expanded in Figure 1 to show that they may share similar secondary structures consisting of a bulge containing at least

the two adenosine nucleotides flanked by helical regions. Indeed, inspection of low-energy, suboptimal structures (>95% of the minimum energy) revealed that the isolates containing the conserved sequence can fold in a manner that reduces the bulge to just the two adenosine residues as seen in DA 1/22 (structures not shown). Whether these regions actually adopt the proposed secondary structures, based on the sequence conservation, it is likely that these conserved regions do play similar roles in creating catalytically active folded topologies.

Diels−Alderase Kinetics. Previously, it was reported that representatives from each of the sequence families and the three orphan sequences were catalytically active.[7a] This is somewhat surprising, given the results of other in vitro selections for RNA catalysts where only a fraction of the isolated sequences were catalytically competent or single sequences had unusual activity and were reasoned to be a very rare occurrence in a randomized pool. With a host of new RNA Diels−Alderases to choose from, representatives from each nonclonally derived family (DA 1, 2, 7, 10, 15, 17, 22, 24, and 37) were chosen for further study. The insolubility of dienophile **3** precluded kinetic evaluation at substrate saturation of many of the isolates studied. Nevertheless, observed rate constants (k_{obs}) were determined for each sequence over a concentration range up to 500 μM of **3**. It should be noted that these and the other kinetic results reported in this manuscript were obtained under single turnover conditions. Fitting the data to the Michaelis−Menton equation, the kinetic constants in Table 2 were obtained. The k_{cat} and K_m values for these different RNA catalysts span an approximate 75-fold range, whereas the values for the catalytic efficiency (k_{cat}/K_m) only vary by 5-fold.[16] Sequences appear to have been selected for their ability to facilitate the reaction with an efficiency demanded by the selection conditions, whether it was a faster k_{cat}, a lower K_m, or intermediate values for both. Interestingly, the five isolates that have the 10 base conserved region display the fastest k_{cat} values, while those isolates that lack the conserved region were measured to have the tightest K_m values.

Factors such as reaction time and concentration of **3** may have been critical in determining which RNA catalysts emerged from the vast starting library. However, other more conspicuous variables could have also contributed significantly such as the length of the PEG linker, the incubation temperature, and the reaction medium (metal ion and organic solvent concentrations). Evaluating and understanding selection pressures such as these will help guide future experiments directed toward isolating new RNA catalysts.
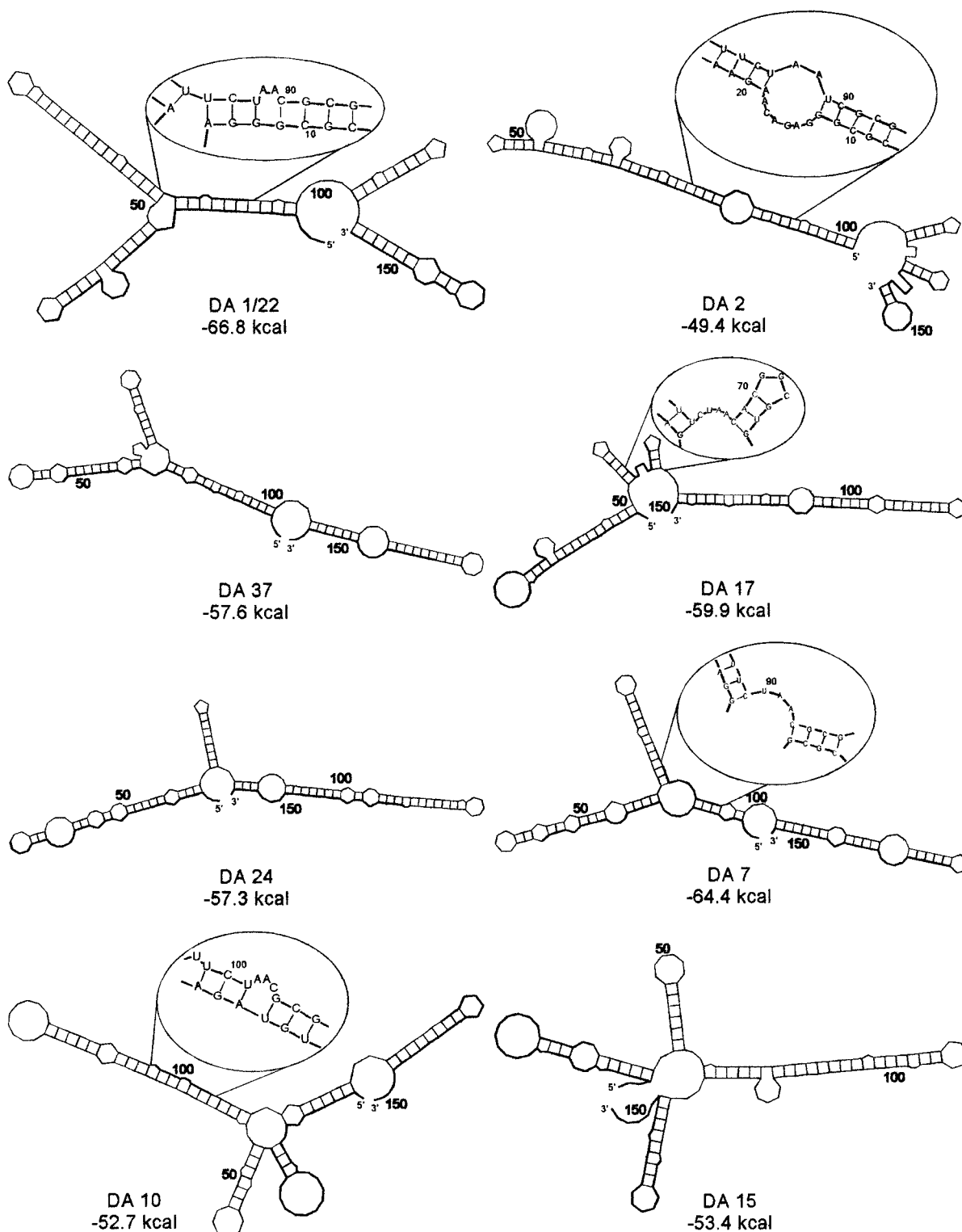
RNA Diels−Alderase Product Inhibition. Product inhibition studies can often be instructive for reactions such as the Diels−Alder cycloaddition because they typically proceed via late transition-states, thus providing an indication of the recognition elements that comprise the active site. Compound **4** has all of

4

(14) Sequences for the isolates characterized can be found in the Supporting Information or in ref 7a.

(15) Matzura, O., Wennborg, A. *Comput. Appl. Biosci.* **1996**, *12*, 247−249.

(16) The uncatalyzed second-order rate constant for the spontaneous Diels−Alder cycloaddition in the presence of random oligonucleotide sequences was measured to be 3.25×10^{-1} M^{-1} min^{-1}. See Experimental Section and ref 7a for details.

Figure 1. Proposed secondary structures of the nine RNA Diels−Alderase isolates studied. Each structure is labeled with the isolate number and the calculated energy at 25 °C for the fold shown. The 5′ and 3′ constant regions are shown in bold. The 10 base conserved sequences found in five of the eight unique sequences are also shown in bold and have been enlarged to show secondary structural details.

the structural elements of the RNA-conjugated product with a two ethylene glycol unit of the PEG attached through a carbamate linkage. Each of the RNA Diels−Alderases was evaluated for inhibition by **4**. In all cases, product inhibition was observed to varying degrees (Table 2). There appears to be a trend between k_{cat} and K_{iapp} where a faster k_{cat} correlates to a tighter K_{iapp}. This is consistent with the notion that the Diels−Alder transition state is product-like, where a faster k_{cat} results from a tighter recognition of the transition state, and

thus the product would be expected to be a better inhibitor. There are two exceptions to this apparent trend. One example, isolate 37, has a relatively low K_{iapp} but displays an anemic k_{cat}. This disparity may suggest that functionality remote to the reaction center contributes significantly to binding of the product. Alternatively, the structure and dynamics of the active site in 37 could be more closely coupled to ground-state product conformers that do not resemble the cycloaddition transition state.[17] The other anomalous example, isolate 7, has a relatively

Table 2. Kinetic and Inhibition Constants

isolate	$k_{cat} \times 10^2$ (min^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} min^{-1})	K_{iapp} (μM)
22a	66 ± 12	2300 ± 500	287	32.5 ± 2.6
1	42 ± 11	1920 ± 650	219	29.7 ± 6.0
17	24 ± 5	1220 ± 320	197	44.6 ± 7.5
2	23 ± 10	1800 ± 1000	128	59.9 ± 6.1
7	18.9 ± 0.1	1640 ± 20	115	222 ± 29
10	6.4 ± 0.4	830 ± 500	77.1	135 ± 9
37	4.3 ± 0.4	118 ± 30	364	56 ± 17
15	1.7 ± 0.2	69 ± 22	246	108 ± 32
24	0.89 ± 0.03	29 ± 4	307	165 ± 46

^{a} Data for isolate 22 were reported previously in ref 7a.

fast k_{cat} (0.189 min^{-1}) but the highest K_{iapp} (222 μM). It should be noted that the product **4** was prepared by the spontaneous reaction of the cognate diene and dienophile and that all analytical methods used to characterize the molecule indicate that it is exclusively the endo product. Thus, if DA 7 preferentially catalyzes the formation of the exo product, **4** would not be expected to be a particularly good inhibitor. Studies are currently underway to address the stereospecificity of these RNA active sites.

Diels—Alderase Metal Dependence. All of the Diels—Alderase isolates surveyed displayed an absolute dependence on the presence of cupric ion. No activity was observed in the absence of Cu^{2+}. This is remarkable, given the diversity of metal ions included in the selection (Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}) capable of forming structural elements or Lewis acidic centers. As noted previously for isolate 22,[7] this binary dependence on a single metal ion suggests that these Diels—Alderases achieve a significant portion of their rate acceleration via Lewis acid catalysis. Since at least eight unique sequences were identified to have notable catalytic activity, it appears that there are multiple solutions to accelerating the rate of this cycloaddition through the use of cupric ion Lewis acids.[18]

Diels—Alderase Substrate Specificity. Substrate specificity, a crucial feature in any biocatalyst, was used to further characterize these new RNA catalysts. The Diels—Alderase activity for each isolate was determined by substituting either of two alternative dienophiles and one diene (Table 3).[19] For all of the isolates, no significant activity was observed in the presence of either the fumaramidate dienophile **5** or the cyclic diene **7**. These two substrates are clearly structurally distinct from the cognate diene and dienophile used to select these Diels—Alderases, indicating some modest level of molecular discrimination by the different RNA active sites. However, the dienophile **6**, also being a maleimide, shares a considerable amount of structural and functional homology with the original dienophile **3**. Indeed, they are identical out to the carbon β to the maleimide functionality. Despite these close similarities, all of the RNA Diels—Alderases demonstrated a distinct discrimination in favor of the dienophile **3**. This discrimination ranged from approximately 2- to 100-fold, with the majority of catalysts showing 20-fold or greater selectivity. While the magnitude varies somewhat from one isolate to the next, specificity such as this is remarkable, given the fact that no experimental

pressures were included in the selection procedure that demanded substrate discrimination. It appears that these RNA Diels—Alderases possess inherent substrate specificity as a common feature among the entire catalytic population.

Conclusions

For the first time characterization of different RNA Diels—Alderases obtained from a single in vitro selection experiment has been completed. Several interesting features of these catalysts were revealed in the analysis of their sequence phylogeny and Diels—Alderase characteristics. Clonal family base composition analysis clearly showed no bias for or against incorporation of the modified U (26.5%), consistent with studies of modified base incorporation by the enzymes utilized. A 10 base conserved sequence was found in five of the eight unique sequences, while the other three nonclonally derived catalysts did not share any sequence homology. The conserved region apparently represents a relatively abundant and general solution to the problem of catalysis, while there are clearly numerous less obvious sequence variations that determine overall catalytic efficiencies.

Representatives from each of the families as well as the each of the orphan sequences were catalytically active, spanning only a 5-fold range in catalytic efficiency (k_{cat}/K_m). This narrow range was perhaps the result of selection conditions such as reaction time, metal ion composition, or concentration of **3**. Consistent with the notion that the Diels—Alder transition state is product-like, representative Diels—Alderases from each family demonstrated a general trend in product inhibition relative to the kinetic constant k_{cat}.

Perhaps most striking was the high degree of substrate specificity exhibited by these RNA catalysts. Despite the fact that no selection pressures were applied to demand dienophile substrate specificity, all of Diels—Alderases showed significant selectivity, even for dienophiles of comparable spontaneous reactivity. This discovery reveals the degree of molecular recognition possible by suitably modified RNA catalysts, which we are continuing to explore to determine if it is an inherent property of these biopolymers or unique to this Diels—Alder chemistry.

There now exists a fundamental framework for the design of experiments to further unravel the mechanisms by which these pyridyl-modified RNA catalysts operate. Efforts are underway to reveal new RNA catalysts for difficult small molecule transformations of synthetic utility for assembling pharmaceuticals. To date we have selected not only Diels—Alderases but also amide synthases[20] and peptide transfer RNA[21] catalysts. In all cases, the majority of nonclonally derived sequences showed pronounced catalytic activity. Further understanding of the relationship between sequence, structure, and function will provide the groundwork necessary for advancing our ability to design and create novel RNA catalysts on demand. Because RNA is easily mutated, selected, and sequenced, there exists the opportunity to create and access the enormous amount of information necessary to understand these fundamental yet extremely complex relationships. The progress in experimental tools for in vitro selection and spectroscopic investigation bode well for this new era in RNA catalysis.
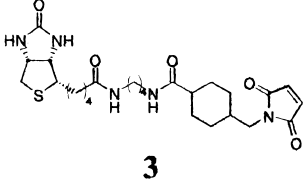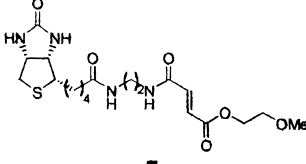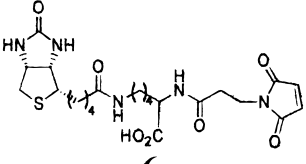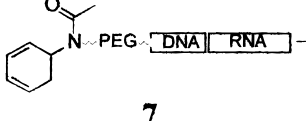
Experimental Section

RNA Library Design and Synthesis. The initial RNA library was designed and synthesized from double-stranded DNA (dsDNA) as

(17) Lighthouse, F. C., Bruice, T. C. *J. Am. Chem. Soc.* **1996**, *118*, 2595–2605.

(18) Despite the fact that there are numerous copper complexes that catalyze the Diels—Alder reaction in organic solvents, only one example of a pyridyl—copper complex catalyzing this cycloaddition in water has been reported. In this case, the pyridine—copper Lewis acid was conjugated to the dienophile. (a) Otto, S.; Bertoncin, F.; Engberts, J. B. F. N. *J. Am. Chem. Soc.* **1996**, *118*, 7702–7707. (b) Otto, S.; Engberts, J. B. F. N. *Tetrahedron Lett.* **1995**, *36*, 2645–2648.

(19) Values for isolate 22 can be found in ref 7b.

(20) Wiegand, T. W.; Janssen, R. C.; Eaton, B. E. *Chem. Biol.* **1997**, *4*, 675–683.

(21) Nieuwlandt, D.; West, M.; Eaton, B. E., unpublished results.

Table 3. RNA Diels−Alderase Substrate Specificity

Substrate	Isolate 1*	Isolate 2*	Isolate 17*	Isolate 24*	Isolate 7*	Isolate 10*	Isolate 15*	Isolate 37*
3	100	100	100	100	100	100	100	100
5	0	0	0	0	0	0	5	0
6	4	3	5	44	1	3	31	9
7	0	0	0	1	0	0	1	0

* Values are background corrected, adjusted to the total amount of active catalyst, and relative to the catalyzed reaction between the acyclic diene **2** and the biotin maleimide **3**. Values of zero are below the assay detection limits (<1).

described previously[20] with the exceptions noted below. A 100 nucleotide random region (100N) was employed with the pyridyl methyl-modified UTP (**1**) substituted for native UTP during T7 transcription, using standard reaction conditions.[20] The resulting 147 nucleotide (nt) RNA library was passed through successive Nap columns (Pharmacia, Uppsala, Sweden), concentrated on a 30 000 molecular weight cutoff (MWCO) filter, washed with water and concentrated (1 × 2 mL), extracted twice with phenol/chloroform/isoamyl alcohol (25:24:1), and extracted once with chloroform, and ethanol was precipitated. The resulting pellet was dissolved in water and quantitated by UV ($\epsilon_{260} = 40\ \mu$g/mL) and specific activity. The DNA 10-mer-PEG−diene was ligated by combining the RNA library (1 nmole), PEG modified DNA 10-mer (2 nmoles),[22] and the bridging oligomer (3 nmoles) in 20 μL of 10× T4 DNA ligase buffer, 16 μL of T4 DNA ligase storage buffer, and water to a final volume of 200 μL. The mixture was heated to 72 °C for 3 min and then allowed to come to room temperature over 10 min. RNAsin (160 units, 39 units/μL, Promega, Madison, WI) and T4 DNA Ligase (80 units, 5units/μL, Boeringer Mannheim, Mannheim) were then added, and the mixture was incubated at 37 °C for 4−16 h. The ligation product was purified on a 6% denaturing polyacrylamide (19:1 cross-linking) and quantitated by UV and specific activity.

Incubation Conditions. All RNA incubations were conducted under the following conditions except as noted: 50 mM HEPES pH 7.0; 500 nM pyrU RNA; 200 mM NaCl; 200 mM KCl; 1 mM CaCl$_2$; 1 mM MgCl$_2$; 10 μM each of Al(lactate)$_3$, Ga$_2$(SO$_4$)$_2$, MnCl$_2$, FeCl$_2$, CoCl$_2$, NiCl$_2$, CuCl$_2$, and ZnCl$_2$; 10% ethanol (EtOH); and 2% dimethyl sulfoxide (DMSO). The dienophile **3** was added as a stock solution in DMSO such that the **3** aliquot always accounted for the entire 2% DMSO required. The concentration of **3** varied in the isolate characterization experiments but was held constant at 100 μM throughout the selection. RNA, HEPES, monovalent metals, CaCl$_2$, and water were combined and heated to 72 °C for 6 min after which time the remaining metal ions were added, and the solution was allowed to cool slowly to ambient temperature over 25 min. The EtOH and dienophile in DMSO were then added, and the reaction mixture was incubated at 25 °C for the time periods indicated. Incubations were terminated by the addition of β-mercaptoethanol (BME) to a final concentration of 5 mM or passing by the solution over two successive Nap columns to remove excess **3**.

Reaction Assay and Partitioning. The extent of reaction and partitioning of reacted and unreacted RNA molecules was accomplished using a streptavidin (SA)-dependent gel shift. The amount of SA added to the RNA sample varied according to how much **3** was used in the incubation and whether Nap columns had been used following the incubation. If Nap columns (two) were employed, then SA was added such that the ratio of **3** in the incubation to added SA was 20. If Nap columns were not utilized, the ratio of SA to **3** was 4:1. The RNA was incubated with the SA for 20 min at ambient temperature, followed by the addition of 0.5 vol of formamide-loading buffer. The sample was then heated at 72 °C for 3 min, loaded on a preheated (∼50 °C) 6% denaturing polyacrylamide gel, and eletrophoresed. The shifted and unshifted bands were visualized by autoradiography and phosphorimaging, the later being used for quantitation. For partitioning, shifted bands were excised, and the RNA−SA complex was extracted, desalted, and subjected to reverse transcription and PCR amplification according to previously published procedures.

Cloning and Sequencing. Cleavage sites for the restriction enzymes Hind III and BamH I were incorporated into the fixed sequences during PCR amplification of the cDNA pool following round 12. The resulting double-stranded DNA was inserted into the plasmid puc9, transformed into DH5α cells via electroporation, and plasmid DNA isolated using a PERFECT prep plasmid DNA kit (5′-3′ Inc., Boulder, CO). The purified DNA was bidirectionally sequenced by means of standard fluorescent sequencing techniques and an automated sequence analyzer (ABI).

Kinetic Analyses-k_{uncat}. The uncatalyzed second-order rate constant for the Diels−Alder cycloaddition was determined by measuring the second-order reaction rate constant for the reaction between the diene ligated to the starting mixture of random, pyrU modified RNA and **3**. This rate constant represents the combined effect of [4 + 2] cycloaddition and presumably Michael additions of internal nucleophiles with

(22) Tarasow, T. M.; Tinnermeier, D.; Zyzniewski, C. *Bioconjugate Chem.* **1997**, *8*, 89−93.

the maleimide. The second-order rate constant for the Michael addition was determined using the PEG-OMe ligated to random RNA. The difference between the two measured rate constants equals the second-order rate constant for the cycloaddition. All second-order rate constants were extracted from pseudo-first-order rate constant data measured over a range of **3** concentrations.

$$k_{\text{uncat/Diels−Alder}} = k_{\text{uncat/combined}} - k_{\text{uncat/Michael}} = 6.97 \times 10^{-3}\,\text{s}^{-1}\,\text{M}^{-1} - 1.55 \times 10^{-3}\,\text{M}^{-1}\,\text{s}^{-1} = 5.42 \times 10^{-3}\,\text{M}^{-1}\,\text{s}^{-1}$$

K K_m **and** k_{cat}**.** Kinetic assays were performed as described above. All data were obtained at 500 nM RNA and the indicated amounts of **3**. k_{obs} values were determined by fitting the fraction of unreacted RNA to the following equation for first-order kinetics: $R = \alpha R_0 \exp(-k_{obs}t) + (1 - \alpha R_0)$ where R is the fraction of RNA unreacted, αR_0 represents the fraction (α) of functional RNA (R_0), t is time in minutes, and k_{obs} is the observed first-order rate constant.

Apparent K_i Determination. Apparent K_i values for the cycloaddition product **4** were determined at 500 μM **3** by fitting the observed first-order rate constants to the following equation for inhibition: $k_{obs} = (k_{obs0}/2)(\alpha E - I - K_i + ((K_i + \alpha E - I)^2 + 4IK_i)^{1/2})$ where k_{obs} is the measured rate constant in the presence of **3**, k_{obs0} is the observed rate constant in the absence of **3**, αE represent the fractional (α) concentra-tion of functional active sites (E), I is the concentration of **3**, and K_i is the apparent inhibition constant.[23]

Alternative Substrate Experiments. Biotin maleimides **2** and **6** were purchased from Pierce, (Rockford, IL) and were used without further purification. Fumaramidate **5** and diene **12** were prepared and characterized as previously described.[7b,22] Incubations were conducted as described in the kinetics section at 500 μM dienophile, 500 nM RNA, and 25 °C.

Supporting Information Available: The random region sequences of the RNA Diels−Alderases studied (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA992448+

(23) Williams, J. W.; Morrison, J. G. *Methods Enzymol.* **1979**, *63*, 437−467.