J|A|C|S
ARTICLES
Published on Web 09/03/2004

The Effect of Mutation on RNA Diels−Alderases

Theodore M. Tarasow, Elizabeth Kellogg, Benjamin L. Holley, Daniel Nieuwlandt, Sandra L. Tarasow, and Bruce E. Eaton*

Contribution from the Department of Chemistry, College of Physical and Mathematical Sciences, North Carolina State University, Raleigh, North Carolina 27695-8204

Received February 2, 2004; E-mail: Bruce_Eaton@ncsu.edu

Abstract: Chemical mutagenesis of a previously reported RNA Diels−Alderase (DA22) was followed by in vitro selection based on [4 + 2] catalysis. New mutated families of RNA Diels−Alderases closely related in sequence space were obtained. The mutated Diels−Alderases selected showed significant improvements in catalytic efficiency (k_{cat}/K_m) as compared to the original DA22. The improvement in catalytic activity was primarily due to a decrease in K_m, but modest increases in k_{cat} were also observed. The increase in catalytic activity of these new Diels−Alderases was found not to negatively affect their dienophile specificity. Surprisingly, one of the most active Diels−Alderases (DAM 40), a subtle sequence mutant of DA22, was found to show a new metal dependence and could function with Ni^{2+} as the only transition-metal ion. Truncation experiments of DA22 showed that the region shown to be hypervariable at the 3′-end of the structure could be deleted without a significant decrease in the relative rate of Diels−Alder catalysis.

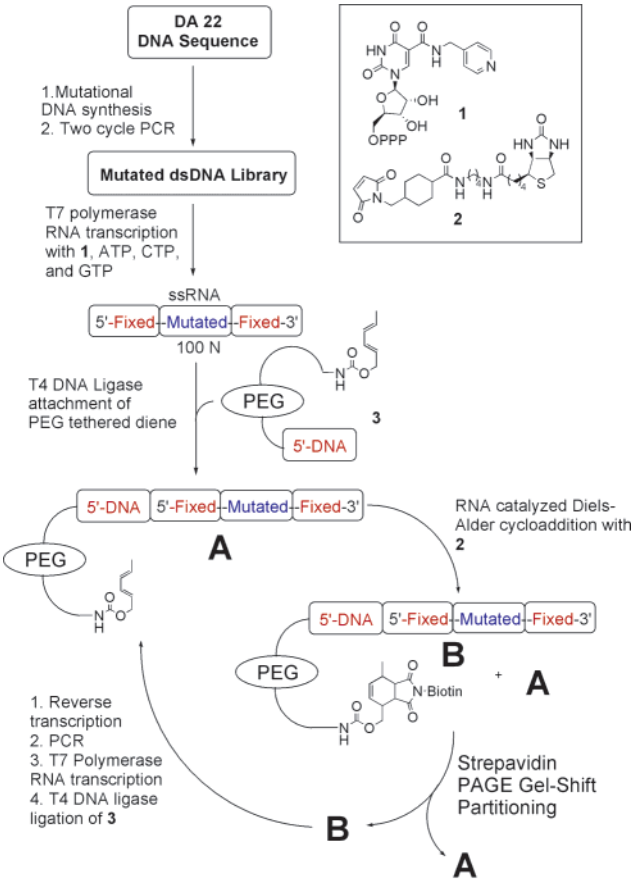
Introduction

Recently, an expanding list of RNA and DNA catalysts have been discovered through in vitro selection techniques. Examples to date include acyl transfers,[1−6] amide synthesis,[7] urea synthesis,[8] Diels−Alder cycloadditions,[9−13] Michael additions,[14] nucleophilic substitutions,[15,16] porphyin metalations,[17,18] enzyme cofactor synthesis[19,20] and nucleotide cofactor-dependent redox reaction.[21] Recent advances in the chemical and enzymatic modification of RNA and DNA are being used to increase the functional diversity in nucleic acids, thereby tuning the nucleic acid for the desired catalytic outcome.[22−32] In addition, RNA catalysts have demonstrated an inherent substrate specificity[12,33,34] and sometimes stereospecificity in product formation.[35] However, as the number and type of organic reactions catalyzed by nucleic acids expand so do the number of mechanistic and evolutionary questions. Herein, we seek to better understand how unique previously reported RNA Diels−Alderases are with regard to the sequence space they were selected from and if better RNA catalysts can be found.

An RNA Diels−Alderase (DA22)-containing modified uridine base had been selected previously[9] and served as the starting sequence for the mutagenesis research described here. DA22 was selected from 10^{14} RNA sequences that contained

(1) Lohse, P. A.; Szostak, J. W. *Nature* **1996**, *381*, 442−444.
(2) Suga, H.; Lohse, P. A.; Szostak, J. W. *J. Am. Chem. Soc.* **1998**, *120*, 1151−1156.
(3) Lee, N.; Bessho, Y.; Wei, K.; Szostak, J. W.; Suga, H. *Nat. Struct. Biol.* **2000**, *7*, 28−33.
(4) Illangasekare, M.; Sanchez, G.; Nickles, T.; Yarus, M. *Science* **1995**, *267*, 643−647.
(5) Illangasekare, M.; Yarus, M. *J. Mol. Biol.* **1997**, *268*, 631−639.
(6) Kumar, R. K.; Yarus, M. *Biochemistry* **2001**, *40*, 6998−7004.
(7) Wiegand, T. W.; Janssen, R. C.; Eaton, B. E. *Chem. Biol.* **1997**, *4*, 675−683.
(8) Nieuwlandt, D.; West, M.; Cheng, X. Q.; Kirshenheuter, G.; Eaton, B. E. *ChemBioChem* **2003**, *4*, 651−654.
(9) Tarasow, T. M.; Tarasow, S. L.; Eaton, B. E. *Nature* **1997**, *389*, 54−57.
(10) Tarasow, T. M.; Tarasow, S. L.; Tu, C.; Kellogg, E.; Eaton, B. E. *J. Am. Chem. Soc.* **1999**, *121*, 3614−3617.
(11) Seelig, B.; Jaschke, A. *Chem. Biol.* **1999**, *6*, 167−176.
(12) Tarasow, T. M.; Tarasow, S. L.; Eaton, B. E. *J. Am. Chem. Soc.* **2000**, *122*, 1015−1021.
(13) Stuhlmann, F.; Jaschke, A. *J. Am. Chem. Soc.* **2002**, *124*, 3238−3244.
(14) Sengle, G.; Eisenfuhr, A.; Arora, P. S.; Nowick, J. S.; Famulok, M. *Chem. Biol.* **2001**, *8*, 459−473.
(15) Wecker, M.; Smith, D.; Gold, L. *RNA* **1996**, *2*, 982−994.
(16) Wilson, C.; Szostak, J. W. *Nature* **1995**, *374*, 777−782.
(17) Li, Y. F.; Sen, D. *Nat. Struct. Biol.* **1996**, *3*, 743−747.
(18) Li, Y. F.; Sen, D. *Biochemistry* **1997**, *36*, 5589−5599.
(19) Huang, F. Q.; Bugg, C. W.; Yarus, M. *Biochemistry* **2000**, *39*, 15548−15555.
(20) Jadhav, V. R.; Yarus, M. *Biochemistry* **2002**, *41*, 723−729.
(21) Tsukiji, S.; Pattnaik, S. B.; Suga, H. *Nat. Struct. Biol.* **2003**, *10*, 713−717.
(22) Benner, S. A.; Battersby, T. R.; Eschgfaller, B.; Hutter, D.; Kodra, J. T.; Lutz, S.; Arslan, T.; Baschlin, D. K.; Blattler, M.; Egli, M.; Hammer, C.; Held, H. A.; Horlacher, J.; Huang, Z.; Hyrup, B.; Jenny, R. F.; Jurczyk, S. C.; Konig, M.; von Krosigk, U.; Lutz, M. J.; MacPherson, L. J.; Moroney, S. E.; Muller, E.; Nambiar, K. P.; Picciilli, J. A.; Switzer, C. Y.; Vogel, J. J.; Ricert, C.; Roughton, A. L.; Schmidt, J.; Schneider, K. C.; Stackhouse, J. *Pure Appl. Chem.* **1998**, *70*, 263−266.
(23) Dewey, T. M.; Zyzniewski, C.; Eaton, B. E. *Nucleosides Nucleotides* **1996**, *15*, 1611−1617.
(24) Gourlain, T.; Sidorov, A.; Mignet, N.; Thorpe, S. J.; Lee, S. E.; Grasby, J. A.; Williams, D. *Nucleic Acids Res.* **2001**, *29*, 1898−1905.
(25) Lutz, S.; Burgstaller, P.; Benner, S. A. *Nucleic Acids Res.* **1999**, *27*, 2792−2798.
(26) Matulic-Adamic, J.; Daniher, A. T.; Karpeisky, A.; Haeberli, P.; Sweedler, D.; Beilgelman, L. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1299−1302.
(27) Perrin, D. M.; Garestier, T.; Helene, C. *J. Am. Chem. Soc.* **2001**, *123*, 1556−1563.
(28) Sakthivel, K.; Barbas, C. F., III. *Angew. Chem., Int. Ed.* **1998**, *37*, 2872−2875.
(29) Teramoto, N.; Imanishi, Y.; Ito, Y. *Bioconjugate Chem.* **2000**, *11*, 744−748.
(30) Thum, O.; Jager, S.; Famulok, M. *Angew. Chem., Int. Ed.* **2001**, *40*, 3990−3993.
(31) Vaish, N. K.; Fraley, A. W.; Szostak, J. W.; McLaughlin, L. W. *Nucleic Acids Res.* **2000**, *28*, 3316−3322.
(32) Lermer, L.; Roupioz, Y.; Ting, R.; Perrin, D. M. *J. Am. Chem. Soc.* **2002**, *124*, 9960−9961.
(33) Seelig, B.; Keiper, S.; Stuhlmann, F.; Jaschke, A. *Angew. Chem., Int. Ed.* **2000**, *39*, 4576−4579.
(34) Sengle, G.; Eisenfuhr, A.; Arora, P. S.; Nowick, J. S.; Famulok, M. *Chem. Biol.* **2001**, *8*, 459−473.
(35) Stuhlmann, F.; Jaschke, A. *J. Am. Chem. Soc.* **2002**, *124*, 3238−3244.

Scheme 1



the modification 5-(4-pyridylmethyl)carboxamide-uridine (**1** of Scheme 1) in place of uridine. The pyridyl modification expands the functional group diversity of the RNA by providing additional hydrogen bonding, hydrophobic and dipolar interactions, as well as an alternative σ-donor ligand for metal ion binding. Molecular replacement experiments have shown that the precise positioning of pyridyl groups can be crucial for DA22 Diels−Alderase activity.[10] Consistent with a well-defined catalytic active site, it was previously found that DA22 catalytic activity absolutely depended on the presence of cupric ion (Cu^{2+}). Moreover, DA22 demonstrated a high degree of substrate specificity even for other maleimide dienophiles of comparable reactivity in the spontaneous Diels−Alder reaction. It was of interest to determine if DA22 was unique or if other sequences closely related in sequence space would show similar requirements for catalytic activity. What is the density of catalytically active sequences for this particular diene and dienophile combination? Herein we describe experiments in which a library of 10^{14} RNA sequences was generated by chemical synthesis mutation of RNA Diels−Alderase (DA22) at a rate of 25% per nucleotide in the variable region. The in vitro selection strategy shown in Scheme 1 was used to select mutated RNA Diels−Alderases that could catalyze the same reaction as DA22. Individual RNA catalysts were isolated and sequenced, the sequence data were examined for structural information, and the individual RNA isolates were kinetically characterized to determine catalytically active rate enhancements, the mechanistic features of metal dependence, substrate specificity, and product inhibition.

Results and Discussion

DA22 Mutation/Library Synthesis and in Vitro Selection. A 25% mutation rate of the DA22 sequence was achieved by chemically synthesizing the corresponding DNA sequence on an automated DNA synthesizer (ABI 392) using 9:1:1:1 phosphoramidite ratios of the cognate nucleotide relative to the three noncognate nucleotides at each nucleotide position in the nonpriming region between sequence positions 34 and 130 in Figure 2. This chemical synthesis gave a mutated single-stranded DNA (ssDNA) library that was subjected to two cycle PCR (Scheme 1). The mutated double-stranded DNA (dsDNA) library created by PCR was used in T7 polymerase RNA transcription to give an ssRNA library (ca. 10^{14} sequences). The resulting ssRNA library was estimated to contain 10 copies of the original DA22 sequence and approximately 10^{13} mutated sequences. T4 DNA ligase was used to attach DNA-PEG-diene **3** to the 5′-end of the ssRNA library, giving **A**. Incubation of the library **A** with **2** gave a small fraction of active Diels−Alderases **B** and mostly unreacted **A**. Partitioning of **B** from **A** was achieved by denaturing polyacrylamide (6%) gel-shift electrophoresis (PAGE gel-shift) and isolation of RNA with a gel mobility consistent with binding to Strepavidin. The RNA was isolated from the gel slices by the freeze squeeze technique and subjected to reverse transcription with AMV reverse transcriptase to give cDNA of the isolated RNA sequences. The cDNA was PCR-amplified under standard conditions, and the resulting dsDNA was used as a template for T7 polymerase RNA transcription and the beginning of the next cycle.[7,9]

The initial RNA library (**A** of cycle 1 in Scheme 1) showed no catalytic activity; however, after two cycles of selection and amplification detectable catalytic activity was observed. It should be noted that in the original selection that yielded DA22 eight cycles of selection and amplification were required to enrich the population to the point of observing catalytic activity.[9] Clearly, the mutated RNA library contained a greater abundance of sequences that could catalyze this particular Diels−Alder reaction. Although activity was observed after only two cycles of selection, 11 cycles were conducted to isolate the most active catalysts. Increasingly stringent partitioning was required following cycle three to drive the selection and limit the amount of RNA carried forward. For example, in cycle 1, 2.7% of **A** reacted after 480 min (100 μM **2**) and by cycle 11, 1.2% **A** reacted after 1 min (25 μM **2**) as determined by phosphorimaging of Strepavidin dependent PAGE gel-shift RNA. Necessary reaction controls showed that the gel-shifted RNA was also dependent on **2** and **3**. This increase in catalytic activity for the RNA library pool indicated that mutation of the DA22 sequences had resulted in a population of RNA catalysts with increased activity. Following cycle 11, increases in *pool* activity appeared to plateau and the pool was subsequently cloned and sequenced using previously published procedures.[12]

Primary Sequence Analysis. From the cycle 11 pool, cloning and sequencing procedures allowed for the isolation of 51 individual RNA sequences (isolates). Of these 51 isolates, 49 were unique sequences (Figure 1). Importantly, *none* had precisely the same sequence as DA22. Figure 1 shows that many new Diels−Alderases were obtained from mutating DA22, making it likely that less efficient catalysts with similar sequences were removed during the selection. Clearly, for this single [4 + 2] cycloaddition with this specific diene and

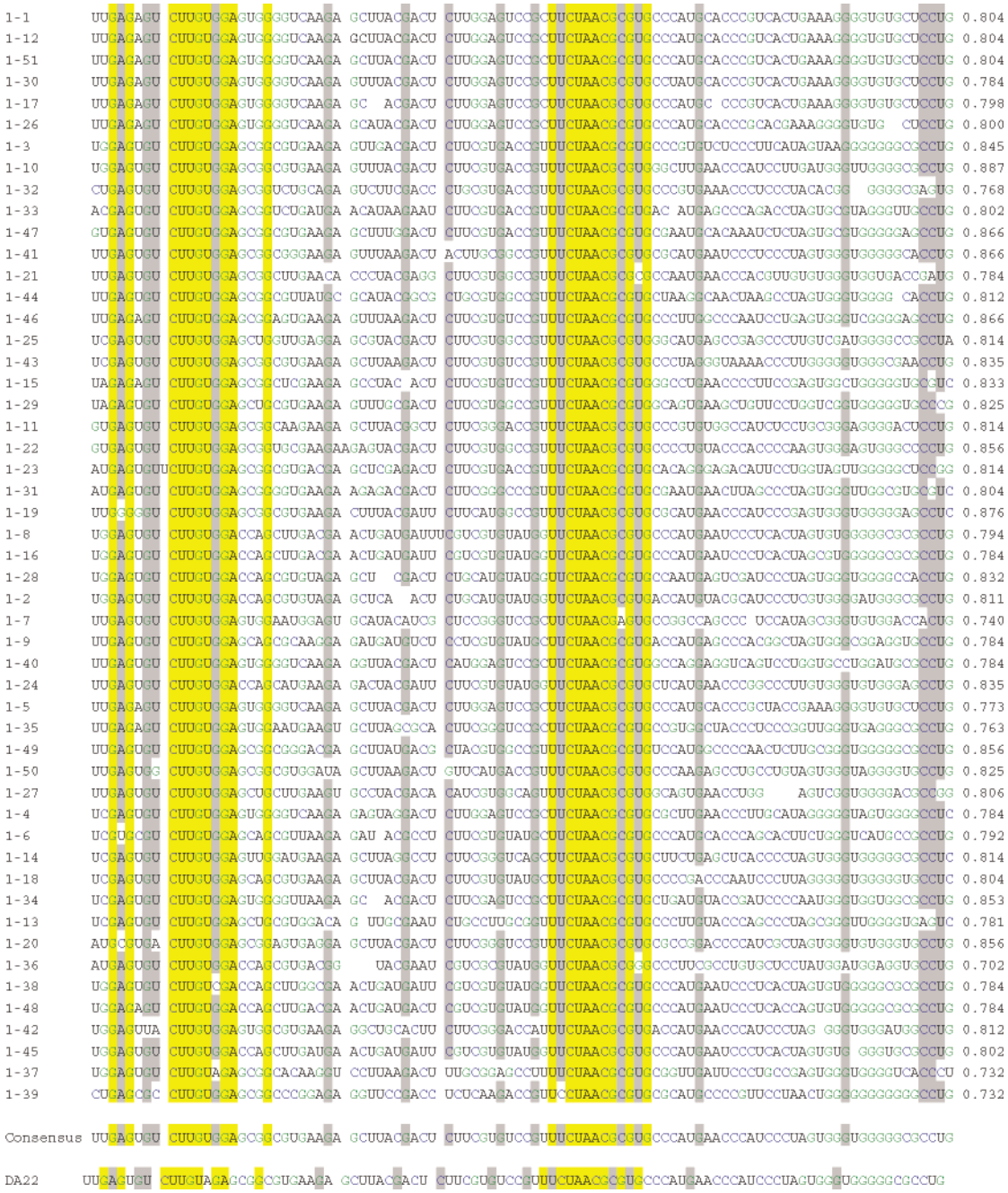


Figure 1. Aligned sequences of 51 RNA isolates identified. Fixed sequence regions are omitted for clarity. Completely conserved residues are highlighted in yellow, while those that were conserved in >90% of the sequences are highlighted in gray. The proportion similarities to the original DA22 sequence taking all residues into consideration are shown at the far right.

dienophile pair, there are many catalytic sequence solutions contained within the variants of the consensus RNA Diels–Alderase framework.

Primary sequence alignment of the 51 RNA isolates and analysis of the mutated region (97 nucleotides) revealed that 19 of the bases were absolutely conserved in all isolated sequences and that an additional 17 nucleotides were highly conserved in >90% of the sequences. The fully conserved bases were found shifted toward the 5′-end of the variable region while the highly conserved residues were more evenly distributed. With only scattered point mutations, both the completely and highly conserved residues corresponded to the nucleotides found in DA22, indicating that these residues were necessary for Diels–Alderase function in this particular structural motif. Interestingly, in 49 of the 51 isolates the one highly conserved residue not found in the original DA22 sequence was a mutation (at position 47 near the 5′-end) of an A to G. The prevalence of this mutation in the population suggests that it confers a distinct selective advantage in catalytic activity. The importance of this mutation (A to G) on function is further underscored by its location to a putative key structural feature (vide infra).

Analysis of the number of mutations observed in the isolates provides insight as to the diversity of the Diels–Alderase sequence space around the highly conserved core sequence and the putative active site structure. It should be noted that the original selection that gave DA22 started with 10^{14} sequences.

For this original selection, 100 consecutive nucleotides were randomly distributed. This gives a total potential sequence space of 100^4 or approximately 10^{60}. Clearly, selecting from only 10^{14} sequences is a sparse sampling of Diels−Alderases that could exist in this potential sequence space. It was of interest to determine if the chemical mutation rate of 25% resulted in a more comprehensive search for Diels−Alderases for this particular diene and dienophile pair. Analysis of the percent similarity to DA22 of each isolate is shown in Figure 1 (right column). The average similarity to DA22 for all unique sequences is 80.7% (19.3% mutation rate), representing an average of approximately 19 mutations per isolate. However, this mutation rate includes 19 bases of the 97 base variable region that were absolutely conserved throughout all sequences and 17 bases that were conserved in more than 90% of the sequences. Removing these conserved bases from the mutation rate calculation gives an average mutation rate of 30.8%, significantly higher than the 25% rate used in the DNA synthesis. These results suggest that a mutation rate of >30% could be used to further investigate the potential sequence space and that additional Diels−Alderases could be found. These regions of RNA sequence variability demonstrate that there are many unique catalytic solutions to this single diene and dienophile [4 + 2] cycloaddition reaction. It remained to be determined if this apparent increase in the number of active Diels−Alderases for this particular diene and dienophile resulted in a corresponding decrease in substrate specificity (vide infra).

Potential Secondary Structure Motifs. The mutated RNA sequences isolated from this selection contain the 5-position pyridyl modification of **1** in place of the 5-position hydrogen in uridine. There are no folding algorithms that have been parametrized to account for this modification. The added hydrophobicity of this pyridyl group, differences in metal ion coordination compared to unmodified RNA, and the new hydrogen bonding possibilities for the amide linkage to the modified uridine would all be expected to contribute to the RNA structure. For this reason, secondary structure predictions may be of limited value and are only included to stimulate thought and to provide a potential picture of these Diels−Alderases consistent with all known Diels−Alderases for this diene dienophile pair.

The consensus sequence shown in Figure 1 was folded using the Zucker algorithm-based program MFOLD,[36−38] and the lowest energy (−74.0 kcal/mol) secondary structure obtained is shown in Figure 2. Base-pair analysis and covariation across all of the mutational isolates is consistent with this secondary structure. Many of the completely and highly conserved residues identified in the primary sequence analysis could be base-paired with the 5′ fixed sequence region of the molecules as shown in Figure 2. Sequence conservation and co-variation analysis are also consistent with formation of the second stem and the stem that connects the two three helix junctions. However, the proposed 3′-end of the structure has only a few highly conserved points, making this portion of the structure uncertain. The 3′-end comprises a significant portion of the variable sequence region. It appears that this portion of the RNA structure is more tolerant of changes that still afford catalytic active sites and there are many structural motifs that are possible.

Most interestingly, only one highly conserved mutation was found in comparing the conserved region of the original DA22 sequence to the consensus sequence found for the DAM isolates. The high degree of sequence conservation found for the point mutation A to G found at position 47 is consistent with this residue being a key structural component of these DAM Diels−Alderases. Only isolate 38 contained A at this position. Position 47 is at the three-helix junction in Figure 2 where a high number of pyridine-modified uridine residues are also found. The high proximity of conserved pyridine functionality may indicate that this three-helix junction comprises a portion of the catalytic pocket where the pyridine groups form cupric ion Lewis acid sites.

Diels−Alderase Kinetic Characterization. Six DAM Diels−Alderases obtained from the mutagenesis/selection experiment were kinetically evaluated. Under single turnover conditions, the kinetic constants k_{cat} and K_m were determined and are shown in Table 1, where they are compared to values of the original Diels−Alderase DA22. The k_{cat} values for the DAM Diels−Alderases range from 0.32 to 1.55 min^{-1}, representing an approximate 2-fold decrease to a 2.3-fold increase in the catalyzed reaction rate compared to DA22.

When compared to the uncatalyzed second-order rate constant (k_{uncat}) of 0.325 M^{-1} min^{-1}, the k_{cat} values represent an effective molarity range of 0.98−4.7 M. K_m values range from 119 to 460 μM for the DAM Diels−Alderases and in all cases represent significantly stronger substrate binding contributing to the observed rate enhancement as compared to those of DA22. The decrease in K_m values for the DAM Diels−Alderases relative to those of DA22 more than offset the lower k_{cat} values obtained for some of the isolates and lead to higher observed reaction rates, as measured by k_{cat}/K_m, for all of the DAM Diels−Alderases relative to those of DA22. Comparing the uncatalyzed second-order rate constant, k_{uncat}, to the range of k_{cat}/K_m values shows these Diels−Alderases to have achieved 2400 to 40,000-fold rate enhancements over the uncatalyzed cycloaddition. Clearly, this mutational approach of creating new biased libraries of RNA sequences is a viable method to find Diels−Alderases with enhanced catalytic activity. Given the high mutation frequency observed for these DAM isolates it appears that further mutation could reveal more active catalysts, but will the substrate specificity of these Diels−Alderases be compromised as their activity increases?

Substrate Specificity. Previously, DA22 had been shown to possess a high degree of substrate specificity even when presented with dienophile substrates that were structurally and electronically very similar to the maleimide dienophile used in its selection. This property seemed to be inherent to DA22 as no selective pressure was applied to drive for substrate specificity. The mutated Diels−Alderases DAM **1**, **7**, **16**, **40**, **43**, and **46** were assayed for their substrate specificity to determine if the sequence changes required to increase catalytic activity relative to DA22 had resulted in catalyst structures that were compromised in their ability to discriminate between similar dienophiles. These DAM Diels−Alderases were incubated with either of the two alternative biotin maleimides dienophiles (**4** or **5**, 20 μM) shown in Table 2. The relative extent of reaction

(36) Jaeger, J. A.; Turner, D. H.; Zuker, M. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 7706−7710.
(37) Zuker, M. *Science* **1989**, *244*, 48−52.
(38) Zuker, M. *Curr. Opin. Struct. Biol.* **2000**, *10*, 303−310.

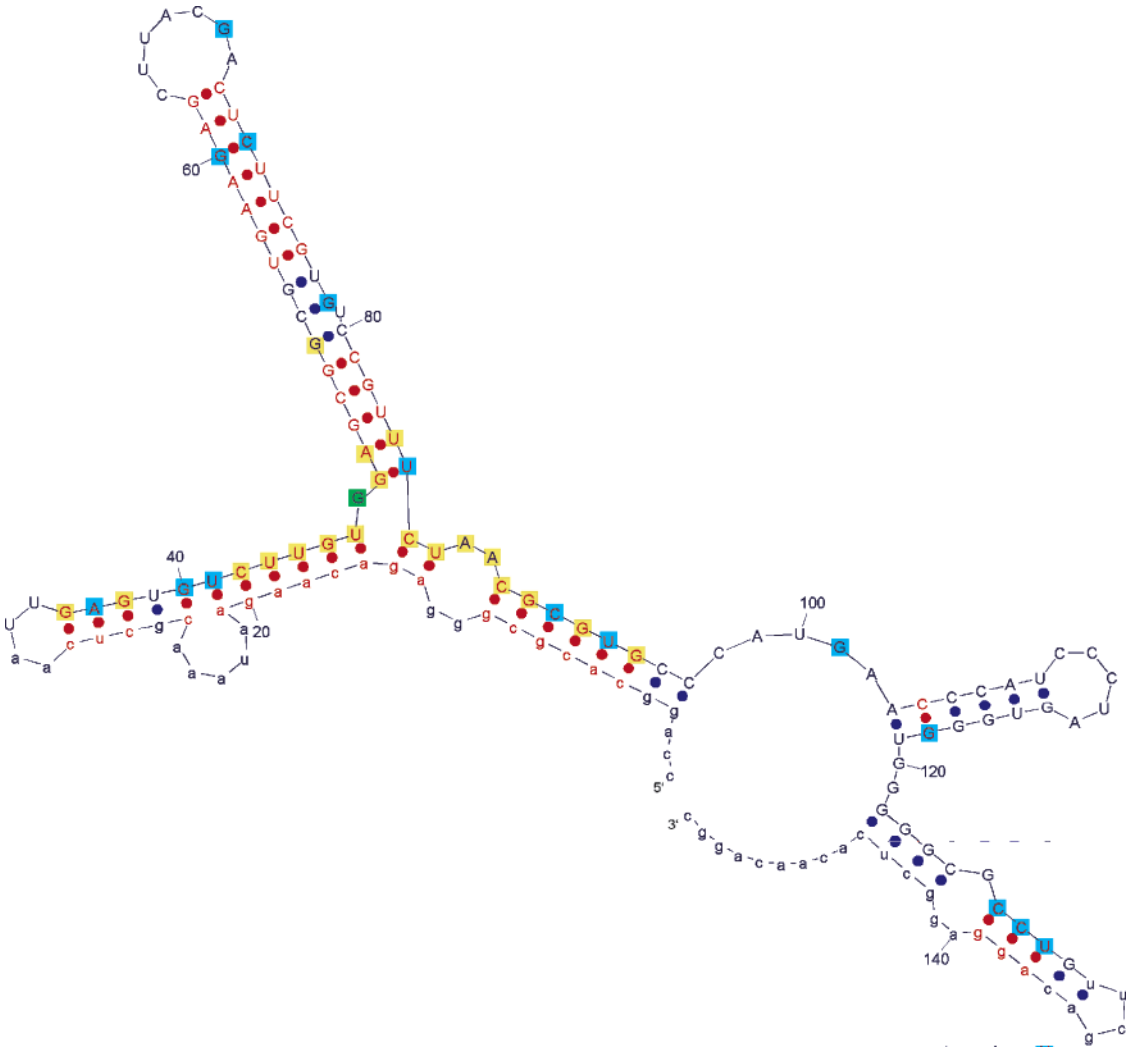


Figure 2. Lowest energy (−74.0 kcal/mol) secondary structure as predicted by the MFOLD program for the consensus sequence shown in Figure 1 but including fixed sequences. Fixed regions of sequence are shown in lower case with the mutated sequence region corresponding to Figure 1 in upper case. Residues that were completely conserved in all isolates are highlighted in yellow, while those that were conserved in >90% of the isolates are highlighted in blue. The single highly conserved inherited mutation of A to G at position 47 is highlighted in green. Highly conserved base pairs as determined by residue conservation or co-variation are shown in red.

Table 1. Kinetic and Inhibition Constants for DA22 and Mutated DAM Diels−Alderases

RNA isolate	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m ($\mu M^{-1}\,min^{-1}$) × 10^2	K_{iapp} (μM)
DA22[a]	0.66 ± 0.12	2300 ± 500	0.0287	32.5 ± 2.6
DAM **1**	1.55 ± 0.10	119 ± 20	1.30	8.40 ± 0.18
DAM **7**	0.65 ± 0.16	460 ± 110	0.14	36.3 ± 4.5
DAM **16**	1.02 ± 0.09	317 ± 55	0.32	22.6 ± 4.1
DAM **40**	1.25 ± 0.12	278 ± 52	0.45	12.3 ± 3.3
DAM **43**	0.31 ± 0.04	409 ± 89	0.078	39.2 ± 5.2
DAM **46**	0.39 ± 0.06	367 ± 59	0.101	39 ± 12

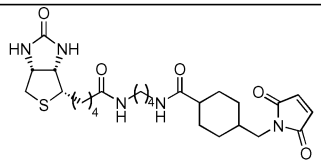
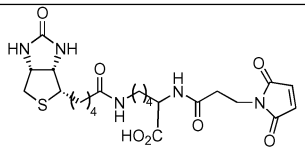
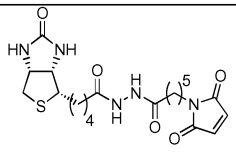
[a] Data for DA 22 from ref 9.

observed during the incubation period for each Diels−Alderase is summarized in Table 3 and shows that, in general, the DAM Diels−Alderases demonstrate similar levels of substrate specificity when compared to the original RNA Diels−Alderase DA22. However, DAM **43** and DAM **46** did show low but significant levels of reaction with maleimide **5**. Interestingly, these two Diels−Alderases had the lowest k_{cat} values of all the catalysts studied. In no case was significant reaction observed with dienophile **4**.

Although alternative substrates **4** and **5** did not in general serve as substrates for the catalyzed cycloaddition, each could bind at the active site in a catalytically unproductive mode especially in light of their similarity to maleimide **2**, either because of the biotin or the maleimide or both. When each Diels−Alderase was incubated with **2** (20 μM) in the presence of a large excess of **4** (500 μM), no measurable decrease in the amount of reaction product was observed (data not shown), indicating that **4** does not effectively compete for binding to the active site. These results further demonstrate that the biotin moiety of **2** is not important for substrate recognition since **4**, **5**, and **2** all contain biotin. The specificity data support the notion that through mutation, RNA catalytic activity can be increased without sacrificing substrate specificity.

Diels−Alder Product as a [4 + 2] Cycloaddition Transition-State Analogue. These mutated Diels−Alderases expand the number of kinetically characterized RNA catalysts for this particular diene and dienophile [4 + 2] cycloaddition to 15. Together, these 15 Diels−Alderases provide an opportunity to

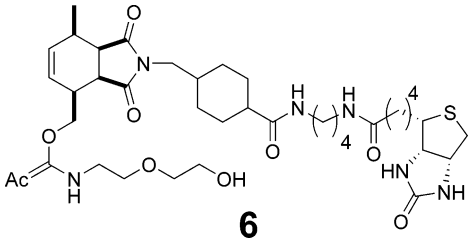
Table 2. Dienophile Substrate Specificity for the DA 22 and the DAM Diels−Alderases **1**, **7**, **16**, **40**, **43**, and **46**[a]

Substrate #		DA 22	DAM **1**	DAM **7**	DAM **16**	DAM **40**	DAM **43**	DAM **46**
2		1	1	1	1	1	1	1
4		0.04	0.04	0.04	0.04	0.03	0.02	0.03
5		0.06	0.04	0.05	0.03	0.05	0.12	0.21

[a] Values are background-corrected, normalized to the total amount of active catalyst, and relative to the catalyzed reaction with **2**. The values are below the assay detection limits.

Table 3. Kinetic Results from the Original DA and Mutated DAM Diels−Alderases

Diels–Alderase	K_{iapp} (μM)	K_m (μM)	K_m/k_{cat} (mM s)
DAM **1**	8.40	119	4.61
DAM **7**	36.3	460	42.5
DAM **16**	22.6	317	18.6
DAM **40**	12.3	278	13.3
DAM **43**	39.2	409	77.2
DAM **46**	39	367	55.3
DA 22[b]	32.5	2300	209
DA 1[b]	29.7	1920	274
DA 17[b]	44.6	1220	305
DA 2[b]	59.9	1800	469
DA 7[b]	222	1640	520
DA 10[b]	135	830	778
DA 37[b]	56.0	118	165
DA 15[b]	108	69	243
DA 24[b]	165	29	195



6

[a] K_{iapp} values were determined using 6 as an inhibitor. [b] Previously reported.[12]

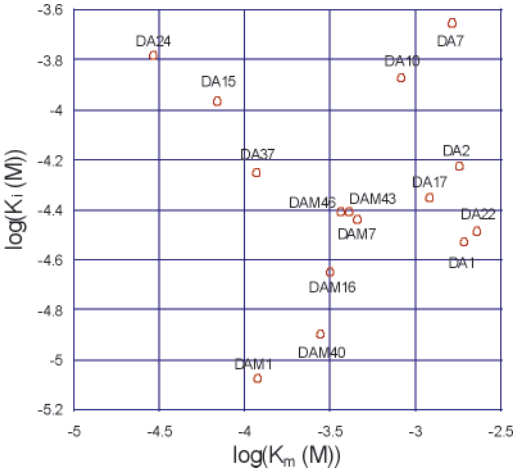


Figure 3. Plot of log(K_i) versus log(K_m) for all the data in Table 2.

examine the transition-state analogue character of the cycloaddition product across a number of unique but related RNA catalysts. In analogy to previous work on proteins, a number of enzymes have been examined to determine the extent of transition-state stabilization in catalysis and how well transition-state analogue inhibitors take advantage of this stabilization. Many of these studies have generally focused on a single enzymatic system and a range of substrates and corresponding transition-state analogue inhibitors where a linear correlation between log K_i and log(K_m/k_{cat}) indicates inhibition by transition-state mimicry of the varied inhibitors.[39]

Combining the original RNA Diels−Alderases (DA series) and the DAM Diels−Alderases, it can be determined if a correlation exists between K_{iapp} and K_m/k_{cat}. All 15 RNA sequences were kinetically characterized, and the K_{iapp} values of the Diels−Alder product **6** were determined for each. The data are summarized in Table 3 and plotted in Figures 3 and 4. The data displayed in Figure 3 (log K_{iapp} vs log K_m) clearly show no correlation for these Diels−Alderases consistent with the product **6** not acting as a ground-state inhibitor. However, from Figure 4 there appears to be some correlation ($R = 0.82$) between log(K_{iapp}) values for the inhibitor and log(K_m/k_{cat}) values for the dienophile substrate, indicating that differences in the structure and active site within these catalysts produce similar effects on enzymatic rate enhancement and transition-state analogue binding. It should be noted that preliminary data

(39) Mader, M. M.; Bartlett, P. A. *Chem. Rev.* **1997**, *97*, 1281−1301.

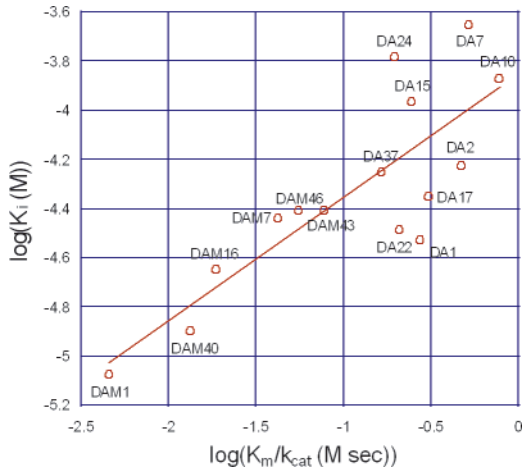


Figure 4. Plot of $\log(K_i)$ versus $\log(K_m/k_{cat})$ for all the data in Table 2.

suggest that DA 7 in particular may be a Diels−Alderase that produces exo-product, and the K_{iapp} value generated using the endo-product would not be expected to correlate well with K_m/k_{cat}. The correlation of $\log(K_{iapp})$ to $\log(K_m/k_{cat})$ is significantly higher ($R = 0.97$) when only the DAM Diels−Alderases are plotted (elimination of all DA points in Figure 4), suggesting convergence in active site structure toward a similar mechanism. These results are consistent with these Diels−Alderases providing cycloaddition reaction catalysis via stabilization of a late [4 + 2] endo transition state. More thorough kinetic and structural investigation of these Diels−Alderases will be required to unravel the details of how [4 + 2] cycloaddition is catalyzed by these related RNA molecules.

Metal Dependence. The original in vitro selection that produced the original RNA Diels−Alderases (e.g., DA22) included a mixture of metal ions (Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Al^{3+}, Ga^{3+}). This provides the pyridyl-modified RNA the opportunity to utilize any metal or metal combination in structural or functional roles. Previous work demonstrated that the catalytic activity of DA22, and all the other original Diels−Alderases, was absolutely dependent on the presence of cupric ion which, combined with results from molecular replacement experiments, suggested that organization of certain pyridine groups is required to form specific copper ion binding sites. Consistent with known pyridyl−cupric complexes, these RNA active sites could function by Lewis acid catalysis in binding to the dienophile carbonyl(s). To determine if new metal-dependent activity could emerge on further exploration of sequence space, the mutated Diels−Alderases were incubated with the same mixture of metal ions that was used in the original selection. Of the six DAM Diels−Alderases characterized, all showed a dependence on Cu^{2+} ion. However, unlike the original DA selection, only five DAM Diels−Alderases (**1**, **7**, **16**, **43**, and **46**) of the six demonstrated an exclusive dependence on Cu^{2+}, with comparable catalytic activity observed in the presence of the metal mixture or just Cu^{2+} and no catalytic activity observed in the absence of Cu^{2+}. Interestingly, full activity of isolate **40**, as defined by that observed with the entire metal mixture, was found to be dependent on the presence of both Cu^{2+} and Ni^{2+}. Partial catalytic activity (20−40%) was observed with isolate **40** in the presence of *either* Cu^{2+} or Ni^{2+}, but to observe full activity both transition metals were required. This could indicate either two metal binding sites specific for each metal or a new bimetallic site. The importance of this new Ni^{2+} metal dependence is underscored by the fact that isolate **40** was one of the most active Diels−Alderases found to date. The unique metal dependence of isolate **40** demonstrates that RNA sequences with new structural and/or mechanistic features can be generated through mutation of existing RNA catalysts.

Sequence Truncation and Relative Catalytic Activity. Comparison of the DAM Diels−Alderase sequences and the parent sequence of DA22 showed that the 3′-region past position 96 of Figure 2 had few highly conserved nucleotides. It was of interest to determine if this variability in sequence was because this portion of the RNA was not essential for activity. Since all the DAM Diels−Alderases were descendent from DA22 and contained the same highly conserved consensus region, DA22 was subjected to a series of truncation experiments. Two RNA Diels−Alderase truncates DA_TR96 and DA_TR119 were prepared by T7 RNA polymerase transcription of dsDNA templates produced via PCR amplification of DA22 with the following primers: 5′-GCACGCGTTAGAAACGGACACG (producing DA_TR96), 5′-ACCCACTAGGGATGGGTTCATG (producing DA_TR119). DA_TR96 corresponds to the structure shown in Figure 2 where the sequence is truncated by ending at position 96. This truncate results in the elimination of all of the 3′-end hypervariable region. DAM_TR119 corresponds to the structure in Figure 2 ending at position 119. This truncate contains only one of the stem loops found at the 3′-end of Figure 2. All truncates gave excellent transcription and ligation yields. Kinetics were determined as previously described using gel-shift electrophoresis strepavidin partitioning to measure the fraction of ^{32}P body labeled RNA reacted. The kinetics were measured over four concentrations (250, 125, 62.5, 31.25, and 15.63 μM) of **2** and the k_{cat} and k_{cat}/K_m values determined in a procedure identical to that used for the DAM isolates. The relative rates found for DA_TR96 and DA_TR119 compared to DA22 were significantly different. Relative to DA22 ($k_{cat(DA\_TR96)}/k_{cat(DA22)}$), DA_TR96 showed a 1.7 increase in k_{cat} and a 1.3 increase in K_m ($K_{m(DA\_TR96)}/K_{m(DA22)}$), providing a modest *increase* in k_{cat}/K_m for this truncate of 1.3. Surprisingly, DA_TR119 gave no significant increase in k_{cat}/K_m relative to DA22. Clearly, *all of the sequence in DA22 past position 96 of Figure 2 is nonessential for catalytic activity*. DA_TR96 and DA_TR119 show that the highly conserved region of these truncates is only slightly affected by the alteration of the 3′-end of these Diels−Alderases. There appears to be a limited interaction between the structural elements at the 3′-end of these RNA catalysts with the highly conserved catalytic core.

Conclusions

When obtaining a perspective on Diels−Alderases and the results found to date, it is important to note that there are currently no naturally occurring RNA or protein examples for comparison. The challenges of [4 + 2] cycloaddition catalysis include low affinity of acyclic dienes for the biopolymer catalysts and the potential problem of product inhibition limiting catalytic efficiency. In the research reported here, it has been found that mutation of a RNA Diels−Alderase can generate numerous related catalysts for a specific [4 + 2] cycloaddition. A picture is beginning to emerge as to how primary sequence and concomitant secondary structure relates to RNA Diels−Alderase activity for this particular diene and dienophile pair.

Kinetic analysis of the DAM Diels−Alderases clearly demonstrated that catalytic efficiency was enhanced and that further mutation could lead to greater improvements in catalytic activity. Improvements in both K_m and k_{cat} were observed, but only K_m was consistently improved. While the K_m values of these Diels−Alderases are in the range of protein enzymes, it is expected that protein Diels−Alderases would have significantly greater catalytic activity, most likely accomplished by higher k_{cat} values. It remains to be determined if additional modification of the RNA or further mutation and selection would result in RNA Diels−Alderases with significant improvements in k_{cat}. It should be noted that the modified RNA libraries used in the selection for DAM Diels−Alderases only contained a single modification and that of the total RNA sequence space possible ($4^{100} \approx 10^{60}$) only a small fraction has been investigated. It appears that we have yet to test the catalytic limits of RNA Diels−Alderases.

Notably, substrate specificity was maintained in the mutated Diels−Alderases for the diene and dienophile pair used in the selection. Even maleimide dienophiles of comparable reactivity in the spontaneous Diels−Alder reaction could not be utilized effectively as substrates. Related to this result, the catalytic activity increase for the DAM Diels−Alderases was largely due to a decrease in K_m. The dienophile specificity and the decrease in K_m observed in the DAM Diels−Alderases are consistent with mutation of the RNA active site providing higher affinity and potentially specificity[40] in saturable substrate binding. Many RNA catalysts were generated for the diene and dienophile pair studied, providing the opportunity to correlate observed catalysis with inhibition by the cycloaddition product. Correlation between K_{iapp} for the product and K_m/k_{cat} for the DAM Diels−Alderases indicated that the active sites were similar, and when considered in the context of the DAM families, clustered together in sequence space. It seems reasonable to propose that further mutation could lead to further convergence on superior highly specific [4 + 2] catalysts for this diene dienophile pair. In addition, it is clear that the 3′-end hypervariable region found for these Diels−Alderases is not essential for catalytic activity, and future work should be focused on these truncated [4 + 2] RNA catalysts. The question remains, however: is this the only acyclic diene and dienophile pair that can be catalyzed by a RNA Diels−Alderase? Furthermore, can RNA Diels−Alderases be used as general [4 + 2] cycloaddition catalysts?[41] Clearly, much work remains before these questions have been rigorously addressed.

One new feature of the selection reported here is that mutation of a RNA catalyst can give sequence families clustered in sequence space but with surprises in metal dependence. As reported previously, the original Diels−Alderases showed an absolute cupric metal dependence. Interestingly, Diels−Alderase DAM **40** revealed a new dependence on Ni^{2+} and Cu^{2+}. This is the first demonstration of RNA structure and/or function being dependent on Ni^{2+}. Apparently, additional metalloenzyme possibilities exist for these pyridyl-U-modified RNA Diels−Alderases. The superposition of the various metal ions, their concentrations in selection, functional modification of the RNA, and density of concomitant catalytically active sequences, make the landscape of RNA [4 + 2] cycloaddition catalysis enormous. Fortunately, the density of RNA catalysts in this landscape is relatively high. Further research is in progress to better understand the breadth of RNA catalysis for other [4 + 2] cycloadditions and catalytic role of the metal ions in these RNA structures.

(40) Eaton, B. E.; Gold, L.; Zichi, D. A. *Chem. Biol.* **1995**, *2*, 633−638.
(41) Jaschke, A.; Seelig, B. *Curr. Opin. Chem. Biol.* **2000**, *4*, 257−262.

Experimental Section

Mutated RNA Library Design Synthesis. The mutated RNA Diels−Alderase library was synthesized from dsDNA as described previously[9] with the exceptions noted below. The initial mutated RNA pool of sequences was created from a dsDNA library based on the sequence of the previously characterized RNA Diels−Alderase DA22. The dsDNA library was prepared by synthesizing the corresponding cDNA using an ABI 392 oligonucleotide synthesizer, 2000 Å CPG support (Glenn Research, Sterling VA) and the standard reagents (Glenn Research, Sterling VA) and coupling procedures. The fixed regions were not mutated and were synthesized as the same sequences described for DA22.[9] The sequence positions in DA22 that corresponded to the randomized region in the original RNA library (positions 34−130 in Figure 2) were mutated at a rate of 25% by using 9:1:1:1 molar ratios of the cognate phosphoramidite (Glenn Research, Sterling, VA) to each of the remaining three phosphoramidites at each base sequence position. The cDNA library was cleaved from the CPG solid support and deprotected by incubating in saturated NH_4OH at 55 °C for 10 h, followed by purification using denaturing 6% polyacrylamide electrophoresis (PAGE, 19:1 cross-linking). The corresponding dsDNA and subsequent RNA libraries were prepared from this cDNA library as previously reported, including the substitution of modified uridine **1** for UTP during T7 RNA polymerase transcription. The resulting 144 nucleotide mutated 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), extracted once with chloroform, and ethanol precipitated. The resulting pellet was dissolved in water and quantitated by UV (ϵ_{260} = 40 μg/mL) and specific activity. The DNA 10-mer-PEG-diene was ligated by combining the RNA library (1 nmol), PEG modified DNA 10-mer **3** (2 nmole), and the bridging oligomer (3 nmol) 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 u/μL, Promega) and T4 DNA ligase (80 units, 5u/μL, Boeringer Mannheim) were then added, and the mixture was incubated at 37 °C for 4−16 h. The ligation product **A** was purified on a 6% denaturing polyacrylamide (19:1 cross-linking) and quantitated by UV spectroscopy and specific activity.

Incubation/Reaction Conditions. All RNA incubations were conducted under the following conditions except as noted: 50 mM HEPES pH 7.0, 500 nM modified RNA, 200 mM NaCl, 200 mM KCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 10 μM each Al(lactate)$_3$, $Ga_2(SO_4)_2$, $MnCl_2$, $FeCl_2$, $CoCl_2$, $NiCl_2$, $CuCl_2$, and $ZnCl_2$, 10% ethanol and 2% dimethyl sulfoxide (DMSO). The dienophile (**2**) was added as a stock solution in DMSO such that the **2** aliquot always accounted for the entire 2% DMSO required. The concentration of **2** was held constant at 100 μM throughout the first seven selection cycles and reduced to 25 μM for cycles 8−11. The 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 ethanol and dienophile **2** 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 and/or passing the solution over two successive Nap columns to remove excess **2**.

Reaction Assay and Partitioning. Measurement of the extent of reaction and partitioning of reacted and unreacted RNA molecules were

accomplished using a Strepavidin (SA)-dependent PAGE gel-shift. The amount of SA added to the RNA sample varied according to how much **2** was used in the incubation and whether Nap columns had been used following the incubation. If Nap columns (2) were employed, then SA was added such that the ratio of **2** in the incubation to added SA was 20. If Nap columns were not utilized, the ratio of SA to **2** was 4:1. Control experiments confirmed that these amounts of SA were sufficient to bind all of the biotinylated (reacted) RNA (data not shown). The RNA was incubated with the SA for 20 min at ambient temperature, followed by the addition of 0.5 volume of formamide loading buffer. Samples were then heated at 72 °C for 3 min, loaded on a preheated (~50 °C) 6% denaturing (19:1 cross-linking) polyacrylamide gel and electrophoresed. The shifted and unshifted bands were visualized by autoradiography and/or phosphorimaging, the latter 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. The PCR Script Amp Cloning Kit (Stratagene, La Jolla, CA) was used to clone the dsDNA obtained following cycle 11. All procedures and supplies were used without modification. Plasmid DNA was obtained from the clones using a PERFECT prep plasmid DNA kit (5′-3′ Inc., Boulder, CO). The purified DNA was bi-directionally sequenced using standard fluorescent sequencing techniques.

Kinetic Analyses k_{uncat}. The uncatalyzed second-order rate constant for the Diels−Alder cycloaddition was determined previously ($k_{uncat/Diels-Alder} = 5.42 \times 10^{-3}$ $M^{-1}s^{-1}$).[9]

K_m and k_{cat}. Kinetic assays were performed as described above. All data were obtained at 500 nM RNA and up to 500 μM of **2** (500 μM solubility limit for **2**). The 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. k_{cat} and K_m were determined by fitting the k_{obs} data to the Michaelis−Menton Equation: $k_{obs} = k_{cat}*[S]/(K_m + [S])$ where k_{cat} is the first-order rate constant, [S] is the dienophile substrate concentration, and K_m represents the substrate concentration at which half the maximum rate is observed. All RNA sequences surveyed demonstrated dienophile saturation kinetics.

Apparent K_i Determination. Apparent K_i values for the endocycloaddition product **6** were determined at 500 μM **2** by fitting the observed first-order rate constants to the following equation for inhibition: $k_{obs} = (k_{obs0}/2)\{\alpha E - I - K_i + [(K_i + \alpha E - I)^2 + 4K_i]^{1/2}\}$ 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 represents the fractional (α) concentration of functional active sites (E), I is the concentration of **3**, and K_i is the apparent inhibition constant.[42]

Alternative Substrate Experiments. Biotin maleimides **2** and **4** were purchased from Pierce, while **5** was purchased from Sigma. All were used without further purification. Incubations were conducted as described in the kinetics section. The data displayed in Table 3 was generated with dienophile **2**, **4**, or **5** (20 μM) and Diels−Alderase (500 nM) at 25 °C, while the competition experiment was conducted with **2** (20 μM), **4** (500 μM), and Diels−Alderase (500 nM) at 25 °C.

DA22 dsDNA Template Truncation. PCR amplification was performed under hot start conditions in 100 μL with DA22 dsDNA template (10 pmol), dNTP's (200 μM each), KOD XL DNA polymerase (2.5 U), 1X SYBR green I, 5′-primer (1 μM) and either 3′-primer (1 μM) (5′-GCACGCGTTAGAAACGGACACG (DA_TR96) or 5′-ACCCACTAGGGATGGGTTCATG (DA_TR119). Cycle parameters for the PCR were 96 °C for 15 s, followed by 68 °C for 30s. All DNA was purified by electrophoresis on a 6% polyacrylimide gel, and the excised bands were isolated by the freeze-squeeze technique. The purified dsDNA templates were transcribed, ligated, and assayed for their Diels−Alderase activity as described vide supra for all the other RNA catalysts.

Acknowledgment. We wish to thank all of the talented people at NeXstar Pharmaceuticals and Invenux who contributed to our work on RNA catalysis. A special thanks to Jeff Carter at SomaLogic for preparation of truncation primers.

JA0494149

(42) Williams, J. M.; Morrison, J. F. *Methods Enzymol.* **1979**, *63*, 437−467.