# RNA Aptamers That Bind Flavin and Nicotinamide Redox Cofactors

**Charles T. Lauhon and Jack W. Szostak***

*Contribution from the Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114*

*Received September 2, 1994*[⊗]

**Abstract:** RNA molecules that specifically bind riboflavin (Rb) and β-nicotinamide mononucleotide (NMN) have been isolated by *in vitro* selection. A simple structural motif containing intramolecular G-quartets was found to bind tightly to oxidized riboflavin ($K_d$ = 1−5 μM). DNA versions of the consensus sequence also bind, but with weaker affinity. DMS protection experiments show that the quartet structure of these aptamers is stabilized by interaction with the flavin. As a measure of their redox specificity, the aptamers do not show significant differential binding between oxidized and reduced forms of a 5-deazariboflavin derivative that is a close structural analog of riboflavin. In contrast to the lack of redox specificity of the riboflavin aptamers, RNAs selected for binding to the nicotinamide portion of NAD discriminate between NAD and NADH in solution by over an order of magnitude. A mutagenized pool based on one of the NMN aptamer sequences was used to reselect for NMN binding. Comparison of the reselected sequences led to the identification of the binding region of the aptamer. A complex secondary structure containing two interacting stem-loops is proposed for the minimal NMN-binding RNA. The same mutagenized pool was used to select for increased discrimination between NMN and NMNH. From these reselected sequences, a mutation within the binding region was identified that increases specificity for NMN. These experiments show that RNA can bind these cofactors with low micromolar affinity and, in the case of nicotinamide cofactors, can discriminate between the two redox states. These cofactor binding motifs may provide a framework for generating new ribozymes that catalyze redox reactions similar to those found in basic metabolic pathways.

## Introduction

The discovery of RNA molecules with catalytic activity[1,2] has considerably strengthened hypotheses[3] of an early stage of life in which catalytic functions were largely fulfilled by RNA molecules or ribozymes. The current prevalence of metabolic coenzymes that contain nonreacting, structural components of RNA has been a major supporting observation for this hypothesis, in that these compounds may be the remnants of the transition from an earlier "RNA world".[4,5] The role of RNA in protein synthesis and the fact that 2′-deoxyribonucleotides are synthesized solely from ribonucleotides suggest a primary position for RNA in the history of both biocatalysis and genetics. Because much of the "fossil record" of primordial metabolism has been lost due to evolutionary drift, further support for these hypotheses must come from experimental assessment of the catalytic versatility of RNA. The recent application of *in vitro* selection[6] to isolate nucleic acid molecules with unique binding and catalytic properties thus has direct application in evaluating the feasibility of ribozyme-based life.

*In vitro* selection techniques using both RNA and DNA have produced a number of nucleic acid molecules with interesting properties. These include tight binding to proteins such as thrombin[7] and HIV reverse transcriptase[8] as well as specific binding to small molecules such as the amino acids D-tryptophan[9] and L-arginine[10,11] and ribose-containing cofactors such as ATP,[12] cyanocobalamin,[13] and, recently, FMN and FAD.[14] *In vitro* selection has also been utilized to alter the catalytic activity of existing ribozymes[15,16] and to find new RNA catalysts.[17,18] In many cases, binding discrimination approaches that displayed by protein enzymes. For example, the RNA aptamer for ATP binds at least 100-fold less tightly to 2′-deoxy-ATP, while an RNA aptamer for theophylline shows a $10^4$-fold lower affinity for caffeine, which has a methyl group substituted for a hydrogen at N7.[19]

Reduction−oxidation (redox) reactions comprise nearly one quarter of the total number of classified enzymatic reactions in modern primary metabolism.[20] The organic compounds that mediate the majority of these reactions are the flavin (FMN and FAD) and nicotinamide (NAD and NADP) cofactors (Figure 1). Both of these cofactors have RNA components in their structure; in fact, they are essentially modified ribonucleotides. Thus, from a prebiotic evolutionary standpoint, it is interesting

[⊗] Abstract published in *Advance ACS Abstracts,* January 1, 1995.

(1) Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech. T. R. *Cell* **1982**, *31*, 147−157.
(2) Guerrier-Takado, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. *Cell* **1983**, *35*, 849−857.
(3) Gilbert, W. *Nature* **1986**, *319*, 618.
(4) White, H. B., III *J. Mol. Evol.* **1976**, *7*, 101−104.
(5) Benner, S. A.; Ellington, A. D.; Traver, A. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 7054−7058.
(6) (a) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818−822. (b) Turek, C.; Gold, L. *Science* **1990**, *249*, 505−510. (c) Szostak, J. W. *Trends Biochem. Sci.* **1993**, *17*, 89−93.
(7) Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J. *Nature* **1992**, *355*, 564−566.
(8) Tuerk, C.; MacDougal, S.; Gold, L. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 6988−6992.
(9) Famulok, M.; Szostak, J. W. *J. Am. Chem. Soc.* **1992**, *114*, 3990−3991.
(10) Famulok, M. *J. Am. Chem. Soc.* **1994**, *116*, 1698−1706.
(11) Connel, G. J.; Illangeskare, M.; Yarus, M. *Biochemistry* **1993**, *32*, 5497−5502.
(12) Sassanfar, M.; Szostak, J. W. *Nature* **1993**, *364*, 550−552.
(13) Lorsch, J. R.; Szostak, J. W. *Biochemistry* **1994**, *33*, 973−982.
(14) Burgstaller, P.; Famulok, M. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 1084−1087.
(15) Green, R.; Szostak, J. W. *Science* **1992**, *258*, 1910−1915.
(16) Pan, T.; Uhlenbeck, O. *Biochemistry* **1992**, *16*, 3887−3895.
(17) Bartel, D. P.; Szostak, J. W. *Science* **1993**, *261*, 1411−1418.
(18) Lorsch, J. R.; Szostak, J. W. *Nature*, in press.
(19) Jenison, R. D.; Gill S. C.; Pardi, A.; Polisky, B. *Science* **1994**, *263*, 1425−1429.
(20) White, H. B., III In *Pyridine Nucleotide Coenzymes*; Academic Press, Inc.: New York, 1982; Chapter 1.

0002-7863/95/1517-1246$09.00/0 

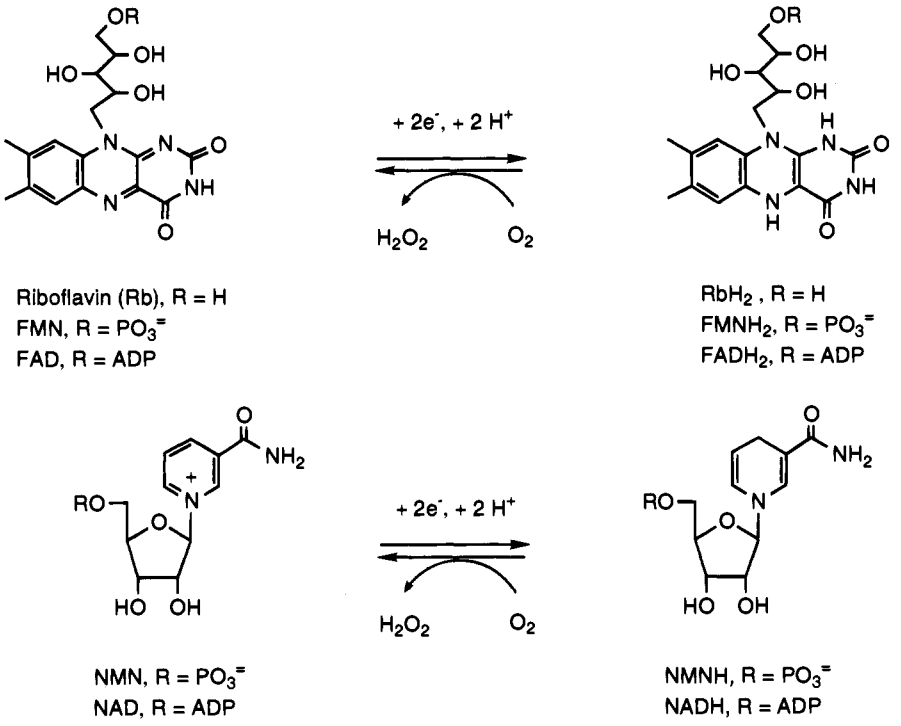



**Figure 1.** Structure of flavin and nicotinamide cofactors.

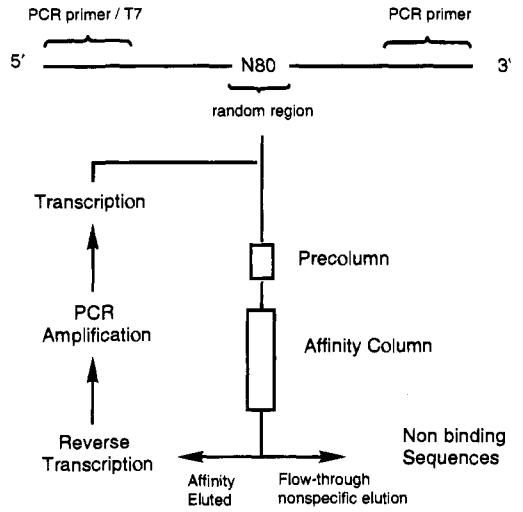



**Figure 2.** Selection scheme used for obtaining RNAs that bind small molecule ligands.

to determine how these cofactors may interact with RNA. Here we describe the selection of RNA molecules that specifically bind riboflavin and NMN.

## Results and Discussion

The basic protocol for aptamer selection experiments has been described in detail[12,13] and is shown schematically in Figure 2. Selection for RNAs that bind to a small molecule is achieved by using an affinity column containing the immobilized ligand.[6] A precolumn is used to select against those sequences that bind to the support matrix or to the linker arm and typically contains the underivatized support or the support derivatized with the linking moiety. The site of attachment of the ligand to the matrix can be used as an element of design to favor solvent accessibility of this site when the ligand is bound to the selected binding pocket. Binding specificity can also be tailored by eluting nonselective RNAs with structural analogs of the ligand or (less efficiently) by using additional precolumns. Using these methods, we have selected for RNA molecules that are able to bind the oxidized forms of riboflavin (Rb) and nicotinamide mononucleotide (NMN) (Figure 1).

**Selection For RNAs That Bind Oxidized Riboflavin.** The affinity ligand chosen for the oxidized riboflavin selection was N3-carboxymethyl riboflavin[21] (1), attached at the carboxyl group via a hydrazide linkage to adipic dihydrazide-modified agarose using a water soluble carbodiimide (EDC) as the coupling reagent (Figure 3). The yield of the coupling reaction was proportional to the concentration of 1 over a wide range and could be used to control the riboflavin concentration on the matrix, from 1−2 $\mu$mol/mL gel used in selection rounds 1−5, to 0.3 $\mu$mol/mL gel in later rounds to select for RNAs with greater affinity. The initial pool[16] of RNA contained a random sequence region of 72 nucleotides (nt) flanked by constant regions of 20 nt for PCR primer binding sites and had a complexity of $5 \times 10^{14}$ unique molecules. After five rounds of enrichment, 10% of the RNA was bound and eluted with riboflavin. The calculated complexity of the enriched pool was relatively high ($\sim 10^6$ molecules), based on the amount of purification from each round. Therefore, further rounds of selection were used to select for the most specific RNAs. Selection against nonspecific binding to aromatic compounds was achieved by initial elution with quinoxaline-2-carboxylic acid (**2**, Figure 3) before elution with riboflavin (rounds 6 and 7). By round 9, 30% of the RNA bound specifically to riboflavin and was retained on a column of 0.3 mM ligand concentration for at least 20 column volumes. Additional rounds of selection using lower concentrations of immobilized riboflavin (0.05 $\mu$mol/mL) failed to give further enrichment for RNAs of greater affinity.

RNA from the round 9 pool was cloned and sequenced. The majority of sequences revealed a common G-rich motif flanked by bases of Watson−Crick complementarity (Figure 4). Analysis of individual clones indicated in every case that riboflavin binding required monovalent cation with a strong preference for potassium over sodium, while magnesium ion was not required. The pattern of guanosines in the G-rich loop and the binding dependence on monovalent cation suggested an in-

(21) Ying-Hsiueh, F.; MacKenzie, R. E.; McCormick, D. B. *Biochemistry* **1970**, *9*, 2219−2224.

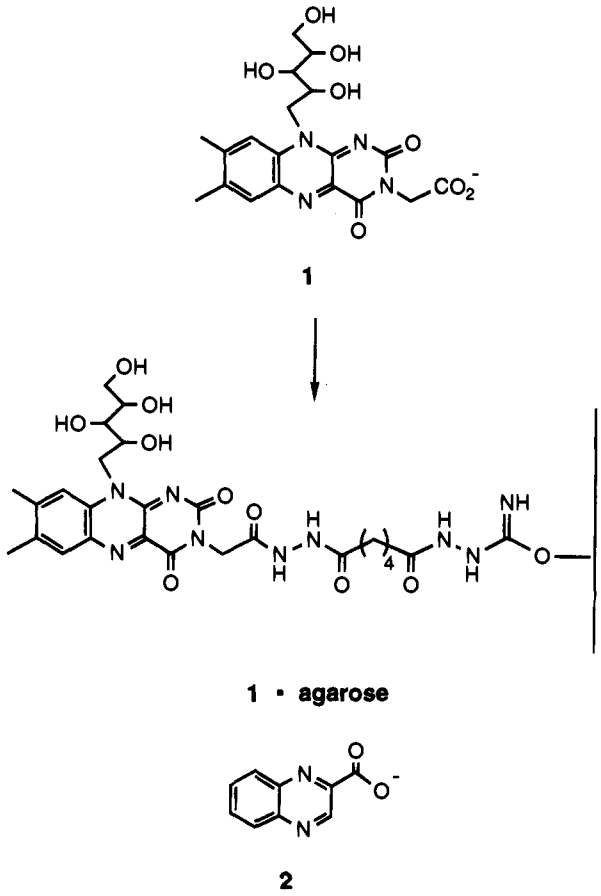



**Figure 3.** Preparation of immobilized N3-(carboxymethyl)riboflavin (**1**).

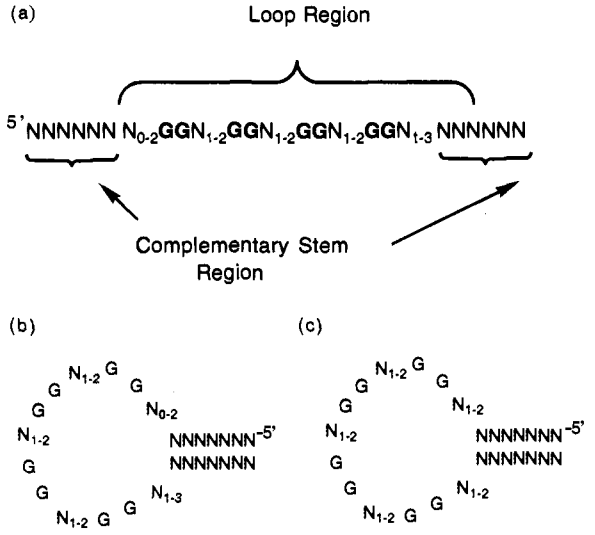



**Figure 4.** (a) Consensus sequence for riboflavin aptamers based on sequences of round 9 pool, (b) secondary structure of consensus sequence, and (c) structure of randomized pool based on major subset of the consensus sequence.

tramolecular G-quartet structure,[22] such as that shown in Figure 5. RNA G-quartet tetraplex structures are known,[23] although there have been no reports of either intramolecular RNA G-quartets or quartet structures in which the loop regions joining

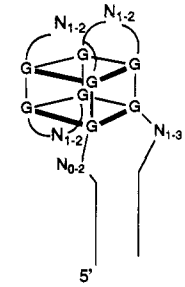



**Figure 5.** Proposed tertiary fold of Rb aptamer consensus sequence into intramolecular G-quartet.

the guanosines contained fewer than two nucleotides. An interesting feature of the proposed structure is that only a single nucleotide spans one of the "wide grooves" [22b] of the G-quartet and likely results in a compact or even distorted quartet structure. Inspection of the sequences of clones from the round 9 pool indicates a preference for a single base at this position, whereas for the other two loop positions there is no obvious preference.

Dimethyl sulfate (DMS) protection experiments[24] with representative clones showed no protection of the N7 atoms of guanosines in selection buffer without riboflavin (Figure 6a). However, in the presence of near-saturating concentrations of riboflavin ($\sim 50 \times K_d$), the guanosines in the putative quartet region showed increased protection from methylation by DMS, whereas those in the stem and loop regions joining the quartet guanosines showed no increased protection. This protection pattern was not observed with riboflavin in sodium buffer, in which the RNA does not bind as shown by both affinity column chromatography and by fluorescence assay ($K_d > 100\ \mu M$, see below). Further evidence for a G-quartet structure was found by synthesizing a three-tiered molecule with an additional guanosine added to each G doublet in the loop. This molecule binds riboflavin with essentially the same $K_d$ value as the original structure and is exceedingly stable, showing complete methylation protection and similar riboflavin affinity in both potassium and sodium buffers (data not shown). DMS protection of the three-tiered quartet shows that it is folded even in relatively low (10 mM) concentrations of $Li^+$ as the only abundant monovalent cation. This lack of monovalent cation specificity differs from intramolecular DNA G-quartets, which normally show greater stability in $K^+$ than in $Na^+$.

When added to riboflavin in binding buffer, the aptamer RNAs were found to completely quench riboflavin fluorescence. $K_d$ values were obtained by measuring the fluorescence quenching as a function of RNA concentration and were on the order of $1-5\ \mu M$, which was consistent with $K_d$ estimates based on analysis of the column binding of the pool RNA.[25] A Scatchard plot indicates that there is only one flavin binding site for both the two-tiered and three-tiered RNAs (Figure 7). The high wavelength maximum of the UV spectrum of RNA-bound riboflavin is redshifted by 5 nm (from 443 to 448 nm) relative to free riboflavin, and the complex exhibits a weak shoulder at 482 nm. This shift can be compared to that of a monoclonal antibody that binds oxidized flavin tightly in which a redshift of 15 nm and a significant shoulder at 482 nm is observed in the bound complex.[26] The aptamer-induced shift is equal to the wavelength shift in FAD ($\lambda_{max} = 448$), which is known to exist predominantly in a stacked conformation. Solutions of FAD also show much weaker fluorescence when compared to

(22) For references regarding the structure of DNA intramolecular g-quartets, see: (a) Macaya, R. F.; Schultze, P.; Smith, F. W.; Roe, J. A.; Feigon, J. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3745–3749. (b) Wang, K. Y.; McCurdy, S.; Shea, R. G.; Swaminathan, S. S.; Bolton, P. H. *Biochemistry* **1993**, *32*, 1899–1904.

(23) (a) Kim, J.; Cheong, C.; Moore, P. *Nature* **1991**, *351*, 331–332. (b) Cheong, C.; Moore, P. *Biochemistry* **1992**, *31*, 8406–8414.

(24) (a) Peattie, D. A.; Gilbert, W. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 1760–1764. (b) Peattie, D. A.; Gilbert, W. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 4679–4682.

(25) Arnold, F. H.; Schofield, S. A.; Blanch, H. *J. Chromatogr.* **1986**, *355*, 1–12.

(26) Shokat, K. M.; Lehmann, C.; Schultz, P. G. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 1172–1174.

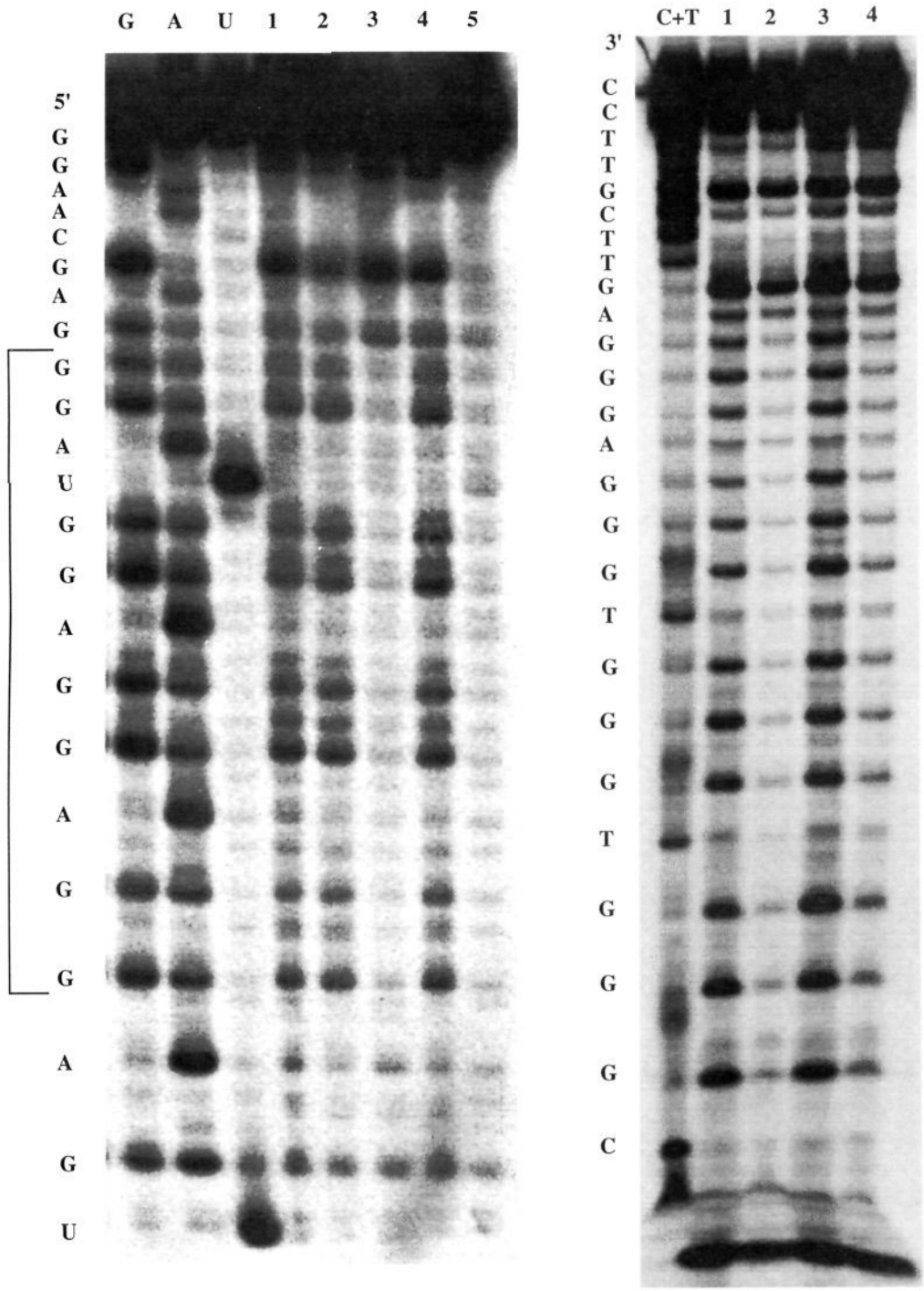



**Figure 6.** DMS protection of guanosine N7′s with and without added riboflavin. (a, left) Originally selected two-tiered RNA of sequence 5′-GGAACGAGGGAUGGAGGAGGAGUCGUUCC-3′: lane 1, DMS in $K^+$; lane 2, DMS in $Na^+$, lane 3, DMS in $K^+$ with 250 $\mu$M Rb (50 × $K_d$); lane 4, DMS in $Na^+$ with 250 $\mu$M Rb; lane 5, aniline cleavage only. The putative quartet region is indicated by a bracket. (b, right) Three-tiered DNA aptamers of sequence 5′-GGAACGACGGGUGGGUGGGAGGGACUCGUUCC-3′: lane 1, DMS in $Li^+$ with 200 mM Rb (75 × $K_d$); lane 2, DMS in $K^+$; lane 3, DMS in $Na^+$; lane 4, DMS in $Na^+$ with 200 $\mu$M Rb (75 × $K_d$).

riboflavin or FMN, an observation attributed to quenching by the adenine group. Thus, it is likely that the riboflavin is bound to these aptamers in a stacked orientation, possibly with a guanosine in the quartet. Such an interaction could give the observed stabilization of the quartet structure by riboflavin.

In order to assess the role of the RNA ribose 2′-hydroxyls in binding to riboflavin, DNA oligonucleotides based on two of the original two-tier consensus sequences were synthesized. These DNAs were found to bind riboflavin, but with $K_d$ values over an order of magnitude greater than those of the analogous RNA molecules. Three-tiered DNA quartets with this motif are significantly more stable and bind with affinities nearly identical to those of the corresponding three-tiered RNA in $K^+$ buffer. Their ion dependence differs from the RNA, however, in that they are less stable in $Na^+$ buffer (Figure 6b). Similar to the originally selected two-tiered RNA aptamers, the stability of the three-tiered DNA structures in $Na^+$ buffer can be increased with added riboflavin. The $K_d$ values for both the two and three-tiered RNA and DNA versions of aptamers derived from two selected clones are shown in Table 1. This data provides further support for the existence of G-quartets in these aptamers and suggests that the 2′-hydroxyl groups of the

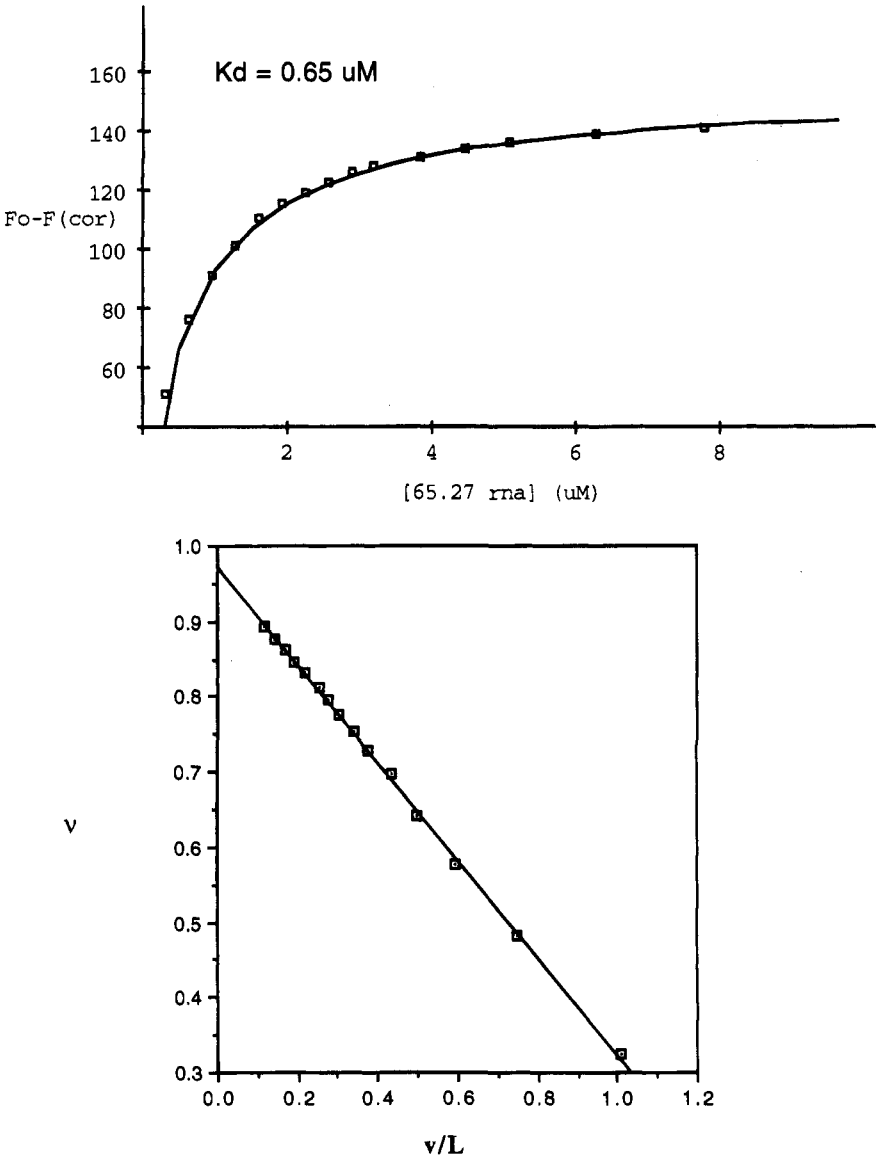



**Figure 7.** (a) Nonlinear regression analysis of fluorescence binding curve of three tiered Rb RNA aptamer (sequence (b) in Table 1). (b) Scatchard plot of same data showing a single RNA-riboflavin binding site.

**Table 1:** Solution $K_d$'s of RNA and DNA Two- and Three-Tiered Aptamers Based on Two Selected Loop Sequences

| 5'-...G(G)$_n$AU(G)$_n$A(G)$_n$A(G)$_n$AG...[a] | | | |
|---|---|---|---|
| | | $K_d$[b] ($\mu$M) | |
| | | K | Na |
| 2-tier | RNA | 4.5 | >100 |
| ($n = 2$) | DNA | ~160 | >100 |
| 3-tier | RNA | 3.8 | 5.6 |
| ($n = 3$) | DNA | 3.3 | >100 |
| 5'-...C(G)$_n$U(G)$_n$U(G)$_n$A(G)$_n$AGA...[a] | | | |
| | | $K_d$[b] ($\mu$M) | |
| | | K | Na |
| 2-tier | RNA | 2.8 | >100 |
| ($n = 2$) | DNA | 60 | >100 |
| 3-tier | RNA | 1.2 | 2.7 |
| ($n = 3$) | DNA | 1.2 | 2.8 |

[a] Stem sequence is 5'-GGAACGA-loop-UCGUUCC-3'.

RNA do not interact directly with the ligand but are important for the stability of the RNA structures.

The binding specificity of these molecules was assessed using a column binding assay (Figure 8). The originally selected RNAs bind to 3-carboxymethyl-5-deazariboflavin (**5**, see below) and riboflavin with equal affinity (confirmed by fluorescence quench measurements) but do not bind tightly to simple single ring or fused ring aromatics such as benzoate or quinoxalines. Larger aromatic ring systems, such as benzyl viologen, also do not elute bound RNA. Uracil is effective in eluting bound RNA at a concentration 5-fold higher than the riboflavin concentration on the column, indicating that the RNAs may interact with the carbonyl groups of the isoalloxazine ring. Ribitol is found to elute an observable amount (1−2%) of bound aptamer at 5-fold the column concentration. This indicates a weak but significant interaction with the ribityl side chain, consistent with the observation that these aptamers will not bind to a column in which the riboflavin is attached via the ribose hydroxyls, although the 5'-modifications of FMN and FAD are tolerated.

Covalent linkage of the flavin at N3 does not necessitate solvent accessibility at N5 of the bound complex, which is the chemically reactive site of redox chemistry. In order to probe the solvent accessibility of this position, a 5-alkyl-5-deazaflavin (**9**) was prepared by the route shown in Scheme 1. Lewis acid cyclization of known deazariboflavin intermediate **7**[27] with 3-carbomethoxypropionyl chloride gave a modest yield of the desired tricycle **8** as well as recoverable starting material.

(27) Ashton, W. T.; Brown, R. D.; Tolman, R. L. *J. Het. Chem.* **1978**, *15*, 489−491.

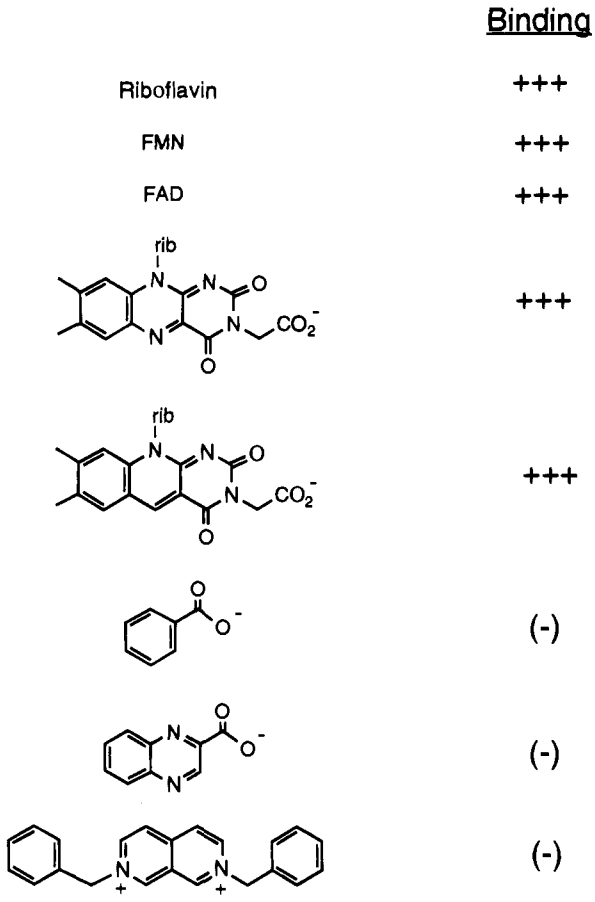



**Figure 8.** Binding specificity of round 9 pool Rb RNA based on column assay (see Experimental Section).

Longer reaction times led to the formation of byproducts that were difficult to remove. Alkaline hydrolysis of the acetates and methyl ester gave 5-(carboxyethyl)-5-deazariboflavin (**9**). Compound **9** was shown to bind to the RNA aptamers by fluorescence assay ($K_d$ = 1.7 $\mu$M for the three-tiered RNA in Table 1, bottom), indicating that the N5-position of the riboflavin is not sterically restricted. In addition, sulfite reactivity of the RNA-bound flavin was investigated. At sulfite concentrations below that which adds to pyrimidines (<200 mM), and RNA concentration 10-fold above the $K_d$, bound riboflavin was found to react to give a sulfite adduct. The dissociation constant of flavin–sulfite adducts has been shown to correlate with the redox potential of the flavin.[28] The dissociation constant for sulfite addition to aptamer-bound riboflavin was not significantly different from that with free riboflavin. This result suggests that the redox potential of the flavin is not significantly altered when bound to the RNA.

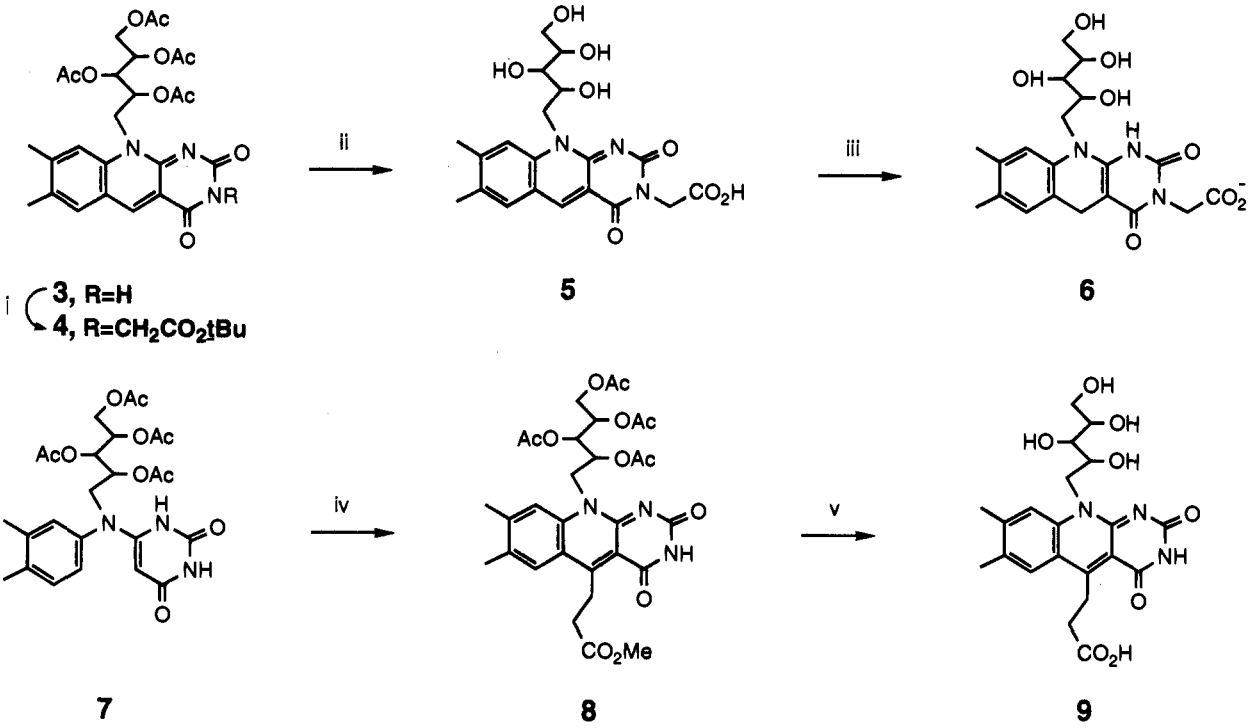
The ability to selectively bind oxidized vs reduced flavin was investigated using the 5-deaza analog **5**, which in the oxidized state is bound with the same affinity as riboflavin. N3-(Carboxymethyl)-5-deazariboflavin (**5**) was synthesized from 5-deazariboflavin[29] (Scheme 1) by a route analogous to that used for the 5-aza compound[21] and reduced with $NaBH_3CN$ to produce the 1,5-dihydro-5-deazaflavin **6**, a more oxygen-stable analog of reduced riboflavin.[42] It has been reported by Bruice and co-workers[30] that even careful reduction of some 5-deazaflavins with $NaBH_4$ results in ~50% overreduction to products that are not reoxidized by $O_2$. In our hands, $NaBH_3CN$ was a superior reagent for the specific reduction of 5-deazaflavins (notably those that are N3-alkylated) to the 1,5-dihydro forms. The reduced 5-deazaflavins are quantitatively reoxidized by $O_2$ back to the original oxidized structures (as observed by $^1H$ NMR—see Experimental Section). The RNAs selected for binding to oxidized riboflavin were found to bind well to immobilized **5** that has been reduced to the 1,5-dihydro form **6**. A similar percentage of aptamer RNA was bound and eluted with equal volumes of similar concentration of **6**, indicating that the difference in affinity between the two is small (<10-fold). This result again suggests that these aptamers do not possess the requisite binding specificity for differentiation between the oxidized and reduced forms of riboflavin itself.

Attempts to select for RNAs that preferentially bind to the oxidized form of flavin were made using a small randomized pool based on a major subset of molecules comprised by the original consensus sequence (Figure 4c). The pool was prepared at the DNA level using an oligonucleotide synthesizer by the following method. The 3′ primer binding site and the 3′ part of the stem were set as defined sequence. The loop region was begun by adding a random nucleotide. Half the resin of the synthesis column was removed, another random base was added, and then the removed resin was added back to the column for the addition of the first set of two G's. This step was repeated for each $N_1$-2 position in the loop to generate a pool of DNA molecules with varying loop size of 13–18 bases and a Watson–Crick stem 10 bp long, with constant regions for PCR and transcription into RNA. Cloning and sequencing confirmed that both the number and sequence of the $N_1$-2 positions were random. The calculated pool complexity was $3 \times 10^6$. After four rounds of selection using a precolumn of immobilized 1,5-dihydro-5-deaza-3-carboxymethylriboflavin, 3–4% of the RNA was specifically bound and eluted, although a large proportion of RNA remained bound to the precolumn as well. The ratio of RNA bound to the oxidized flavin column versus the reduced deazaflavin precolumn could not be increased in further rounds. Thus, it appears that although this simple RNA motif is capable of binding oxidized flavins tightly and specifically, it is unable to show enough specificity to distinguish between the different redox states.

**Selection for RNAs That Bind to NMN.** Nicotinamide cofactors are structurally closer to ribonucleotide bases than the flavins, having a true glycosidic linkage to the nitrogen of the pyridine "base", with a hydrogen bonding pattern similar to that of guanosine or adenosine, depending on the rotational configuration of the primary amide at C-3.[31] Addition of nucleophiles[32] has been shown to be favored chemically at C4, the biological site of reduction, although isomeric mixtures can arise during reduction with certain reagents such as borohydride ion.[33] The active biological cofactors NAD(H) and NADP(H) contain nicotinamide mononucleotide covalently attached to AMP by a 5′-5′ pyrophosphate linkage (Figure 1). NMN itself has been shown to be an alternative substrate for at least one dehydrogenase, but with a significantly lower V/K value[34] (the value of V with NMN is increased by the addition of adenosine). It is not clear *a priori* why NMN is not utilized by enzymes for redox reactions since it has essentially the same redox potential as NAD(P).[35] It is known that many dehydrogenases contain a

(28) Muller, F.; Massey, V. *J. Biol. Chem.* **1969**, *247*, 4007–4016.
(29) Janda, M.; Hemmerich, P. *Angew. Chem., Int. Ed. Engl.* **1976**, *15*, 443–444.
(30) Chan, R. L.; Bruice, T. C. *J. Am. Chem. Soc.* **1978**, 7375–7382.
(31) Visser, C. M. *Origins of Life* **1982**, *12*, 165–179.
(32) (a) Everse, J.; Zoll, E. C.; Kahan, L.; Kaplan, N. O. *Bioorg. Chem.* **1971**, *1*, 207–233. (b) Johnson, L. L.; Smith, K. W. *Biochemistry* **1976**, *15*, 553–559.
(33) Chaykin, S.; King, L.; Watson, S. G. *Biochim. Biophys. Acta* **1966**, *124*, 13–25.
(34) Sicsic, S.; Durand, P.; Langrene, S.; Le Goffic, F. *FEBS Lett.* **1984**, *176*, 312–324.
(35) Srinivasan, R.; Fischer, H. *Arch. Biochem. Biophys.* **1983**, *223*, 453–457.

**Scheme 1.** Synthesis of 5-Deazariboflavin Analogs[a]




[a] (i) $BrCH_2CO_2tBu$, $K_2CO_3$, DMF, 50 °C, 80%; (ii) (a) TFA, 0 °C, quant., (b) 0.1 N NaOH, 0 °C, 65%; (iii) $NaBH_3CN$, 0.1 M KPi, pH 6.0, 4 h, >95%; (iv) $ClOC(CH_2)_2CO_2CH_3$, $SnCl_4$, $ClCH_2CH_2Cl$, reflux, 55% (based on rec. sm); (v) 20% $CH_3CN$ in 0.1 N NaOH, room temperature, dark, 12 h, 55%.

highly conserved region for binding the dinucleotide cofactors,[36] and this may be evidence for a strong evolutionary bias for using the dinucleotides. Because the affinity of most dehydrogenases for NAD(P) is only in the low micromolar range, it is unlikely that the dinucleotide is required merely for increased affinity.

We therefore decided to select for RNAs that bind NMN, by applying random sequence RNA to a column of immobilized NAD and then eluting with NMN, and thus selecting primarily for interaction with the nicotinamide ribose portion of the molecule. For this selection, a new pool of DNA was synthesized containing a random region 80 nt long, with a complexity of $4 \times 10^{15}$ molecules. Many sequences were found that could bind in a nonspecific manner to the aromatic portion of NAD, as observed in the previous selection with oxidized riboflavin. Elution as before with a high concentration (10 mM, 3× the concentration of NAD on the matrix) of quinoxaline-2-carboxylic acid (**2**) eventually eliminated these nonspecific molecules, so that by round 8 nearly 50% of the RNA was eluted only by NMN and not by **2**. Cloning and sequencing of 14 of the clones revealed one major (11 of the 14) and one minor clone (1 out of the 14) that bound NAD and were specifically eluted by NMN. Both clones showed 70% binding and specific elution by affinity column assay, but the major clone was more efficiently transcribed by T7 RNA polymerase and was thus more abundant at each round of selection.

Both clones show very similar binding specificity, although there is no observable similarity of primary or secondary structure. The binding specificity was determined by column binding assay (Figure 9) and shows that the major binding determinant resides in the positively charged nicotinamide ring. *N*-methyl nicotinamide elutes NAD-bound RNA nearly as well as NMN and significantly better than nicotinamide, indicating that the positively charged nitrogen is an important recognition element. The amide $NH_2$ of NMN is more important for binding than the carbonyl oxygen, based on the greater ability of thio-NAD to elute bound RNA compared to acetylpyridine adenine dinucleotide (APAD) or nicotinic acid mononucleotide (NAMN). It is possible that the negative selection with the quinoxaline derivative **2**, which has a structure similar to the base portion of NAMN, contributed to the observed specificity against NAMN.

Both clones bind NAD in preference to NADH as shown by column binding competition. NAD at a concentration equal to that of NAD on the column will elute 5-fold more bound RNA than will an equal concentration of NADH. However, there is very little difference between the amounts of NAD-bound RNA eluted by NMN versus NMNH. One explanation for this specificity is the observation based on NMR experiments that the nicotinamide ring is significantly puckered in the stacked conformation of NADH but is planar in NAD and NMNH.[37] Thus, these aptamers may distinguish between a planar and puckered nicotinamide ring conformation.

The minor clone was chosen for further study because of its slightly better stability in terms of amount of RNA bound to an NAD column and because it showed greater specificity for NAD over NADH. Dissociation constants in solution were measured for RNA binding to the dinucleotides by equilibrium gel filtration with $^{14}C$-labeled NAD and enzymatically prepared $^{14}C$- or $^{32}P$-NADH.[38] The minor clone binds to NAD with a $K_d$ of $2.5 \pm 0.5$ $\mu$M, compared to a value of $37 \pm 3$ $\mu$M for NADH (Figure 10). Many dehydrogenases[39] and other reductases[40]

(36) Rossman, M. G.; Moras, D.; Olsen, K. W. *Nature* **1979**, *277*, 598–600.

(37) Oppenheimer, N. J.; Arnold, L. J.; Kaplan, N. O. *Biochemistry* **1978**, *17*, 2613–2619.

(38) Viola, R. E.; Cook, P. F.; Cleland, W. W. *Anal. Biochem.* **1979**, *96*, 334–340.

(39) (a) For horse liver alcohol dehydrogenase (SS isozyme): Ryzewski, C. N.; Pietruszko, R. *Biochemistry* **1980**, *19*, 4843–4848. (b) For D-β-hydroxybutyrate dehydrogenase: Fritzsche, T. M.; McIntyre, J. O.; Fleischer, S.; Trommer, W. E. *Biochim. Biophys. Acta* **1984**, *791*, 173–185.

(40) For dihydrofolate reductase: Adams, J. A.; Fierke, C. A.; Benkovic, S. J. *Biochemistry* **1991**, *30*, 11046–11054.

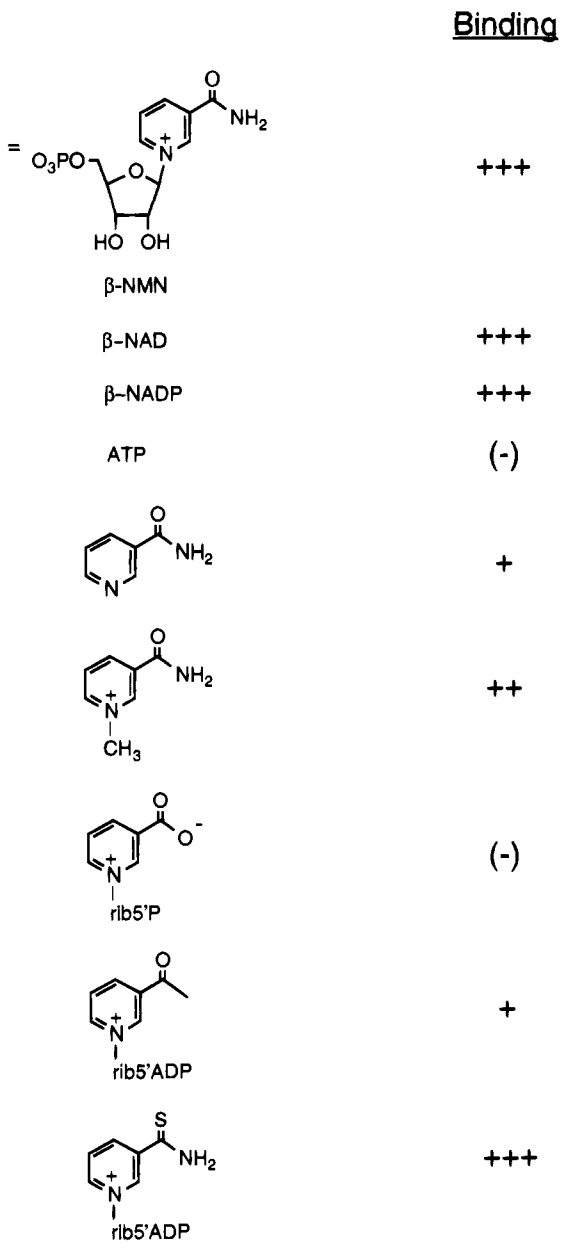



**Figure 9.** Binding specificity of single NMN aptamer based on column assay.

have $K_d$ ratios for NAD/NADH binding that are of similar magnitude (i.e., <100).

To elucidate the bases in the aptamer sequence that are important for binding NMN, a mutagenized pool was prepared based on the sequence of this clone. Each position of the random region of the aptamer sequence was mutagenized to give a level of 30% mutagenesis at each base by automated oligonucleotide synthesis using mixtures of amidites. The DNA pool was amplified by PCR and transcribed to yield an RNA pool with a complexity of approximately $1 \times 10^{15}$ molecules which was then used for reselection for NMN binding. After five rounds of selection, the enriched pool showed 47% specific binding to an NAD column and elution with NMN, and RNAs from this pool were cloned and sequenced. The sequences of these clones were aligned with the sequence of the parent clone to determine the regions of the molecule that were critical for binding. An internal region of about 65 nt was relatively well conserved (Figure 11). A small RNA molecule corresponding to this stretch of the parent sequence was found to bind to an NAD column and was specifically eluted by NMN. Only one-third as much of the small RNA bound to and specifically eluted

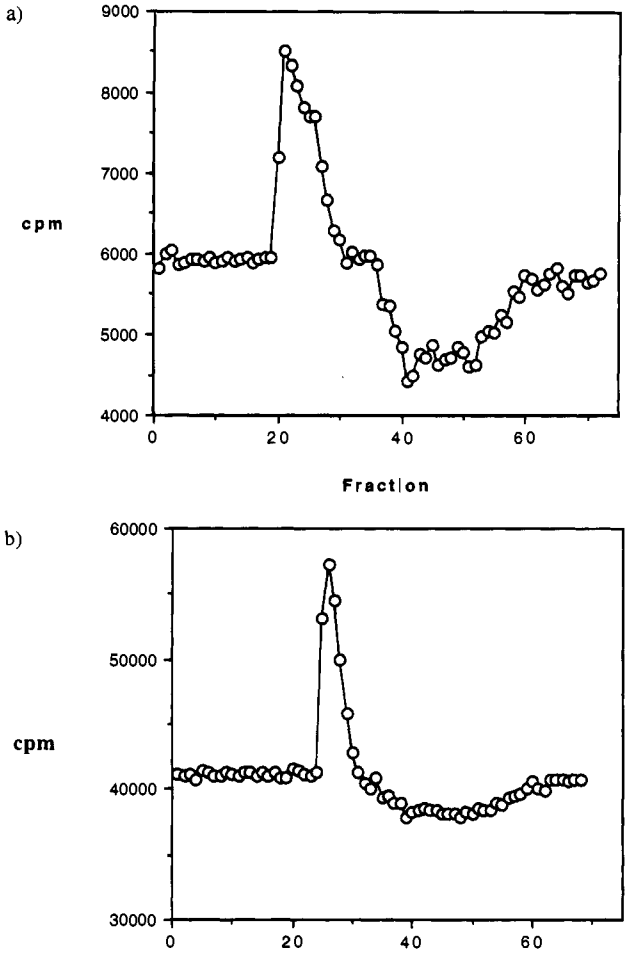



**Figure 10.** Equilibrium gel filtration measurement of solution $K_d$'s of NMN aptamer with (a) 1 $\mu$M $^{14}$C-NAD and 1 nmol RNA ($K_d$ = 2.5 $\mu$M) and (b) 0.5 $\mu$M $^{14}$C-NADH and 5 nmol RNA ($K_d$ = 37 $\mu$M).

from an NAD column relative to the full length clone (22% vs 70%), although this amount could be increased to 37% by increasing the $Mg^{2+}$ concentration from 5 to 20 mM. The ligand specificity of this shorter RNA was found to be the same as the full length clone, suggesting that the smaller RNA contains all of the bases that interact with the ligand but is not as structurally stable as the parent molecule.

Analysis of the sequence variations of the reselected pool indicated an abundance of covariations in the conserved region indicative of base-paired regions. A secondary structure model based on these covariations is shown in Figure 12a. Separate molecules corresponding to each of the putative stem loop regions do not bind to the ligand, suggesting that tertiary interactions between bases in each stem-loop (either with each other or with the ligand) are important for binding and/or stability of the folded structure. There are many observed covariations in the larger stem, while the sequence of the shorter stem is strictly conserved. There are also many clones with deletions that decrease the loop size in the smaller stem loop, while in every case preserving the two CG base pairs of the stem. It is likely that the smaller tetraloops are significantly more stable than the larger loop in the original clone.[41] Although there were no corroborating covariations in the smaller stem, small RNAs in which the two cytosines of this stem were omitted did not bind observably to an NAD column. There were also no covariations observed that were indicative of a pseudoknot structure involving pairing of bases in the loops.

The mutagenized pool was also used to reselect for new variants with increased discrimination between the oxidized and

(41) Uhlenbeck, O. C. *Nature* **1990**, *346*, 613–614.

```
WT   TCTCGAGTTATTCCAACCAGCCCAACTAGGCGTTTGAGGGGATTCGGCCACGGTAACAACCCCTCGGCTTGCGGATTAT

N1   ------------T--TAA-AA--T---G--------G-----------------------------T-------TACG-
N2   -TG------GG---CC-   AG-CCAC-----------------------------------------T------TG--
N4   -------------------A---------------------- ---------------------------A-T-C--G-
N6   --CGCT----G---------A----T---------CG-------------------------CG-TG--A---TG---
N8     --CTAG-CGG--G-GT-C---G----C-T--AA--A---C-----T-T-A-----G---A--------- -------
N9   -TCGCG--- --------T--- ----C----------------------T-------------------CG-TG-GTC
N10  G-A-------C------T- ------A---------G----G-A-------------------C---------T---G-
N13  C--------T-------------G--C----- -------T---GC---G----TT---GG-T--G-C------TC-CGT
N14  ---G----G------ ------- - C---------GC---GA-------T-----------N------T-----GT-
N15  ----A----T--- --  -------A-T-------------------------------------------C--A--
N16  -G-T---C-T-----T---T-----T----------------G-----------------------------A--C-C
N18  --A-A----TG   -----AG--T--- ---------------------------------------------A------
N19  -G--T----G-----C---A-----T---------------G-------------------------A--T--TG-CT-
N22  ---G-C-AGTAG----    ------------------------------------------------------C--C
N25  -----T---G-C------      ---A-------------------------------------C-A-G--G-C----
N101 -----T---G-C------      ---A---------------A-------------------------A-CT-GT-C
N102 --G------GG-------TT-G-----T----------------------C--------T-C-T----A-
N103 ----C-AGCT-GT  ---        ----------AG-------------G-A--------------A----GAT----
N105 ----A-A-CCCC---T-TGCGG---------------------GT-T--T---C----C-T--C-A-A----G---T--
N118 ------TC-T----G--------------CGC--G-C---T-AA-TCAA------G---A-----------TC-G----
N120 CAG-ACT-------C--GG-A------T-C------C--CC----T-A---TT----GC-G--G---T-----A-C-G-
N123 ---------------  --  T-----AA--------------G-------------------------C---C-GTC
N124 ---GC--A----A---GATC----C-----------C-------------CA---------G--CTT-CGGT-C----
N125 CTA-----------T-G-T  ------T-------------------------------------A--A-TT--G--
```

**Figure 11.** Alignment of sequences obtained from mutagenized pool reselected for NMN binding. A relatively conserved region used for further studies is shown in italics.

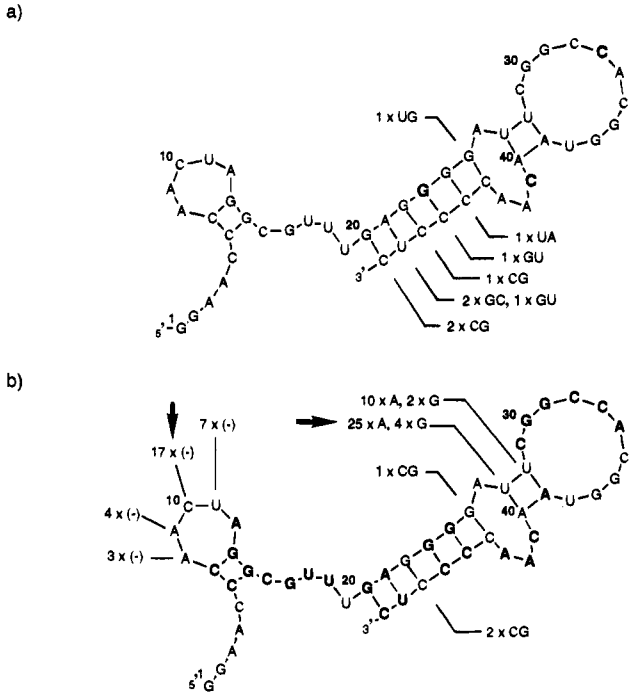



**Figure 12.** (a) Proposed secondary structure of shortened NMN aptamer based on conserved region of sequences in Figure 11. Covariations in larger stem are shown (i.e., 2 × CG indicates two clones found with CG bp at this position, etc.) and strictly conserved residues are in bold. (b) Structure of short NMN aptamer with two major conserved mutations (shown by arrows) obtained upon reselecting for differential binding between NMN and NMNH. Deletions are indicated by (—).

the reduced cofactor at the mononucleotide level. This selection involved eluting first with NMNH to remove molecules that showed no selective binding and then eluting with NMN as before. Negative selection with NMNH was possible because, unlike flavins and 5-deazaflavins,[42] the disproportionation of reduced and oxidized nicotinamides is relatively slow.[43] After seven rounds of this selection, nearly 40% of the RNA eluted with NMN, while only 10% eluted with NMNH. The fact that this selection required two more rounds of enrichment relative to the initial reselection for binding suggests that there are fewer sequences in the mutagenized pool that possess the additional binding specificity. Cloning of RNAs from this pool and testing of individual clones revealed RNA molecules that showed significant selective binding for oxidized NMN as determined by column assay.

Inspection of the sequence alignments of this enriched pool (Figure 13) indicated the same generally conserved area as before, with a few notable changes in the sequence of the binding region. These changes are shown in the context of the proposed secondary structure in Figure 12b. Specifically, a uridine residue (U27 in the short RNA) that is nearly completely conserved in the reselection for NMN binding is mutated to a purine (mostly A) in nearly every reselected clone. In cases where it is unchanged, the adenosine residue opposite it in the putative short stem, A40, is mutated to a uridine or guanosine. The adjacent uridine (U28) in this stem is also mutated in nearly half of the sequences. These mutations destroy potential base pairing and would allow for other interactions with bases in the smaller stem loop or with NMN itself. An RNA construct like that shown in Figure 12b containing the mutations of the reselected clone -N39 (Figure 13) was found to show greater specificity (threefold less elution of RNA from an NAD column with NMNH) with an A at position 27 than with U. Thus, the U to A mutation contributes to the observed selectivity for NMN. As was observed for some clones in the original doped pool selection, in every reselected clone that showed significant binding, there was a single nucleotide deletion in the smaller loop giving a presumably more stable tetraloop. Comparing shortened versions of -N39 with both the four- and five-base loop (with or without residue C10) revealed that this deletion does not affect the selectivity of NMN vs NMNH but does increase the fraction of bound RNA and hence is likely to stabilize this structure.

It is possible that the smaller loop is involved in a tertiary contact with the larger stem. Such an interaction has been postulated in the tertiary structure of the Tetrahymena group I intron.[44] Although in that case, the tetraloop is of the GNRA

(42) Spencer, R.; Fisher, J.; Walsh, C. T. *Biochemistry* **1976**, *15*, 1043–1053.

(43) Ludwig, J.; Levy, A. *Biochemistry* **1964**, *3*, 373–378.

(44) (a) Michel, F.; Westof, E. *J. Mol. Biol.* **1990**, *216*, 581–606. (b) Jaeger, L.; Michel, F.; Westof, E. *J. Mol. Biol.* **1994**, *236*, 1271–1276.

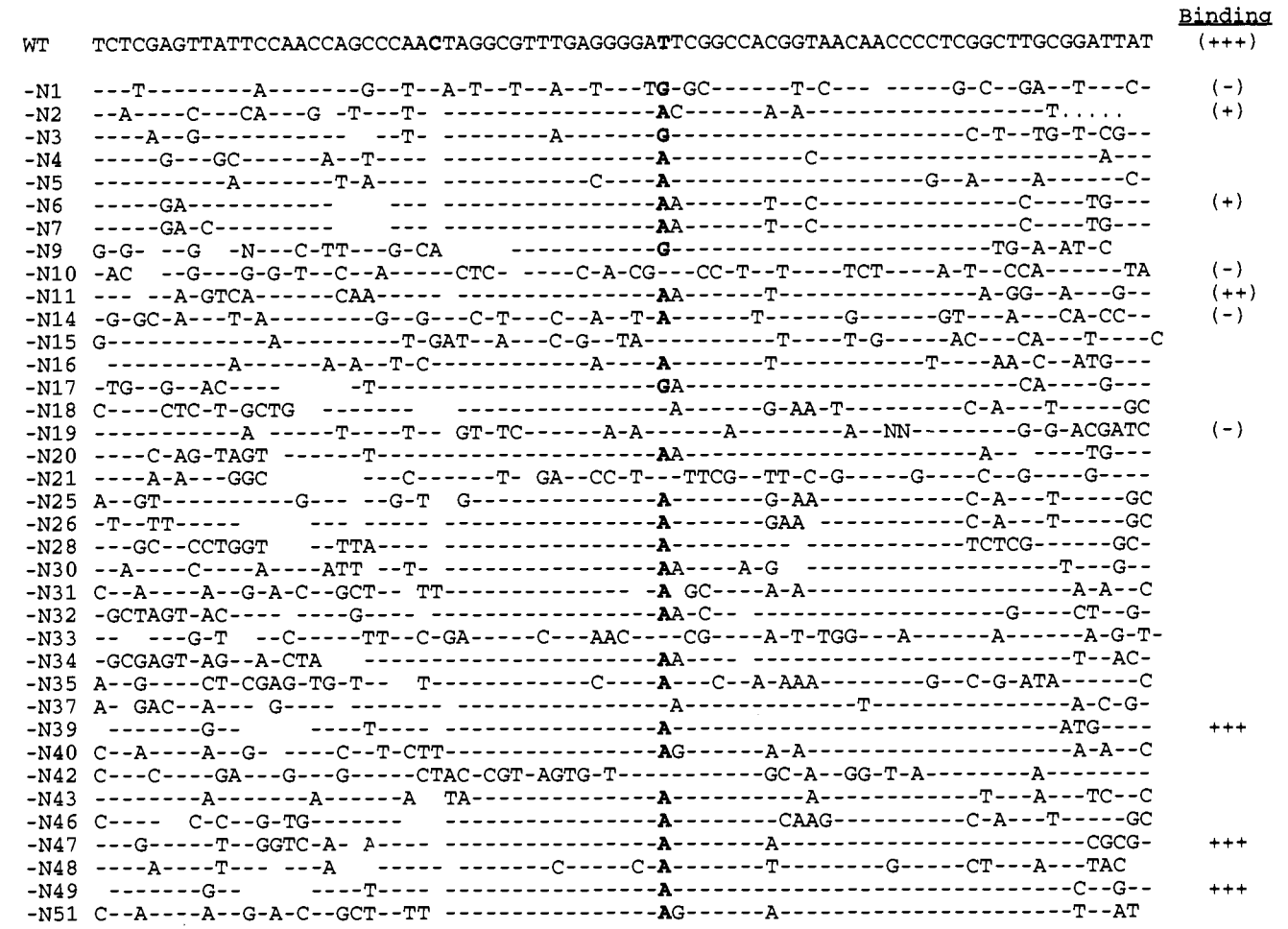



**Figure 13.** Alignment of sequences from pool selected for NMN selectivity over NMNH. A conserved mutation in the binding region is shown in bold.

type, the interaction involves hydrogen bonding of the 3′-terminal A in the tetraloop with a G-C base pair in the minor groove. Thus, it is interesting to note in the proposed structure the presence of a conserved A at the 3′-end of the smaller loop and a pair of highly conserved G-C base pairs in the larger stem.

There are only two other commonly observed mutations in the binding region, that of C35 mutated to a U and U38 mutated to a C. Binding analysis of reselected clones suggests that these mutations do not affect the binding selectivity. There were also a small number of clones from both mutagenized selections that contained mutations that destroyed the Watson−Crick pairing of bases in the larger stem. In every such case, binding to immobilized NAD was severely diminished. We are presently pursuing chemical modification experiments to gather more structural information on this aptamer.

## Conclusion

Using *in vitro* selection we have isolated RNA molecules that can specifically bind ribonucleotide-like redox cofactors with affinities in the low micromolar range. The riboflavin binding aptamers are proposed to form an intramolecular G-quartet structure, in which the majority of the loop regions connecting the guanosines in the quartet contain only a single nucleotide. This structure is novel because all quartet structures to date have loops containing at least two nucleotides. In addition, intramolecular RNA G-quartets have yet to be reported. Results from this work suggest that RNA G-quartets of this type may form more readily in solution than the corresponding DNA structures, and in some cases may be less sensitive to the type of monovalent cation for stability. These molecules also represent the first case in which both the DNA and RNA versions of an aptamer sequence bind the target ligand.

The solution to the riboflavin binding problem is in this case a relatively simple one, in that the consensus sequence given in Figure 4c contains a large number of unique sequences ($10^4$–$10^6$). Burgstaller and Famulok have also found relatively simple motifs for flavin binding. They probably did not find G-quartet solutions of the type reported here in their flavin selections because their selection buffer contained only $Na^+$ as the monovalent cation. RNA sequences containing the two-tiered G-quartet structures, which require $K^+$ for binding, would not have been enriched. The fact that there were no three-tiered RNA quartets enriched in any of the selections is not so easily explained, since a fraction of these sequences, although less abundant (ca. 1 in $10^{12}$), should be present in the starting pools. Sequences with a three-tiered structure were also not enriched with a riboflavin selection using the randomized pool of Figure 4c, which contains a subset of the three-tiered structures. It is possible that these sequences are very poor substrates for one or more of the enzymes involved in the selection scheme (e.g., reverse transcriptase), due to their apparent increased stability.

In the case of the nicotinamide selection, the RNA aptamers show differential binding between the two redox states of NAD, including one that displays over a 10-fold lower solution $K_d$ for NAD than for NADH. The binding region of this aptamer has been defined and evidence for a secondary structure presented based on sequence comparisons from a reselected mutagenized pool. The structure appears to be rather complex with two interacting stem-loop domains and is dependent upon $Mg^{2+}$ for stability. The complexity of the structure is reflected

in the small number of unique sequences that were enriched, i.e., 1−2 out of $4 \times 10^{15}$. Burgstaller and Famulok reported that an attempt to isolate NMN binding RNAs failed with a random pool that had a complexity of $\sim 1 \times 10^{15}$, using buffer conditions that would have allowed the isolation of the aptamers reported here. Thus, it appears that solutions for certain ligands (in a given affinity range) have a frequency that lies on the border of the number of molecules that can feasibly be used for this type of selection.

The differential binding observed for the NMN aptamer at the dinucleotide level has been extended to the mononucleotide level through *in vitro* evolution techniques. Comparison of the reselected sequences have allowed the identification of a mutation that contributes to the additional specificity, as confirmed by binding assays with small RNAs. Incorporation of this mutation into a small RNA representing the binding region of the original aptamer sequence did not further enhance the specificity for NAD over NADH as measured by column assay. Because of its complexity, this structure may serve as a useful starting point for further *in vitro* evolution into ribozymes with oxidoreductase activity.

## Experimental Section

**General Chemical Methods.** NMR spectra were recorded on a Varian Gemini 300 MHz spectrophotometer. $^1$H chemical shifts are reported relative to chloroform at $\delta$ 7.24 or water (residual HOD) at $\delta$ 4.63. $^{13}$C chemical shifts are reported relative to chloroform at $\delta$ 77.0 or to $CH_3OH$ in $D_2O$ at $\delta$ 49.9. Chromatography refers to the method of Still, Kahn, and Mitra[45] using silica gel 60 (E. Merck, Darmstadt). High resolution mass spectra were obtained from the Harvard University Mass Spectroscopy Facility.

**Materials.** $\beta$-NMN, $\beta$-NMNH, $\beta$-NAD agarose, $\beta$-NAD, $\beta$-NADH, adipic dihydrazide agarose, and [$^3$H]-riboflavin were from Sigma. Quinoxaline-3-carboxylic acid, sodium cyanoborohydride, and all anhydrous solvents were from Aldrich. [$^{14}$C]-NAD was from DuPont-NEN, and [$^3$H]-$NaBH_3CN$ and [$^{32}$P]-NAD were from Amersham. N3-(Carboxymethyl)riboflavin was prepared by the method of Wu et al.[46] 6-(*N*-D-ribityl-3,4-xylidino)uracil tetraacetate (**8**) was prepared by the procedure of Yoneda et al.,[47] as modified by Ashton et al.[27] 5-Deazariboflavin tetraacetate was prepared from **8** by the method of Janda and Hemmerich.[29]

***tert*-Butyl 1-Deoxy-1-(N3-(carboxymethyl)-3,4-dihydro-7,8-dimethyl-2,4-dioxopyrimido[4,5-*b*]quinolin-10(2*H*)yl)-D-ribitol Tetraacetate (*tert*-Butyl N3-(Carboxymethyl)-5-deazariboflavin Tetraacetate) (4).** To a solution of 520 mg (0.96 mmol) of 5-deazariboflavin tetraacetate (**3**) in 7 mL of dry DMF was added 0.71 mL (4.8 mmol) of *tert*-butyl bromoacetate via syringe followed by 660 mg (4.8 mmol) of powdered anhydrous potassium carbonate. The mixture was warmed to 50 °C and stirred under argon in the dark for 12 h. The solvent was removed, and the residue was dissolved in $CHCl_3$ and washed with 10% HOAc, dilute aqueous $NaHCO_3$, and brine. The organic phase was then dried with anhydrous $Na_2SO_4$ and filtered, and the solvent was removed to give crude **4** as an orange oil. The crude material was purified by flash chromatography over silica using 2% MeOH in $CH_2Cl_2$ as eluent to give 335 mg (51%) of **4** as a yellow oil: $^1$H NMR ($CDCl_3$) $\delta$ 8.81 (s, 1), 7.63 (s, 2), 5.64 (m, 1), 5.42 (m, 2), 4.70 (s, 2), 4.43 (dd, 1, $J$ = 2.0), 4.28 (dd, 1, $J$ = 5.5), 2.55 (s, 3), 2.42 (s, 3), 2.31 (s, 3), 2.22 (s, 3), 2.08 (s, 3), 1.76 (s, 3), 1.47 (s, 9); $^{13}$C NMR ($CDCl_3$) $\delta$ 170.33, 170.07, 169.94, 169.69, 169.49, 167.06, 161.48, 156.56, 155.56, 146.76, 142.46, 139.22, 134.30, 131.39, 119.83, 116.38, 113.57, 81.65, 70.58, 70.38, 69.91, 69.72, 69.63, 69.06, 61.88, 44.46, 42.84, 27.95, 21.21, 20.87, 20.62, 20.48, 20.12, 18.93; HRMS calcd for $C_{32}H_{39}N_3O_{12}(Na^+)$ 680.2421, found 680.2431.

**1-Deoxy-1-(N3-(carboxymethyl)-3,4-dihydro-7,8-dimethyl-2,4-dioxopyrimido[4,5-*b*]quinolin-10(2*H*)yl)-D-ribitol (N3-(Carboxymethyl)-5-deazariboflavin) (5).** To 300 mg (0.46 mmol) of **4** at 0 °C was added 5 mL of trifluoroacetic acid, and the mixture was stirred at 0 °C under argon for 12 h. The solvent was removed under vacuum, and the residue was dissolved in 10 mL of methanol saturated with potassium carbonate and stirred for 4 h at 0 °C. The solvent was removed, and the residue (100% crude) was dissolved in 5 mL of saturated aqueous $NaHCO_3$ and precipitated by addition of concentrated HCl to pH 2 on ice. The yellow solid was then filtered, washed with water and then ethanol, and air dried to give 141 mg (72%) of **5**: $^1$H NMR ($D_2O$) $\delta$ 8.57 (s, 1), 7.62 (s, 1), 7.51 (s, 1), 4.83 (m, 1), 4.56 (m, 1, partially obscured by HOD), 4.31 (s, 2), 4.16 (m, 1), 3.70 (m, 3), 3.55 (dd, 1, $J$ = 6.0, 11.8), 2.34 (s, 3), 2.20 (s, 3); $^{13}$C NMR ($D_2O$) $\delta$ 182.61, 175.83, 150.11, 144.08, 140.21, 137.19, 131.98, 121.17, 117.99, 74.06, 73.88, 71.82, 63.81, 48.02, 45.90, 24.06, 21.98, 19.50; HRMS calcd for $C_{20}H_{23}N_3O_8$ ($M + Na^+$) 456.1383, found 456.1377.

**Methyl 1-Deoxy-1-(5-(carboxyethyl)-3,4-dihydro-7,8-dimethyl-2,4-dioxopyrimido[4,5-*b*]quinolin-10(2*H*)yl)-D-ribitol Tetraacetate (Methyl 5-(Carboxyethyl)-5-deazariboflavin Tetraacetate) (8).** To a solution of 750 mg (1.41 mmol) of **7** in 20 mL of anhydrous 1,2-dichloroethane under argon was added 1.35 mL (14 mmol) of 3-(methoxycarbonyl)propionyl chloride followed by 1.6 mL (14 mmol) of anhydrous stannic chloride. The resulting mixture was refluxed for 3 h, and the solvent was removed. Longer reflux times gave side reactions which made purification more difficult. The crude material was purified by chromatography using a step gradient of 2−3% methanol in dichloromethane as eluent to give 210 mg (25%, 49% based on recovery of starting material) of **8** as a yellow oil: $^1$H NMR ($CDCl_3$) $\delta$ 8.81 (s, 1), 7.88 (s, 1), 7.59 (s, 1), 5.63 (m, 1), 5.46 (m, 1), 5.39 (m, 1), 4.42 (d, 1, $J$ = 11.3), 4.24 (dd, 1, $J$ = 5.7, 12.1), 3.97 (m, 2), 3.69 (s, 3), 2.70 (m, 2), 2.51 (s, 3), 2.41 (s, 3), 2.25 (s, 3), 2.19 (s, 3), 2.05 (s, 3); $^{13}$C NMR ($CDCl_3$) $\delta$ 172.50, 170.47, 170.17, 169.76, 169.60, 162.71, 159.18, 157.73, 155.54, 146.50, 138.52, 134.41, 127.11, 120.19, 117.06, 111.53, 70.78, 70.06, 69.30, 62.02, 51.82, 45.34, 33.70, 24.41, 20.97, 20.71, 20.57, 20.27, 19.53; HRMS calcd for $C_{30}H_{35}N_3O_{12}$ ($M + Na^+$) 652.2118, found 652.2109.

**1-Deoxy-1-(5-(carboxyethyl)-3,4-dihydro-7,8-dimethyl-2,4-dioxopyrimido[4,5-*b*]quinolin-10(2*H*)yl)-D-ribitol (5-(Carboxyethyl)-5-deazariboflavin) (9).** A solution of 138 mg (0.22 mmol) of **8** in 20 mL of 20% CH3CN in 0.1 N NaOH was allowed to stir in the dark overnight at room temperature. The solution was concentrated under reduced pressure to half the original volume and acidified with 2 N HCl to pH 2 and placed at 4 °C until a yellow precipitate formed. The solid was filtered and dried in vacuo to give 52 mg (50%) of a yellow solid which was a single fluorescent spot by TLC ($CH_3CN/H_2O$/HOAc, 20:4:1): $^1$H NMR ($D_2O$) $\delta$ 7.70 (s, 1), 7.47 (s, 1), 4.83 (m, 2), 4.40 (m, 1), 3.98 (m, 1), 3.67 (m, 3), 3.57 (m, 1), 3.50 (dd, 1, $J$ = 6.5, 11.9), 2.32 (s, 3), 2.35 (s, 3), 2.19 (t, 2, $J$ = 8.5); $^{13}$C NMR (DMSO) $\delta$ 173.25, 163.18, 157.93, 155.76, 145.64, 138.52, 133.83, 126.16, 119.60, 118.25, 111.07, 73.78, 72.92, 69.77, 63.47, 51.88, 47.80, 33.65, 23.88, 20.62, 18.96; HRMS calcd for $C_{21}H_{25}N_3O_8$ ($M + Na^+$) 470.1539, found 470.1533.

**Reduction of 5 and 5-Agarose with $NaBH_3CN$.** To a 5 mM solution of **4** in 0.1 M potassium phosphate, pH 6.0, was added solid $NaBH_3CN$ to give a 0.1 M solution. After shaking to dissolve the $NaBH_3CN$, the mixture was allowed to stand at room temperature in the dark for 4 h. The mixture was then desalted by passing over a SepPack cartridge, washing with water, and then eluting with 50% aqueous $CH_3CN$. The solvent was removed to give **6** as a pale yellow solid [$^1$H NMR ($D_2O$) $\delta$ 6.81 (s, 2), 4.35 (s, 2), 4.09 (m, 1), 3.89 (d, 1, $J$ = 16), 3.75 (m, 3), 3.62 (m, 1), 3.54 (dd, 1, $J$ = 7.2, 14.4), 3.40 (s, 2), 2.07 (s, 3), 2.03 (s, 3)]. Analysis of the UV−vis spectrum indicated that **5** was >95% reduced, and no signals for the oxidized deazaflavin were observed in the $^1$H NMR spectrum. In buffered solution at pH 8.5 there is a clear peak at 263 nm in the UV spectrum indicative of the 1,5-dihydro-5-deazaflavin anion at N1−O2.[42] Reduction of **5**-agarose was achieved by adding solid $NaBH_3CN$ directly to the resin suspended in an equal volume of 0.1 M potassium phosphate, pH 6.0, to give a 0.1 M concentration of $NaBH_3CN$. It was repeatedly observed that immobilized dCMRb was reduced more rapidly than dCMRb in solution, perhaps due to the removal of electrostatic repulsion by the N3-carboxymethyl group upon coupling to the resin. Reoxidation of immobilized **6** back to **5** was observed to be essentially quantitative

(45) Still, W. C.; Kahn, M.; Mitra, A. *J. Org. Chem.* **1978**, *43*, 2923−2925.

(46) Wu, F. Y.-H.; MacKenzie, R. E.; McCormick, D. B. *Biochemistry* **1970**, *9*, 2219−2224.

(47) Yoneda, F.; Sakuma, Y.; Ichiba, M.; Shinomura, K. *J. Am. Chem. Soc.* **1976**, *98*, 81.

by comparing the absorbance at 396 nm of **5**-agarose suspended in 50% glycerol before reduction and after reoxidation.

**General Molecular Biology.** Oligodeoxynucleotides were synthesized on a Millipore Expedite using phosphoramidite technology, deprotected with 20−30% $NH_4OH$ at 55 °C for 12 h, and purified by polyacrylamide gel electrophoresis (PAGE). The band of interest was visualized by UV shadowing, excised from the gel, and the nucleic acid eluted by the crush and soak method in 0.5 M NaCl. RNA was converted to cDNA using reverse transcriptase (Superscript, BRL, 200 units/$\mu$L) in a reaction mixture consisting of 0.2 mM dNTPs, 1 $\mu$M 3′-primer, 10 mM DTT, and 1 $\mu$L enzyme in buffer supplied by the manufacturer. DNA was amplified by polymerase chain reaction using a thermocycler and a reaction mixture consisting of 100 mM Tris pH 9.0, 50 mM KCl, 1.5 mM $MgCl_2$, 0.1% Triton X-100, and 2.5 units of Taq DNA polymerase (Promega) for each 100-$\mu$L reaction. The 5′-primer typically contained the T7-promoter sequence followed by 20 nt identical to the first 20 nt of the synthesized pool DNA, while the 3′-primer contained 18 nt complementary to the 3′ end of the pool DNA. Transcription of PCR DNA was performed with T7 RNA polymerase (USB, 80 units/$\mu$L) in a mixture consisting of 40 mM Tris pH 7.5, 35 mM $MgCl_2$, 5 mM of each of the NTPs, 1 mM spermidine, 0.01% (v/v) Triton X-100, and 4 $\mu$L enzyme/100 $\mu$L reaction volume. Transcription reactions were typically run overnight. Cloning of PCR products was performed with a TA cloning kit (Invitrogen) using the instructions provided by the manufacturer. Dideoxy sequencing was by the method of Sambrook[48] using $^{35}$S-ATP-α-S (Dupont NEN) and Sequenase (USB).

**Pool Construction.** The pool containing a random region of 72 bases has been previously described.[16] For the nicotinamide selection, a random pool of DNA containing an 80 nt random region with constant flanking regions of 18 nt was synthesized on a Millipore Expedite DNA synthesizer and purified as described above. The quality of the synthesized DNA was assessed using a primer extension assay, in which a 20-fold molar excess of pool DNA was incubated with $^{32}$P-labeled 3′-primer and subjected to one PCR cycle. The radioactivity of the full length product was compared to the total incorporated radioactivity. Using this method, approximately 30% of the gel-purified pool DNA was found to be fully extendible. The gel-purified material (235 $\mu$g of extendible template DNA) was amplified 16-fold by large scale PCR. The PCR reaction mixture was then phenol and chloroform extracted, and the DNA was ethanol precipitated and resuspended in water for a large-scale transcription reaction (18 mL). The transcription reaction was terminated after 6 h by addition of 1.8 mL of 0.5 M EDTA, pH 8.0. To a portion (3 mL) of the transcription reaction mixture was added 4 g of solid urea and 1% bromophenol blue to give a final concentration of 0.01%, and the RNA was purified by PAGE (8%) to give ca. 3 mg of random pool RNA (11 copies based on a calculated DNA pool complexity of $4 \times 10^{15}$ molecules).

**Selection Procedures. Riboflavin selection.** The selection buffer contained 20 mM HEPES pH 7.5, 300 mM KCl; and 5 mM $MgCl_2$. Approximately three copies of the initial RNA pool (complexity of $5 \times 10^{14}$ molecules, 96 $\mu$g) was dissolved in selection buffer and allowed to equilibrate for 15 min. The RNA was applied to a 0.3 mL precolumn of adipic dihydrazide agarose. The precolumn was washed with 0.6 mL of buffer onto a 1-mL column of riboflavin-derivatized agarose. The riboflavin column was then washed with buffer for a given number (usually 10−20) of column volumes, and bound RNA was then eluted with a concentration of riboflavin in buffer twice that of the estimated concentration on the column (e.g., 1 mM for rounds 1−5). The affinity eluted RNA was then ethanol precipitated overnight using 5 $\mu$g of glycogen as carrier. The precipitated RNA was collected by centrifugation and resuspended in water for reverse transcription and PCR amplification. Beginning with round 2, the transcription products were treated with RNase-free DNase (1 unit/$\mu$g) to remove PCR DNA. In rounds 6−9, negative selection was effected by elution with 5 mM quinoxaline-3-carboxylic acid in selection buffer for three column volumes prior to affinity elution with riboflavin.

**NMN Selection.** The selection buffer contained 20 mM HEPES, pH 7.5, 200 mM KCl, 100 mM NaCl, and 5 mM $MgCl_2$. For round 1, 800 $\mu$g of pool RNA (3 copies, complexity $4 \times 10^{15}$ molecules) in 0.5 mL of buffer was added to a precolumn of 0.3 mL of ATP-agarose (C-8 diamino linked, cyanogen bromide activated). The eluent of the precolumn was allowed to pass directly onto a column of 1.0 mL of NAD-agarose (adenine C-8 linked analogous to ATP-agarose), and the precolumn was washed with 2 × 0.3 mL of buffer before it was removed. The NAD column was then washed for a given number of column volumes (typically 10) before bound RNA was eluted with three column volumes of a solution of 5 mM NMN in buffer. Affinity eluted RNA was then precipitated and subjected to reverse transcription, PCR, and transcription with T7 RNA polymerase as described above. Negative selection was effected as described above with quinoxaline-3-carboxylic acid (10 mM).

**Column Binding Assays.** Assays for RNA (or DNA) ligand binding and binding specificity studies were performed by a column assay in which 1−5 $\mu$g of $^{32}$P-labeled RNA or DNA in 0.15 mL of selection buffer was loaded onto a 0.3 mL affinity column equilibrated in buffer. The column was washed with buffer for six column volumes and then eluted with the ligand in buffer at a concentration equal to or greater than the ligand concentration on the column for four column volumes. Fractions of one column volume are collected, and the radioactivity quantitated for each. For binding specificity experiments, the amount of radiolabeled RNA eluted by the particular ligand at 5 mM for four column volumes is compared to determine approximate relative affinity.

**Fluorescence Binding Measurements.** Binding constants in the low micromolar range for RNA (DNA) with flavins or 5-deazaflavins could be measured by observing the quenching of flavin or deazaflavin fluorescence by bound RNA or DNA. Small aliquots (1−2 $\mu$L) of a concentrated solution of the RNA (100−500 $\mu$M) in binding buffer were added to a dilute solution (0.1 $\mu$M) of flavin (or deazaflavin) in buffer, and the fluorescence was measured after each addition. Fluorescence data, after correction for dilution, was fit to eq 1 using a nonlinear regression program to give the dissociation constant,

$$\Delta F = F_0/(K_d + [L]) \quad (1)$$

where $\Delta F$ is the change in flavin fluorescence upon added RNA, $F_0$ is the original fluorescence of the solution, and [L] is the RNA concentration at each addition. In each case, over 80% of the binding curve was defined by the experimental data. Complete quenching of the fluorescence was observed for both flavins and 5-deazaflavins. For flavins, the excitation and emission wavelengths used were 450 and 525 nm, respectively, while for the 5-deazaflavins they were 375 and 450 nm, respectively. Scatchard analysis of the number of binding sites was done using eq 2 and a plot of $\nu$ vs $\nu$/[L][49]

$$\nu = n - K_d(\nu/[L]) \quad (2)$$

where $\nu = (F_0 - F)/F_0$.

**Equilibrium Gel Filtration** Dissociation constants were measured for NMN aptamers with $^{14}$C- or $^{32}$P-labeled NAD based on a previously described procedure.[12] A 15 × 0.5 cm column of Sephadex G-25 Superfine was used, and fractions of 110 $\mu$L were collected. Typically, 1−5 nmol of RNA was added to the column, dissolved in column buffer consisting of 20 mM HEPES, pH 7.5, 200 mM KCl, 100 mM NaCl, 5 mM $MgCl_2$, and 1.0 $\mu$M NAD. The fractions were quantitated for radiolabeled NAD, and the area of the peak calculated to determine the amount of NAD bound to the RNA.

**Acknowledgment.** The authors would like to acknowledge Dr. David P. Bartel for the use of the random RNA for the flavin selection. We would also like to thank Jon Lorsch for discussions and critical reading of the manuscript. This work was supported by a fellowship from the Damon Runyon-Walter Winchell Cancer Research Fund to C.T.L. (DRG-1157) and by a grant from Hoechst AG.

JA942925I

(48) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1989.

(49) Scatchard, G. *Ann. N. Y. Acad. Sci.* **1949**, *51*, 660−672.