RNA-Catalyzed CoA, NAD, and FAD Synthesis from Phosphopantetheine, NMN, and FMN†

Faqing Huang,*,‡ Charles Walter Bugg,‡ and Michael Yarus§

Department of Chemistry and Biochemistry, University of Southern Mississippi, Hattiesburg, Mississippi 39406-5043, and Department of Molecular, Cellular, and Developmental Biology, University of Colorado at Boulder, Boulder, Colorado 80309-0347

Received August 31, 2000; Revised Manuscript Received October 17, 2000

ABSTRACT: A novel in vitro selection method was developed to isolate RNA sequences with coenzyme-synthesizing activities. We used size-heterogeneous libraries containing randomized ribonucleotide sequences of four different lengths (30*N*, 60*N*, 100*N*, and 140*N*), all with 5′-ATP initiation. Two RNAs, CoES7 (30*N*) and CoES21 (60*N*), are able to catalyze the synthesis of three common coenzymes, CoA, NAD, and FAD, from their precursors, 4′-phosphopantetheine, NMN, and FMN, respectively. Both ribozymes require divalent manganese for activities. The results support the availability of these coenzymes in an RNA world, and point to a chemical explanation for the complex bipartite structures of many coenzymes.

Coenzymes are densely functionalized small molecules. They play essential roles in metabolism by performing chemistry that is inefficient or impossible for typical amino acids. Three common coenzymes, coenzyme A (CoA), nicotinamide adenine dinucleotide (NAD), and flavin adenine dinucleotide (FAD), carry out a variety of acyl group and electron/hydride transfer reactions. Structurally, CoA, NAD, and FAD are complex (Figure 1), consisting of a ribonucleoside adenosine, a pyrophosphate linkage, and a functional group (pantetheine, nicotinamide, or riboflavin). However, the chemically functional elements within these coenzymes are the simpler moieties: sulfhydryl, nicotinamide, and isoalloxazine, respectively.

It has not been clear why these coenzymes have their complex conserved structures. While adenosine and pyrophosphate serve to anchor these coenzymes to host proteins, it seems unlikely that coenzyme availability and reactivity are uniquely optimized by these groups. Therefore, the universality of adenosine pyrophosphate may be the result of evolutionary descent instead of functional requirement. In fact, the existence of these coenzymes in all kingdoms suggests their persistence since the last common ancestor of life on earth (*1*). Furthermore, the inclusion of ribonucleotide adenosine in these coenzymes may imply the utilization of CoA, NAD, and FAD in a more ancient RNA world (*2*) as parts of a prior generation of RNA enzymes, as suggested by White (*3*, *4*).

To gain insight into the complex nature of coenzyme structures relative to their functions, we have isolated relatively small RNA molecules that synthesize RNA-linked CoA, NAD, and FAD from their corresponding precursors, 4′-phosphopantetheine (pan-p), nicotinamide mononucleotide (NMN), and flavin mononucleotide (FMN). The finding suggests a plausible mechanism of coenzyme synthesis and utilization in the RNA world, and may provide clues about coenzyme origin and evolution.

MATERIALS AND METHODS

Substrate Preparation. Two terminal phosphate-containing molecules (Figure 2) were prepared for use as the substrates in the isolation of active ribozymes. The first compound, biotin-p, was prepared from the reaction of sulfo-NHS−SS−biotin (Pierce) with phosphocolamine (Fluka) as follows. Phosphocolamine (110 mg) and sulfo-NHS−SS−biotin (90 mg) were dissolved in 8 mL of water, and the pH of the solution was adjusted to 8.3 with 1.0 N NaOH. The sample was allowed to react with constant stirring for 4 h at room temperature. Two volumes of EtOH and 2 volumes of acetone were then added to precipitate biotin-p at −20 °C. A collected solid sample was dissolved in 8 mL of water, and the pH was adjusted to 3.7 with acetic acid. The solution was loaded onto a C18 Sep-Pak column (Waters), washed with water, and eluted with methanol. The sample was dried under vacuum and dissolved in 8 mL of solution. The solution pH was adjusted to 6.0 with 1.0 M NaOH. Amino group analysis by ninhydrin gave a negative result, indicating that phosphocolamine was absent. The sample was analyzed by mass spectrometry. The expected molecular weight (MW) is 530. Peaks found by ESI-MS (Mass Consortium, San Diego, CA) were as follows: positive ion at 575 ($M - H^+ + 2Na^+$) and negative ion at 529 ($M - H^+$). No free biotin peak was observed. The concentration of biotin-p was determined by alkaline phosphatase digestion followed by quantitative inorganic phosphate analysis (*5*).

† This work was supported by NSF Grant MCB9974487 to F.H. and NIH Grants GM30881 and GM48080 to M.Y.

* To whom correspondence should be addressed. Phone: (601) 266-4371. Fax: (601) 266-6075. E-mail: faqing.h.huang@usm.edu.

‡ University of Southern Mississippi.

§ University of Colorado at Boulder.

10.1021/bi002061f CCC: $19.00 © 2000 American Chemical Society
Published on Web 11/23/2000

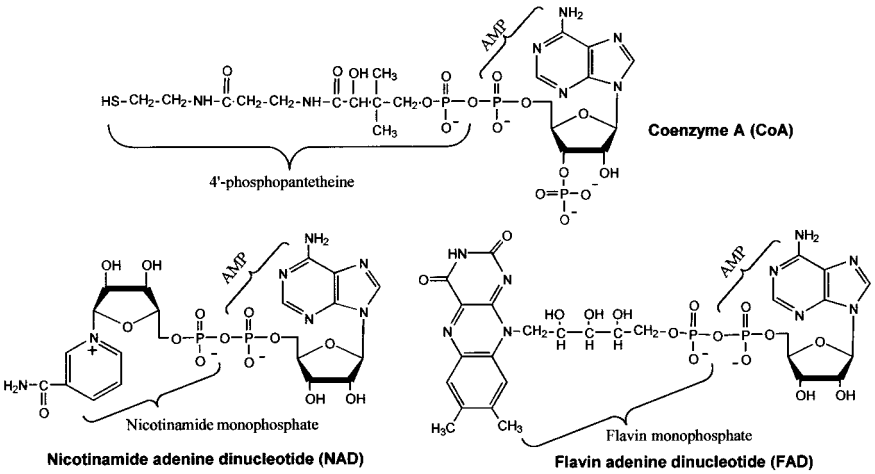


FIGURE 1: Structures of CoA, NAD, and FAD. The coenzymes possess structural similarities, although they have very distinct biochemical functions.

The second substrate (Figure 2B), de-p-CoA−SS−pan-p, was prepared from CoA (Sigma) through multiple steps. First, de-p-CoA was prepared from CoA by phosphatase digestion: 9 mL of 10 mM CoA, 7 units of acid phosphatase (Sigma), and incubation for 5 h at 37 °C. HPLC analysis (by a C18 column) indicated >98% dephosphorylation. After digestion, phosphatase was removed by filtration through a Microcon 3 filter (Amicon). Second, pan-p was obtained from de-p-CoA by digestion with snake venom phosphodiesterase (SVPDE, Sigma), which cuts the pyrophosphate linkage to produce AMP and pan-p. Conditions included 400 μL of 10 mM de-p-CoA, 0.05 unit of SVPDE, 15 mM $MgCl_2$, 0.1 M Tris, pH 8.9, and 1 h at 37 °C. Complete hydrolysis was confirmed by HPLC. SVPDE was removed by Microcon filtration. Third, oxidized de-p-CoA (dimer, de-p-CoA−SS−CoA-de-p) was prepared by H_2O_2 oxidation of de-p-CoA (10 mM de-p-CoA and 200 mM H_2O_2, 1 h at 37 °C) and purified by HPLC. Finally, de-p-CoA−SS−pan-p was made from de-p-CoA−SS−CoA-de-p and pan-p through a disulfide bond exchange reaction (20 mM de-p-CoA−SS−CoA-de-p and 20 mM pan-p, 25 min at 65 °C) and purified by reverse phase HPLC (C18) in a MeOH (20%)/10 mM ammonium formate (80%) solvent at pH 3.7.

Heterogeneous Random DNA/RNA Library Construction. Four different random DNA libraries of 30*N*, 60*N*, 100*N*, and 140*N* (*N* = the number of random nucleotides) were prepared separately by PCR amplification. For each size, 100 pmol of random DNA templates was used. DNA oligonucleotides (synthesized by Operon Tech) used for DNA library construction were P1 (GCGAATTCAGTAATACGACTCACTATT**A**GGGAAGTGCTACC), P2 (TGGCTGGCCGCATGCCCG), N28 [TGGCTGGCCGCATGCCCG$(N)_{28}$GGTCGCACGTGCCG], N30 [TGGCTGGCCGCATGCCCG$(N)_{30}$GGTAGCACTTCCCT], N60 [TGGCTGGCCGCATGCCCG$(N)_{60}$GGTAGCACTTCCCT], N68 [TGGCTGGCCGCATGCCCG$(N)_{68}$GGTCGCACGTGCCG], and N72 [AGGGAAGTGCTACC$(N)_{72}$CGGCACGTGCGACC]. P1 and P2 are PCR primers. The underlined sequence is the T7 class II promoter ϕ2.5 (*6*), with the bold **A** as the initiating nucleotide. The four heterogeneous DNA libraries of different lengths were assembled using the following DNA oligonucleotides: 30*N* library, P1, P2, and N30; 60*N* library, P1, P2, and N60; 100*N* library, P1, P2, N28, and N72; and 140*N* library, P1, P2, N68, and N72. The starting DNA pool was prepared from the mixture of the four random DNA libraries.

The initial size-heterogeneous RNA library was transcribed by standard methods from the random DNA sequences described above (*7*). RNA transcripts were purified by gel filtration through G-25 columns (Pharmacia).

Selection Strategy. Biosyntheses of de-p-CoA, NAD, and FAD occur via coenzyme synthetases that catalyze the linkage of pan-p, NMN, or FMN with ATP to form their respective coenzymes, with concurrent release of pyrophosphate:

$$\text{R-p} + \mathbf{pppA} \rightarrow \text{R-p}\mathbf{pA} + \mathbf{PP_i} \quad (1)$$

where R is either pantetheine, nicotinamide nucleoside, or riboflavin and R-p**pA** represents a product coenzyme, de-p-CoA, NAD, or FAD. We therefore designed in vitro selections for RNA catalysts of the general reaction

$$\text{R-p} + \mathbf{pppA{-}RNA} \rightarrow \text{R-p}\mathbf{pA{-}RNA} + \mathbf{PP_i} \quad (2)$$

To isolate RNA capable of CoA, NAD, and FAD syntheses, we began with 5′-ATP-terminated randomized sequences (i.e., ATP−RNA or pppA−RNA). This was achieved by ATP-initiated RNA transcription of randomized DNA under the T7 RNA polymerase class II promoter (*6*).

To isolate RNA sequences with desired coenzyme-synthesizing activities, we designed a two-stage method as shown in Figure 2. During the first stage (from rounds 1−13), a mixture of four distinct lengths of RNA was used. RNA sequences that reacted with biotin-p were enriched. The reaction buffer contained $MgCl_2$, $CaCl_2$, $MnCl_2$, and imidazole (10 mM each) and NaCl, KCl, and NH_4Cl (20 mM each) at pH 6.0. The concentration of the reactant biotin-p was maintained at 1 mM. The reaction was carried out at room temperature for 18 h for the first three rounds and for 2 h for the following rounds until round 13. After incubation with biotin-p, the RNA solution was filtered through a G-25 column to deplete unreacted biotin-p. The collected RNA solution was then passed through a neutravidin affinity column (Pierce). After the column had been washed with water, RNA was eluted with 0.1 M DTT (30 min at 65 °C), precipitated with EtOH, and recovered in aqueous solution,

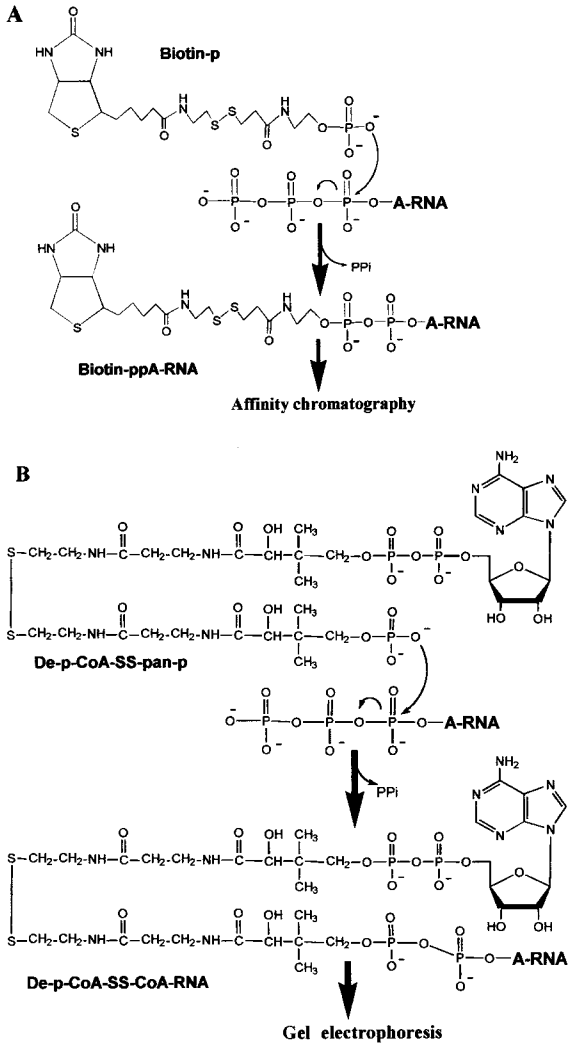


FIGURE 2: Selection methods for isolating coenzyme synthetase ribozymes. (A) Biotin-linked phosphate monoester (biotin-p) reacts with the 5′-terminal ATP of active RNA molecules so that active RNA sequences are tagged with biotin for subsequent purification by biotin−neutravidin affinity chromatography. The reaction mimics the coenzyme synthesis reaction (reaction 2). This method was used in the first stage of coenzyme synthetase ribozyme selection. (B) During the second stage of selection, a CoA precursor, pan-p, was directly used to react with active RNA sequences. An additional de-p-CoA was linked to 4′-phosphopantetheine to facilitate active RNA sequence separation during PAGE (see lane 2 in Figure 5). This second selection ensures that isolated RNA sequences have biotin-independent activities.

which was used for reverse transcription and PCR amplification. In the second stage of selection (rounds 14 and 15), two separate RNA solutions (30*N* and 60*N*; the two larger size groups, 100*N* and 140*N*, had disappeared during the first stage of selection; see Figure 4A) were incubated with 0.5 mM de-p-CoA−SS−pan-p for 2 h in the selection buffer. Active RNA sequences were separated from inactive ones by 8% denaturing polyacrylamide gel electrophoresis (PAGE), eluted, and recovered as described above.

After a total of 15 rounds of selection and amplification, the RNA pools were reverse-transcribed, PCR-amplified, cloned, and sequenced by the standard procedure (*7*).

Identification of RNA-Synthesized Coenzymes. Internally ^{32}P-labeled RNA transcripts were prepared by T7 transcription in the presence of [α-^{32}P]ATP and purified to single-nucleotide resolution by PAGE (*8*). To identify the reaction

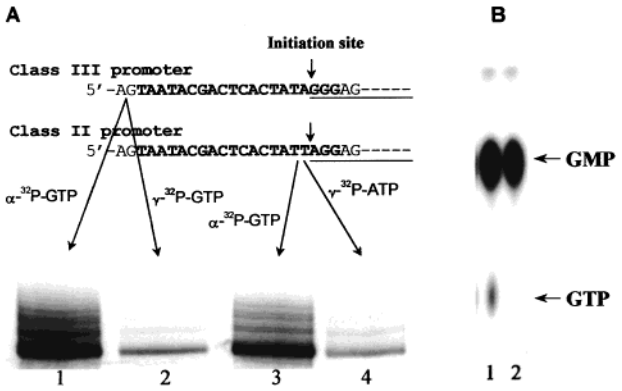


FIGURE 3: RNA transcription under the T7 class III (GTP initiation) and class II (ϕ2.5, ATP initiation) promoters. (A) PAGE analysis of transcription by T7 RNA polymerase. RNA from the class III promoter can be labeled by both [α-^{32}P]GTP (lane 1) and [γ-^{32}P]-GTP (lane 2), while both [α-^{32}P]GTP (lane 3) and [γ-^{32}P]ATP (lane 4) can label RNA transcripts from the class II promoter. Boldface letters represent the T7 promoters, and RNA transcripts are underlined. (B) TLC analysis (*11*) of nuclease P1 digests of [α-^{32}P]-GTP-labeled RNA isolated from panel A (lanes 1 and 3). Thus, transcription under the T7 class II promoter is initiated exclusively with ATP.

products, active RNA sequences were reacted with pan-p, NMN, and FMN. The RNA-linked products were separated by PAGE and recovered by extraction and ethanol precipitation. The coenzyme-linked RNA was digested with nuclease P1 (Pharmacia, at 0.2 unit/μL P1 in water for 3 h at 37 °C) and shrimp alkaline phosphatase (U.S. Biochemicals, at 0.5 unit/μL shrimp alkaline phosphatase in 1× supplier's buffer for 3 h at 37 °C). Digests were resolved by high-performance reverse phase TLC (HPTLC, Merck). Authentic unlabeled commercial coenzymes (Sigma) (served as markers) were mixed with the digests, and their locations on TLC after development were detected by UV fluorescence.

Kinetics. Kinetics for the two most active RNA sequences (one from each size group), CoES7 (30*N*) and CoES21 (60*N*), were conducted by incubating internally ^{32}P-labeled RNA (~0.1 μM) with various concentrations of the substrate biotin-p at room temperature in 20 mM MES (pH 6.0). Optimal divalent metal ion concentrations were used: 1 mM Mn^{2+} and 5 mM Ca^{2+} for CoES7 and 5 mM Mn^{2+} for CoES21. At various time points, aliquots were withdrawn and the reactions quenched with 20 mM EDTA. Products were resolved from unreacted RNA by 8% PAGE and visualized by autoradiography. Product yields were obtained by scanning the film and quantitating using a gel documentation and analysis system (Alpha Innotech Corp.).

RESULTS

Transcription Initiation with ATP. In our selection scheme (Figure 2), 5′-ATP-terminated RNA libraries are required for the RNA to make coenzymes that are covalently linked to RNA (reaction 2). Transcription by T7 RNA polymerase under the T7 class II promoter (*6*) shows that ATP is used to initiate transcription (Figure 3). RNA from both class III and class II promoters is labeled with [α-^{32}P]GTP (lanes 1 and 3 in Figure 3A). After nuclease P1 digestion, RNA from lane 1 (made from the class III promoter) produces both GMP and GTP in a ratio that is expected from the number of G's within the RNA (lane 1 in Figure 3B). On the other

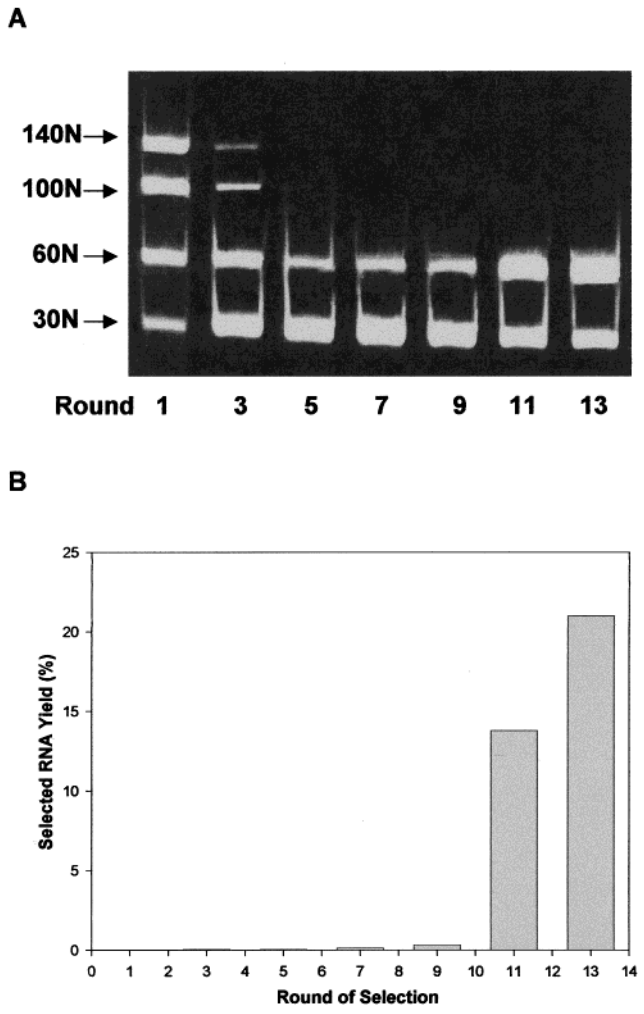


FIGURE 4: Progress of the first stage selection by using biotin-p and affinity chromatography. (A) Library size population (PCR DNA) changes with selection. (B) The level of biotinylated RNA increases with selection. Further rounds of selection after round 13 did not increase the selected yield.

hand, RNA made from the class II promoter (lane 3 in Figure 3A) yields only GMP, indicating that class II transcription does not start with GTP. Rather, ATP is used to initiate RNA transcription, as confirmed through RNA labeling by [γ-^{32}P]-ATP during transcription (lane 4 in Figure 3A). The yield of transcription (A initiation) under the class II promoter is about 70% of that from the more usual class III promoter (G initiation) (*9*).

RNA Selection from Size-Heterogeneous Libraries. To circumvent the effect of a possible RNA size−activity relation, we used size-heterogeneous RNA libraries. An initial DNA library containing four distinct tracts, 30*N*, 60*N*, 100*N*, and 140*N* (each size group with 100 pmol of random sequences), was constructed by PCR amplification (lane 1 in Figure 4A). Transcription of this DNA provided the initial pppA−RNA library. All RNAs contained the same terminal constant sequences so that they could be amplified together.

Two different selection methods were used to isolate general coenzyme synthetases (Figure 2). In the first method, a biotinylated phosphomonoester (biotin-p, Figure 2A) was used as both a substrate for reaction and an affinity tag for neutravidin chromatography purification. During the selection, size distribution within the library changed systematically as revealed by PCR DNA size distribution analysis (Figure 4A). During the first five rounds of selection, the shortest RNA (30*N*) became predominant, while the two longest groups (100*N* and 140*N*) virtually disappeared. However, the level of 60*N* RNAs progressively increased late in selection, from rounds 7 to 11.

The biotinylated RNA yield (Figure 4B) was used to follow the selection. RNA activity rapidly increased at round 11 and moderately increased thereafter. Comparing the two halves of Figure 4 reveals that the shortest sequences (30*N*) have the highest amplification efficiency (defined as the relative yield after cDNA synthesis, PCR amplification, and T7 RNA transcription). Thus, the level of 30*N* RNA increases and becomes dominant before the appearance of catalytic activity. However, from the point where catalytic molecules become a significant part of the population, the level of longer but catalytically more efficient RNAs (60*N*) increases and ultimately dominates the population (Figure 4A, from cycles 9 to 11). In an attempt to gain efficient RNA catalysts, after round 13, the selection was continued with lower biotin-p concentrations (down to 10 μM) and shorter reaction times (down to 8 min). In addition, mutagenic PCR (*10*) was used in two rounds of selection. However, there was no improvement in catalytic efficiency after round 13 as judged by biotinylation of the pool by biotin-p.

In the selection described above, the biotin affinity tag was available as part of the substrate (and was frequently utilized, results not shown). To isolate biotin-independent RNA catalysts, we then continued the selection from round 13 by a second method. The CoA precursor, pan-p, was used to react directly as the substrate to make CoA-linked RNA. Active RNA sequences were isolated by denaturing gel electrophoresis. To increase PAGE mobility differences, a de-p-CoA molecule was linked to pan-p to make de-p-CoA−SS−pan-p (Figure 2B). To facilitate RNA isolation by PAGE, 30*N* and 60*N* size groups of round 13 RNA were separated by gel and used in two separate parallel selections. The RNA product, de-p-CoA−SS−pan-ppA−RNA (or de-p-CoA−SS−CoA−RNA), was separated with about 2 cm resolution by 8% denaturing gel (see Figure 5, lane 2). Since 100*N* and 140*N* RNA groups had disappeared early in the first selection (Figure 4A), no attempt was made to include these in the second selection. During selection rounds 14 and 15, RNA product yields were about 3% from the 30*N* size group and 15−20% from the 60*N* size group. No significant increase was observed from round 14 to 15. Therefore, we cloned and sequenced the RNAs from round 15 pools.

Two dominant sequences (CoES7 and CoES21), one from each of the two RNA size groups, were analyzed for their activities. CoES7 is from the 30*N* size group, isolated 10 times (48% of the final pool). It has the sequence of agggaagtgctaccACAACUUUAGCCAUAAUGUCACUU-CUGCCGcgggcaugcggccagcca. CoES21 was isolated from the 60*N* size group. It repeated five times (31%) among the isolated clones. The sequence is agggaagtgctaccCCUGC-CCGUUGACCGGGGUCUAUGGUAUCUUACUCAU-AGUCGCCCAUUACCACAUAACAcgggcaugcggcca-gcca. Lowercase letters represent primer-related sequences, while uppercase letters are derived from randomized nucleotides.

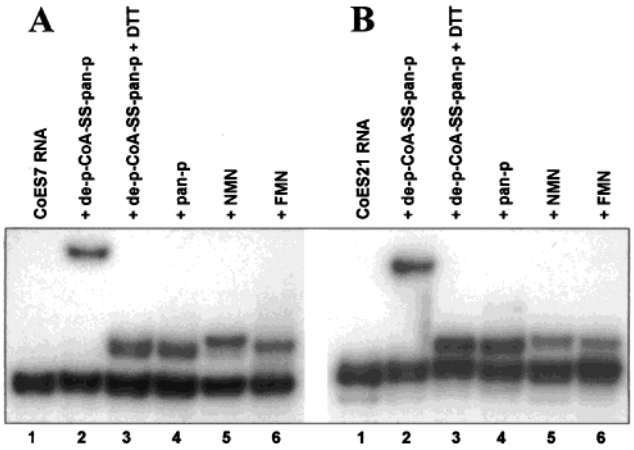


FIGURE 5: Reactions of isolated RNA sequences. RNA was prepared by transcription in the presence of [α-^{32}P]ATP and PAGE-purified to single-nucleotide resolution. Reaction conditions included 0.5 mM reactant and incubation for 1 h in the selection buffer at room temperature. (A) Isolate CoES7 RNA from the 30*N* size group. (B) CoES21 RNA from the 60*N* size group. Products were separated from unreacted RNA by 8% denaturing PAGE. The reaction depicted in lane 2 yielded CoA dimer-linked RNA (de-p-CoA−SS−CoA−RNA, Figure 2B), which migrated much slower than RNA (lane 1) or other RNA products (lanes 3−6). Reduction of de-p-CoA−SS−CoA−RNA by DTT [in 20 mM Tris (pH 8.0) and 50 mM DTT for 5 min at 65 °C] yielded expected CoA−RNA (lane 3), which was the same product from the reaction of CoA precursor pan-p with the RNA (lane 4).

Reaction Center and RNA-Catalyzed Reactions. The reaction site was defined by using 5′-[γ-^{32}P]RNA (*pppA−RNA where the asterisk indicates the ^{32}P), prepared by transcription in the presence of [γ-^{32}P]ATP under standard transcription conditions (*7*). During reactions with phosphate monoesters, pyrophosphate (PP_i) was the only ^{32}P-labeled product released when analyzed by PEI TLC (*11*) (results not shown). Therefore, the α-phosphate of the 5′-terminal ATP is the reaction center.

From the nature of the selection, we expect that both CoES7 and CoES21 react with a variety of terminal phosphate-containing nucleophiles, including the two (biotin-p and de-p-CoA−SS−pan-p) used as the substrates during the selection. Indeed, both RNAs are reactive toward the selection substrate de-p-CoA−SS−pan-p and coenzyme precursors, pan-p, NMN, and FMN as shown in Figure 5. The highly retarded RNA product (lane 2), de-p-CoA−SS−CoA−RNA, forms the basis for the second selection method. As expected, reduction by DTT produces CoA−RNA (lane 3), which is the same as the product from pan-p reaction with RNA (lane 4). Reaction with NMN and FMN results in NAD−RNA and FAD−RNA products, respectively (lanes 5 and 6).

Coenzyme Synthesis by RNA. To confirm CoA-, NAD-, and FAD-linked RNA products, reaction products were isolated from the gel (Figure 5, lanes 4−6) and digested by nuclease P1 and alkaline phosphatase. RNA-synthesized coenzymes were then identified by HPTLC analysis (Figure 6). Unlabeled authentic coenzymes were included in all analyses. Their locations, revealed by UV fluorescence, were marked on the left side of each panel. Panels A and B of Figure 6 establish that de-p-CoA is made by reacting pan-p with CoES7 (similar results with CoES21, not shown), while panels C and D of Figure 6 substantiate the synthesis of NAD and FAD by both CoES7 and CoES21. The following therefore summarize the results depicted in Figure 6:

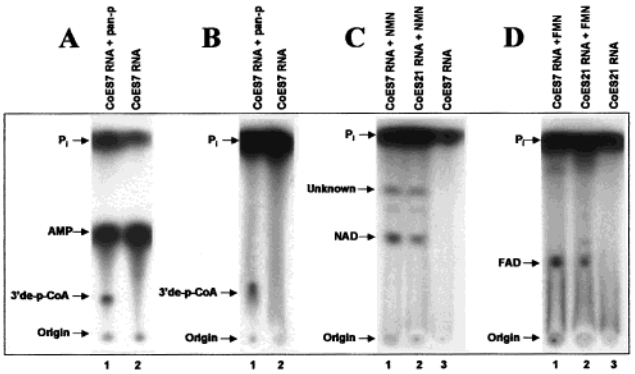
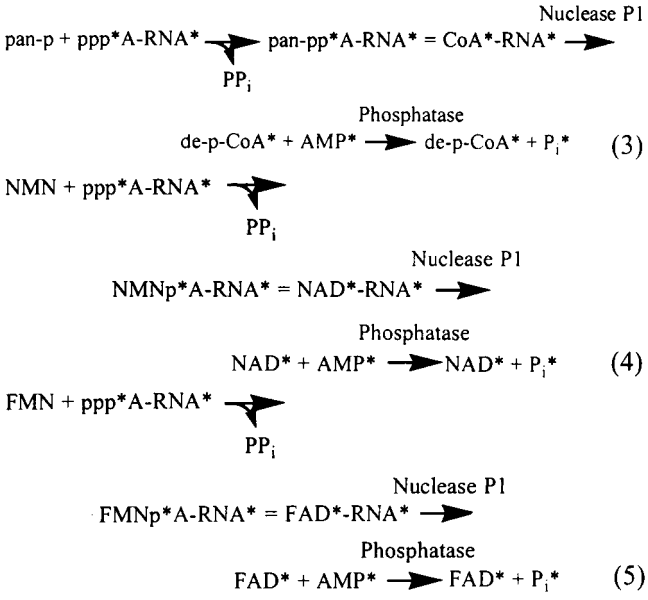


FIGURE 6: Identification of coenzymes synthesized by [α-^{32}P]ATP-labeled CoES7 and CoES21 RNAs. Reaction products from Figure 5 (lanes 4−6) were isolated by 8% denaturing PAGE. The RNA products were digested by nuclease P1 and shrimp alkaline phosphatase. Digests were resolved by high-performance TLC (HPTLC), along with authentic unlabeled coenzymes. In all cases, newly synthesized coenzymes (^{32}P-labeled at the 5′-α-phosphate of AMP) migrated to locations indistinguishable from those of unlabeled authentic coenzymes. (A) CoA identification. The product from the CoES7 RNA and pan-p reaction (lane 1) and RNA (lane 2) was digested by nuclease P1 and analyzed by C18 HPTLC, developed in 50 mM NaH_2PO_4 and 15% MeOH. P_i* resulted from a weak phosphatase activity accompanying the supplier's nuclease P1. Lane 2 should contain a small amount of ATP* (from the 5′-ATP* of RNA* after P1 digestion), but both P* and ATP* ran at the solvent front. Similar results (not shown) from CoES21 have been obtained to confirm that CoES21 also makes CoA from its precursor as shown in lane 4 in Figure 5B. (B) Samples from panel A were further digested by phosphatase and analyzed under the same conditions. (C) NAD synthesis from NMN. The samples were digested and analyzed as described for panel B. The identity of the band between NAD and P_i is unknown, but is believed to be the RNA reaction product from a terminal phosphate-containing contaminant in the reactant NMN. (D) FAD identification. The samples were digested and analyzed as described for panel C except for 30% MeOH in place of 15% MeOH during TLC.



Manganese-Requiring Ribozymes. To establish the requirements for CoES7 and CoES21 RNA, a combination of different buffer compositions was used to assay the catalytic

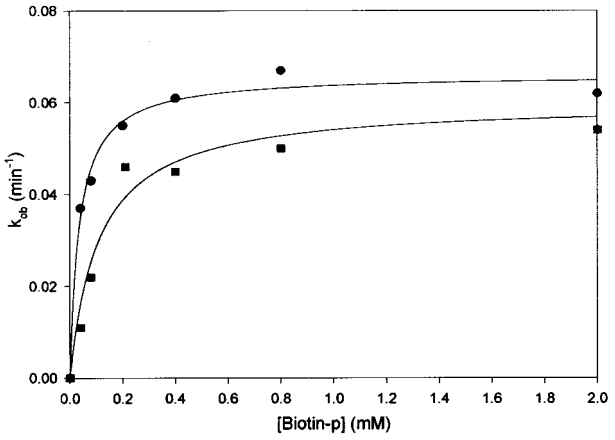


FIGURE 7: Michaelis−Menten plots showing catalytic efficiencies of CoES7 (■) and CoES21 (●) toward the selection substrate biotin-p used during the first stage of selection. Kinetics were determined in 20 mM MES at pH 6.0 and room temperature. The buffer for CoES7 contained 1 mM Mn^{2+} and 5 mM Ca^{2+}, while CoES21 used 5 mM Mn^{2+}.

activity of CoES7 and CoES21 RNA. The analysis started with one component (e.g., Mg^{2+} or Ca^{2+}) of the original selection buffer and systematically increased to include all components. Although both RNAs were isolated in a buffer solution containing Mg^{2+}, Ca^{2+}, Mn^{2+}, and imidazole, in addition to monovalent ions, CoES7 RNA requires only Ca^{2+} and Mn^{2+} for activities, while CoES21 is a solely Mn^{2+}-dependent ribozyme. Neither Mg^{2+} nor imidazole enhances catalysis. Special catalytic capabilities, not found in Mg−RNA species, were previously associated with Ca−RNA and Mn−RNA species (*8*, *12*). While several characterized manganese-dependent ribozymes possess a common GAAA/UUU motif for Mn^{2+} (*13*−*15*), the two CoES ribozymes have no such motif. The origin of particular unusual RNA activities in the presence of particular divalents deserves investigation (*16*). Optimal conditions for CoES7 RNA were 1 mM Mn^{2+}, 5 mM Ca^{2+}, 20 mM MES (pH 6.0) and for CoES21 RNA were 5 mM Mn^{2+} and 20 mM MES (pH 6.0).

Catalytic Efficiency. Kinetic analysis on CoES7 and CoES21 was conducted to assess the catalytic efficiency of these ribozymes. At each substrate (biotin-p) concentration, product yield−time data were fitted (nonlinear curve fit by SigmaPlot, Jandel Scientific) to a first-order kinetic equation to obtain the apparent first-order reaction constant, k_{ob}. Enzymatic parameters k_{cat} and K_M were then derived by fitting k_{ob}−[biotin-p] data to the Michaelis−Menten equation as shown in Figure 7. CoES7 has k_{cat} and K_M values of 0.060 ± 0.005 min^{-1} and 0.11 ± 0.02 mM, respectively, while the corresponding values for CoES21 are 0.066 ± 0.005 min^{-1} and 0.036 ± 0.010 mM. Both RNAs therefore give evidence of a substrate−enzyme complex and, within 3-fold, exhibit similar catalytic efficiencies (k_{cat}/K_M).

There are less frequent active RNA sequences within both 30*N* and 60*N* size groups. Structural and kinetic characterization of these ribozymes is being pursued. This may clarify the relationship between the RNA size and catalytic activity, one of our original goals in using length-heterogeneous random RNA libraries for selection.

DISCUSSION

Through a two-stage in vitro selection, we have isolated RNA sequences (CoES) capable of synthesizing RNA-linked CoA, NAD, and FAD from their corresponding coenzyme precursors. Using 5′-ATP-linked RNA as one of the substrates (as an ATP analogue), these reactions mimic the last steps of biosyntheses of dephospho-CoA, NAD, and FAD catalyzed by the protein enzymes dephospho-CoA, NAD, and FAD pyrophosphorylases, respectively.

We have previously characterized a self-capping ribozyme (Iso6) (*8*) that performs a variety of phosphoryl coupling reactions (*17*). Iso6 is 5′-GTP terminated. Iso6 and the current CoES ribozymes probably catalyze similar chemical reactions. Each links a terminal phosphate-containing nucleophile to the 5′-RNA, forming a diphosphate linkage between the nucleophile and RNA, with concurrent release of pyrophosphate. However, CoES7 and CoES21 differ from prior catalysts in a biologically significant sense. (1) Adenosine and its derivatives play dominant roles in biology (*18*), and CoA, NAD, and FAD all contain an adenosine. (2) When the 5′-GTP of Iso6 was replaced with 5′-ATP, the activity was lost. Therefore, the Iso6 activity cannot make coenzymes CoA, NAD, and FAD (F. Huang and M. Yarus, unpublished results), and it was unclear whether this was a fundamental or an incidental limitation.

Although the three ribozymes Iso6, CoES7, and CoES21 were isolated under similar divalent metal ion conditions (in the presence of Mg^{2+}, Ca^{2+}, and Mn^{2+} at pH ∼6), they appear to be distinct in several ways. First, they were isolated by very different methods. Iso6 was selected on the basis of 5′-pyrophosphate release, and then coupling to bead-linked ATP (*8*). CoES7 and CoES21 were selected by their ability to react with biotin-p (during the first stage selection) and then phosphopantetheine. Second, primary sequence comparison does not reveal sequences common to the three. Third, they have different divalent metal ion requirements. Iso6 is active in the presence of Ca^{2+} (much lower activity with Mn^{2+}) (*8*); CoES7 depends on both Ca^{2+} and Mn^{2+}, while CoES21 requires only Mn^{2+}.

On the other hand, there exist similarities among the three ribozymes. (1) Although all were isolated in the presence of Mg^{2+}, Mg^{2+} neither is required nor enhances activity. In fact, Mg^{2+} is inhibitory for the activity of Iso6 (*11*). (2) The catalytic efficiencies of these ribozymes appear to be similar: Iso6 has a k_{cat} of 0.077 min^{-1} and a K_M of 0.2 mM for GMP (*19*), comparable to those of CoES7 and CoES21 (above).

While transcription by T7 RNA polymerase under the T7 class III promoter (*9*) provides GTP-terminated RNA transcripts that have been widely used in various applications, our ATP-initiated transcription under the T7 class II promoter is comparably efficient. As shown in reaction 2, our selection required such ATP-terminated RNA libraries. Similarly, transcription under the T7 class II promoter may find a wide range of utility in other studies of 5′-A-RNA.

In this work, we used mixed length-randomized libraries to select RNA activities. Such random libraries ensure that different sizes of RNA compete under identical conditions. Therefore, RNA size−activity relationships may be studied. It has been argued that longer randomized regions are uniformly advantageous for selection (*20*), and we do observe

greater coenzyme synthetase activity among 60*N* molecules than in the 30*N* group. However, our results with larger randomized tracts suggest an optimal random tract size for selection of certain activities. That is, our longest randomized regions quickly became rare, and did not reappear. Any advantage of longer sequences may therefore not compensate for the inefficiency of replication, particularly under the primitive conditions of an RNA world. These experiments should be repeated under conditions permissive for longer sequences, to determine if (as we surmise) there are disadvantages of long randomized tracts not related to difficulty in amplification. Selections for other RNA activities and using other selection conditions would also likely be informative.

With coenzyme precursors plausibly available through prebiotic synthesis (*21*−*24*), CoA, NAD, and FAD could have been utilized in an RNA world. RNA-catalyzed acyl group transfer via thioester intermediates therefore is plausible. Such reactions are essential to anabolic processes in which larger molecules are synthesized from CoA-linked subunits. RNA-catalyzed electron/hydride transfer reactions involving forms of NAD and FAD could also have played roles in an RNA-catalyzed energy metabolism. Although none of these RNA reactions have yet been demonstrated, their selections are newly plausible. However, because of the reactivity of CoES7 and CoES21 with varied nucleophiles, we will need to address the problem of forming specific 5′-coenzymes. In view of the substantial K_M's we have observed, it is likely that substrate-specific coenzyme synthetases could now be developed from these unspecific progenitors, CoES7 and CoES21.

Our results suggest an explicit origin for ribonucleotide coenzymes. We suggest that CoA, NAD, and FAD originated according to reactions 3−5. That is, the coenzymes' AMP moiety was initially the 5′-nucleotide of pppA−RNA. Among the common nucleobases, adenine is the most readily formed under plausible prebiotic conditions (*25*). ATP and its derivatives may have acquired their predominant roles in current biology via this circumstance (*18*). ATP-initiated 5′ → 3′ RNA synthesis, as is today universal, necessarily yields pppA−RNA. After incorporation of pan-p, NMN, and FMN into such RNA, the resulting coenzymes CoA, NAD, and FAD would have become prosthetic groups linked covalently to the 5′-RNA terminus.

Such 5′-linkage simultaneously solves several fundamental chemical problems for primitive ribozymes. It solves the problem of reproducibly fixing a unique reactive group with respect to internal active sites, an otherwise difficult problem (for example, compare the extents of incorporation of cofactors during RNA synthesis). While RNA binding sites for coenzymes might be evolved (*26*, *27*), such sites require fixing many nucleotides and bury functional cofactor groups. Covalent 5′-linkages cost only the (small) number of nucleotides required to interact in trans with a cofactor synthetase, and leave most cofactor functionality unhindered. Such a 5′-covalent linkage also couples a reactant (coenzyme) to the enzyme, potentially speeding reactions by reducing molecularity and also by preventing loss of intermediates via diffusion.

With regard to origins, conserved 5′-nucleoside pyrophosphates in coenzymes are particularly informative. Such 5′-linkage can only be 5′-terminal in RNA synthesized by the presently universal route. Indeed, 5′-nucleoside pyrophosphates follow naturally from a synthetic route that uses phosphate attack on an activated RNA 5′-triphosphate terminus, as in our RNAs. Once several ancient coenzyme-centered pathways existed, the structures and functions of coenzymes would have been conserved; any alteration would have disrupted multiple aspects of metabolism. Thus, these otherwise puzzling structural features are explained by conservation of ancestral 5′-RNA reaction products. Such reactions also parallel current pathways for coenzyme synthesis (*28*), suggesting that primordial synthetic logic may still be utilized.

ACKNOWLEDGMENT

We thank Alecia Nix, who conducted kinetic analysis on CoES7, Gordon Cannon and Sabine Heinhorst for the use of their gel documentation and analysis system, and the members of our laboratories for comments on a draft manuscript.

REFERENCES

1. Benner, S. A., Ellington, A. D., and Tauer, A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7054−7058.
2. Gilbert, W. (1986) *Nature 319*, 618.
3. White, H. B. (1976) *J. Mol. Evol. 7*, 101−104.
4. White, H. B. (1982) in *The Pyridine Nucleotide Coenzymes* (Everse, J., Anderson, B., and Yu, K.-S., Eds.) pp 1−17, Academic Press, New York.
5. Chen, P. S., Toribara, T. Y., and Warner, H. (1956) *Anal. Chem. 28*, 1756−1758.
6. Dunn, J. J., and Studier, F. W. (1983) *J. Mol. Biol. 166*, 477−535.
7. Ciesiolka, J., Illangasekare, M., Majerfeld, I., Nickles, T., Welch, M., Yarus, M., and Zinnen, S. (1996) *Methods Enzymol. 267*, 315−335.
8. Huang, F., and Yarus, M. (1997) *Biochemistry 36*, 6557−6563.
9. Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 8783−8798.
10. Cadwell, R. C., and Joyce, G. F. (1992) *PCR Methods Appl. 2*, 28−33.
11. Huang, F., and Yarus, M. (1997) *Biochemistry 36*, 14107−14119.
12. Illangasekare, M., Kovalchuke, O., and Yarus, M. (1997) *J. Mol. Biol. 274*, 519−529.
13. Dange, V., Van Atta, R. B., and Hecht, S. M. (1990) *Science 248*, 585−588.
14. Kazakov, S., and Altman, S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 7939−7943.
15. Landweber, L. F., and Pokrovskaya, I. D. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 173−178.
16. Yarus, M. (1993) *FASEB J. 7*, 31−39.
17. Huang, F., and Yarus, M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 8965−8969.
18. Wilson, J. E. (1984) *J. Theor. Biol. 111*, 615−623.
19. Huang, F., and Yarus, M. (1998) *J. Mol. Biol. 284*, 255−267.
20. Sabeti, P. C., Unrau, P. J., and Bartel, D. P. (1997) *Chem. Biol. 4*, 767−774.
21. Dowler, M. J., Fuller, W. D., Orgel, L. E., and Sanchez, R. A. (1970) *Science 169*, 1320−1321.
22. Miller, S. L., and Schlesinger, G. (1993) *J. Mol. Evol. 36*, 302−307.
23. Miller, S. L., and Schlesinger, G. (1993) *J. Mol. Evol. 36*, 308−314.
24. Keefe, A. D., Newton, G. L., and Miller, S. L. (1995) *Nature 373*, 683−685.

25. Ponnamperuma, C., Lemmon, R. M., Mariner, R., and Calvin, M. (1963) *Proc. Natl. Acad. Sci. U.S.A. 49*, 737−740.
26. Burke, D. H., and Hoffman, D. C. (1998) *Biochemistry 37*, 4653−4663.
27. Roth, A., and Breaker, R. R. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 6027−6031.
28. Begley, T. P., Kinsland, C., Taylor, S., Tandon, M., Nicewonger, R., Wu, M., Chiu, H.-J., Kelleher, N., Campobasso, N., and Zhang, Y. (1998) *Top. Curr. Chem. 195*, 93−142.

BI002061F