In Vitro Selection of RNA Aptamers Specific for Cyanocobalamin†

Jon R. Lorsch‡,§ and Jack W. Szostak*,‡,∥

Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, and Department of Genetics, Harvard Medical School, Longwood Avenue, Boston, Massachusetts 02115

Received August 18, 1993; Revised Manuscript Received November 10, 1993⊛

ABSTRACT: RNA receptors (aptamers) capable of specifically binding cyanocobalamin (vitamin B_{12}) have been isolated by *in vitro* selection from a pool of 5×10^{14} RNAs of random sequence. After eight rounds of selection by affinity chromatography and enzymatic amplification, the pool was dominated by two sequences. The major sequence, comprising 60% of the pool, was studied further. It was found to bind vitamin B_{12} in solution with a dissociation constant (K_d) of approximately 320 ± 90 nM and to bind cobinamide dicyanide with a K_d of 8.8 ± 0.5 μM. The aptamer does not detectably bind adenosylcobalamin (coenzyme B_{12}). The selection was conducted in 1 M LiCl, and binding is dependent on the presence of high concentrations of Li^+ but independent of Mg^{2+}. To define the binding site for cyanocobalamin, a second cyanocobalamin-binding selection was carried out using a pool of sequences derived from the major aptamer sequence randomized at a level of 30%. The sequence data from this selection revealed a 31-base highly conserved region, on the basis of which was synthesized a smaller aptamer of 35 nucleotides. This small aptamer binds cyanocobalamin in solution with a K_d of 88 ± 19 nM and cobinamide dicyanide with a K_d of 20 ± 9 μM. This aptamer has the highest affinity yet reported for a small molecule ligand. A number of covarying positions were found in the conserved region of the sequences from this second, mutagenized pool selection. On the basis of these data, an unusual pseudoknot secondary structure is proposed for the aptamer. Chemical modification protection experiments are consistent with this structure and have demonstrated that the RNA undergoes a conformational change upon binding its ligand. Possible contacts with the cyanocobalamin have also been mapped. A third selection was carried out in which the salt specificity of the aptamer was changed from LiCl to NaCl plus $MgCl_2$. Sequence analysis of the final round pool of RNAs from this selection revealed several conserved changes from the original vitamin B_{12} aptamer sequence.

The discovery of catalytic RNA (Kruger et al., 1982; Guerrier-Takado et al., 1983) has increased the interest in attempts to assess the functional capabilities and limitations of nucleic acids. Use of the newly developed techniques of *in vitro* selection and evolution has greatly facilitated such work [for a review, see Szostak (1992)]. It has been shown that single-stranded nucleic acids are able to specifically recognize a variety of proteins (Tuerk et al., 1992; Bock et al., 1992; Tsai et al., 1992) and small molecules (Ellington & Szostak, 1990, 1992; Famulok & Szostak, 1992; Connel et al., 1993; Sassanfar & Szostak, 1993) and that it is possible to change the catalytic properties of preexisting ribozymes using *in vitro* selection (Green & Szostak, 1992; Beaudry & Joyce, 1992; Lehman & Joyce, 1993). Recently, RNA molecules with novel catalytic activities have been isolated from pools of mutagenized tRNAs (Pan & Uhlenbeck, 1992) and from pools of random sequence RNA (Bartel & Szostak, 1993). We have set out to explore further the ability of RNA to recognize small molecules by selecting for RNA receptor molecules, or aptamers, that specifically bind cyanocobalamin (vitamin B_{12}).

The cobalamin family of molecules (Figure 1) is believed to be of ancient origin (Georgopapdakou & Scott, 1977; Eschenmoser, 1988; Benner et al., 1989) and may have played a key role in the proposed transition from RNA-based life forms to more modern ones based on DNA, RNA, and proteins [Benner et al., 1989; see Reichard (1993) for an opposing view]. The transformation of ribonucleotides into deoxyribonucleotides is catalyzed, in many organisms, by an adenosylcobalamin-dependent ribonucleotide reductase [for reviews, see Stubbe (1990) and Reichard (1993)]. This modern enzyme has been proposed to have had an RNA counterpart in the preprotein world, and if so, it is likely that this ribonucleotide reductase ribozyme would have used cobalamins to perform its function (Benner et al., 1989). A number of other biologically important reactions are carried out by cobalamin-dependent enzymes, including methyl group transfers, carbon skeletal rearrangements, and diol dehydrations (Walsh, 1979). It has been proposed that many of these reactions were also catalyzed by RNA enzymes in the preprotein world (Benner et al., 1989). If this aspect of the RNA world hypothesis is correct, RNA sequences should exist that possess the ability to bind cobalamins tightly and specifically.

We report here the isolation and characterization of aptamers specific for cyanocobalamin. We show that RNA is capable of interacting specifically and strongly with cobalamin derivatives. The secondary and tertiary structures of one of the aptamers have been investigated. The data suggest that the aptamer is an unusual type of pseudoknot and that it undergoes a conformational change upon binding cyanocobalamin. We have also addressed the roles that Li^+ and Mg^{2+} can play in the formation of novel RNA structures. Our results suggest that high concentrations of LiCl may specifically stabilize certain otherwise unfavorable structures.

† This work was supported by a grant from Hoechst A. G. to the Department of Molecular Biology, Massachusetts General Hospital.

‡ Massachusetts General Hospital.
§ Harvard University.
∥ Harvard Medical School.

⊛ Abstract published in *Advance ACS Abstracts,* January 1, 1994.

0006-2960/94/0433-973$04.50/0

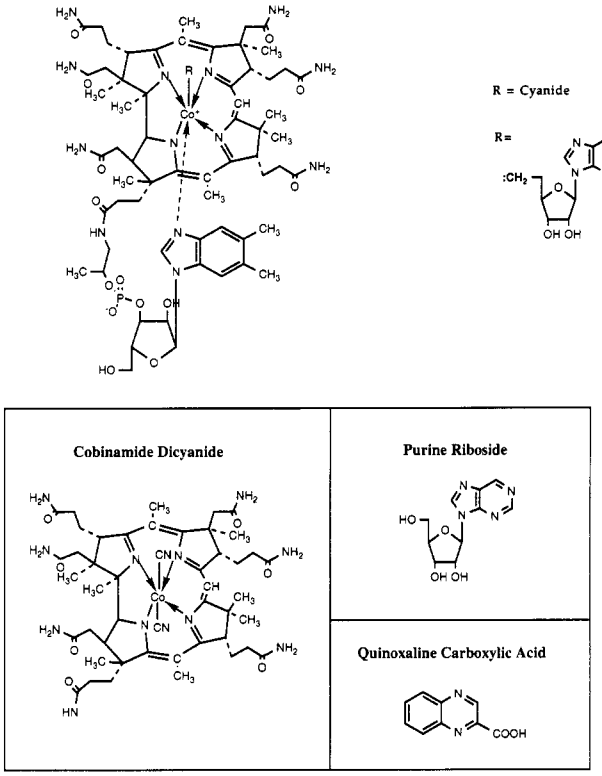

FIGURE 1: Structures of cyanocobalamin (R = cyanide), adenosylcobalamin (R = 5′-deoxyadenosine), cobinamide dicyanide, purine riboside, and quinoxalinecarboxylic acid.

MATERIALS AND METHODS

Materials. Vitamin B_{12} agarose, vitamin B_{12} (cyanocobalamin), cobinamide dicyanide, and coenzyme B_{12} were purchased from Sigma Chemical Company.

General Molecular Biology. Transcriptions from double-stranded PCR DNA and from synthetic oligonucleotides were performed as described previously (Milligan & Uhlenbeck, 1989). T7 RNA polymerase was purchased from United States Biochemicals. Reverse transcriptions were performed using superscript reverse transcriptase (Bethesda Research Labs) and were generally done in 100 μL at 41 °C using the manufacturer's recommended buffer conditions and a 3′-primer concentration of 1 μM. Taq DNA polymerase was purchased from Promega, and the buffer used was that supplied by the company. Cloning of PCR DNA products was conducted using a TA cloning kit (In Vitrogen) or a pT7 Blue T-Vector kit (Novagen). Dideoxy sequencing was performed using sequenase DNA polymerase (United States Biochemicals) and standard methods (Sambrook et al., 1989). Restriction enzymes were purchased from New England Biolabs, and digests were performed using the supplier's recommended protocols. DNA oligonucleotides were made on a MilliGen/Biosearch Model 8750 solid-phase DNA synthesizer. Denaturing polyacrylamide gels were prepared and run as described elsewhere (Sambrook et al., 1989).

Random/Mutagenized Pool Construction. The original random pool was synthesized as described previously (Bartel & Szostak, 1993). The mutagenized pool was constructed as follows: an oligo of the sequence 5′-GGAACCTCTAGGTCATTAGGAACACTATCC GACTGGCACCGCCAGCG GACAAATCCGGTGCGCATAACCACCTCAGTGC GAGCAACGATGGCCACGTCAGAAGGATCCAAG-3′ was synthesized, where the underlined region, corresponding to the sequence of the original aptamer B12.9 from the 5′-constant region to the *Hae*III site, was doped at a level of 30% as described previously (Bartel et al., 1991). In order to compensate for the higher coupling efficiencies of the guanosine and thymidine phosphoramidites relative to the cytosine and adenosine phosphoramidites, a mole ratio of 3:3:2:2 A:C:G:T was used in making up the doped phosphoramidite mix. The oligo was deprotected in concentrated ammonium hydroxide at 55 °C overnight and purified by denaturing polyacrylamide gel electrophoresis on an 8% gel (1.5 mm thick). The band corresponding to the full-length oligo was excised, crushed, and eluted overnight in 0.3 M NaCl. The final yield of purified DNA, after filtration through a 0.2-μm filter and ethanol precipitation, was 236 μg. Of this, 62.5 μg (26.5%) was found to be free of lesions by primer extension analysis. This amount of oligo corresponded to 1×10^{15} different molecules. One-half of this DNA was used for a large-scale (50 mL) PCR reaction (five cycles, four doublings) in which the 5′-PCR primer contained the T7 RNA polymerase promoter sequence. After the completion of the PCR cycles, EDTA was added to a final concentration of 5 mM, and the entire PCR reaction was phenol-extracted, chloroform-extracted, and then ethanol-precipitated. The resulting DNA was used for a large-scale transcription (10 mL). The transcription was stopped after 5 h at 37 °C by the addition of 2 mL of 0.5 M EDTA. One-third of the transcription was purified by denaturing polyacrylamide gel electrophoresis on an 8% gel (5 mm thick). After excision of the band and elution into 0.3 M NaCl, the final yield of RNA was 4.1 mg (130 pool equiv, 5×10^{14} molecules/pool equiv). About 3–5 pool equiv was used for the first round of a selection.

Selection Procedure. The selections were performed as described previously, with some modifications (Sassanfar & Szostak, 1993). ^{32}P-labeled RNA was used to follow all selections. Column volumes were typically 0.5–1 mL. Column buffers were either 1 M LiCl, 5 mM $MgCl_2$, and 25 mM HEPES (pH 7.4) or 1 M NaCl, 5 mM $MgCl_2$, and 25 mM HEPES (pH 7.4). A 0.5-mL precolumn of dihydrazide agarose was used for the initial selection. RNA in column buffer (100–150 μg in 1.0 mL in the first round; 10–20 μg in 0.2 mL in succeeding rounds) was applied to the precolumn and was washed directly onto the cyanocobalamin column with 3 column vol of buffer. RNA was loaded directly onto a cyanocobalamin column with no precolumn for the selections from the mutagnized pool. The cyanocobalamin column was then washed with an additional 12–17 column vol of buffer. Specifically bound RNAs were eluted with 3 column vol of 1 mM ligand in column buffer. The RNA was ethanol-precipitated with glycogen as a carrier. Reverse transcription, PCR, and transcriptions were performed as described above.

Affinity Elution Experiments. Routine tests for ligand binding were carried out using roughly 1–5 μg of ^{32}P-labeled RNA. The RNA was applied to a 0.5-mL cyanocobalamin column in the appropriate buffer, washed with 6–10 column vol of buffer, and eluted with 3 column vol of 1 mM ligand in column buffer. One column volume fractions were quantitated by Cerenkov counting in a scintillation counter. A correction factor of 3 was used for samples containing 1 mM cyanocobalamin, due to quenching of radiated light by the colored ligand. Similar correction factors were determined for other colored ligands.

Equilibrium Dialysis. Equilibrium dialysis measurements were performed using dialysis chambers with 200-μL volumes. The membrane had either a 6000–8000 or 3500 molecular

weight cutoff, depending on the size of the RNA. Typically, an RNA concentration of 10 μM was used. ^{14}C-labeled cobinamide dicyanide was synthesized as follows: 0.46 mg of Na^{14}CN (9.4 μmol, DuPont-New England Nuclear) and 1 mg (0.6 μmol) of cobinamide dicyanide were stirred in 200 μL of water at room temperature, in the dark. After 40 h, the mixture was loaded on a C18 Sep-Pak (Waters-Millipore), which was washed extensively with water. The ^{14}C-labeled cobinamide dicyanide was eluted with 1:1 acetonitrile/water. A stock solution of 4.6 mM [^{14}C]cobinamide dicyanide was used for equilibrium dialysis experiments. All measurements were done at least in duplicate. Measurements of cobinamide dicyanide dissociation constants (K_d[1]) were taken at three different concentrations of ligand: 0.23, 4.6, and 46 μM. The total cpm in each measured sample ranged from 1000 to 230 000, depending on the cobinamide dicyanide concentration. The K_d's for cyanocobalamin were measured by competition at 0.23 μM cobinamide dicyanide and at various cyanocobalamin concentrations. Sample volumes were 100 μL on each side of the membrane. Samples were allowed to equilibrate overnight with rotation, after which aliquots were withdrawn and subjected to scintillation counting with fluor. Since the specific activity of the cobinamide dicyanide was known, it was possible to calculate the concentration of bound ligand directly. K_d's for cyanocobalamin were calculated using the following equation:

$$r_d = 1 + K_b[\text{RNA}]/(K_c[\text{CNCba}] + 1)$$

where r_d is the observed ratio of (bound ligand + free ligand)/free ligand, K_b is the binding constant of the RNA for cobinamide dicyanide, and K_c is the binding constant for the competitor, cyanocobalamin (CNCba; Cantor & Schimmel, 1980). While these measurements give a good value for the relative affinity of the aptamer for cyanocobalamin, they are only as good in absolute terms as the original measurements of the K_d for cobinamide dicyanide. It should also be noted that most systematic errors in these experiments (i.e., dissociation of the cyanide from the cobalamin, impure or incompletely folded RNA, or RNA degradation) would tend to make our measurements overestimate the true K_d's. Our analysis assumes that there is only one binding site per aptamer.

Chemical Modification Experiments. Chemical cleavage experiments were done after the reported methods (Peattie, 1979; Peattie & Gilbert, 1980), with some modifications. 5′-^{32}P-end-labeled RNA (small aptamer; JL67.7-transcribed 35-mer RNA) was used. Diethyl pyrocarbonate (DEPC, 1.25 μL) or 0.5 μL of dimethyl sulfate (DMS) was added to labeled RNA in the appropriate buffer (1 M LiCl or 1 M NaCl plus 5 mM $MgCl_2$ with 40 mM sodium cacodylate (pH 7.4); when LiCl was used, no $MgCl_2$ was added; sodium cacodylate does not interfere with binding of the ligand by the aptamer) to a final volume of 50 μL. When cyanocobalamin was added, the final concentration was 100 μM. The reactions were incubated for 15 and 8 min, respectively, at room temperature. Denaturing controls were done by heating the RNA to 90 °C for 3 min in 40 mM sodium cacodylate (pH 7.4) and then adding the modification reagent and incubating at 80 °C for an additional minute. The reactions were stopped by adding 1 μL of 10 mg/mL tRNa, 5 μL of 3 M sodium acetate (pH 4.5), and 150 μL of cold ethanol. The samples were then frozen on dry ice. After centrifugation and washing with 70% ethanol, the DMS reactions were dissolved in 10 μL of 1 M Tris-HCl buffer (pH 8.4), and 10 μL of freshly prepared 0.2 M sodium borohydride solution was added. The reactions were allowed to proceed for 30 min on ice in the dark, after which time they were stopped by the addition of 200 μL of a cold solution consisting of 0.6 M sodium acetate (pH 4.5) plus 25 μg/mL tRNA. The reactions were then ethanol-precipitated, and the pellets were rinsed with 70% ethanol. Both the DMS and DEPC reactions were next subjected to aniline-induced cleavage. The pellets were dissolved in 20 μL of 1 M aniline–acetate (pH 4.5) and allowed to react at 60 °C, in the dark, for 20 min. The solutions were frozen on dry ice, lyophilized, redissolved in water, and lyophilized a second time. The pellets were dissolved in gel loading buffer (8 M urea, 20 mM EDTA, 2 mM Tris-HCl (pH 7.5), 0.05% bromphenol blue, and 0.05% xylene cyanol) and analyzed on a 20% denaturing polyacrylamide gel.

The reverse transcription chemical modification experiments were performed as described previously (Moazed et al., 1986; Sassanfar & Szostak, 1993). Typically, less than 1 μg of RNA was used for an experiment. The RNA was treated with DMS (0.5 μL of DMS in a 100-μL reaction) for 5–10 min at room temperature. A denaturing control was performed as follows: the RNA was heated in 5 mM sodium cacodylate for 3 min at 92 °C. DMS (0.5 μL) was added, and the reaction was allowed to proceed for 2 min at 92 °C. The reaction was then quenched as described elsewhere (Moazed et al., 1986). Reverse transcriptions were done at 48 °C using AMV reverse-transcriptase (Life Sciences). The RNA strand of the RNA–DNA hybrid was hydrolyzed with 0.2 N NaOH prior to electrophoresis on a 15% denaturing polyacrylamide gel. Dideoxy sequencing was performed as described previously (Moazed et al., 1986; Sassanfar, 1993). Dope.40 RNA (see Results) was used because of its good binding characteristics. It was necessary to use a full-length clone, as opposed to the 35-nucleotide aptamer, because of the need for a 3′-primer binding site. Addition of a 3′-primer binding site to the 35-mer RNA was found to inhibit proper folding of the molecule.

RESULTS

Cyanocobalamin Aptamer Selection. RNA aptamers specific for cyanocobalamin were selected from a pool of RNA molecules with a 72-nucleotide random region flanked by two constant regions for reverse transcription and polymerase chain reaction (PCR) amplification. The total complexity of the pool was approximately 5×10^{14} molecules. Aptamers were isolated by passing the pool RNA over a cyanocobalamin–agarose affinity column (approximately 1 mM ligand concentration), washing extensively to remove unbound and weakly bound RNAs, and then affinity-eluting specifically bound molecules with 1 mM cyanocobalamin. The column buffer was 1 M LiCl, 5 mM $MgCl_2$, and 25 mM HEPES (pH 7.4). High salt concentrations were required to decrease nonspecific interactions (presumably ionic) between the RNA and the column matrix. The resin was a mixture of the monoacid derivatives of cyanocobalamin linked to cyanogen bromide-activated agarose via an eight-atom spacer arm (Sigma Chemical Company, personal communication; Allen & Majerus, 1972). Prior to affinity purification, the RNA was passed over an agarose column. Sequences retained on this precolumn were discarded to prevent the selection of molecules that bind to the affinity column matrix. After affinity column selection, the eluted RNA was reverse-transcribed and PCR-amplified, and a new pool of RNA, enriched in cyanocobalamin-binding sequences, was transcribed from the double-stranded DNA. This new pool was

[1] Abbreviations: CNCba, cyanocobalamin; $Cb(CN)_2$, cobinamide dicyanide; DEPC, diethyl pyrocarbonate; DMS, dimethyl sulfate; K_d, dissociation constant; nt, nucleotide.

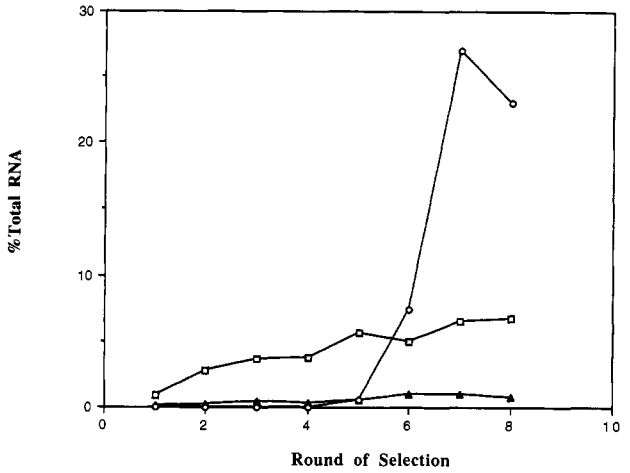

FIGURE 2: Summary of the selection for cyanocobalamin aptamers from the random sequence RNA pool. The curves show the percent of the total RNA loaded on the column that remained on the precolumn (□), was specifically eluted by cyanobalamin (○), or remained bound to the cyanocobalamin column after affinity elution (▲) at each round of selection.

used for the next round of affinity selection and amplification. A significant amount of specifically eluted RNA (~30%) was seen after six rounds of selection and amplification. After eight rounds there was no further enrichment, and the pool was cloned and sequenced. Figure 2 shows a summary of each round of selection. The total purification factor was approximately 5×10^{14}, indicating that very few sequences in the original pool can bind cyanocobalamin tightly enough to be selected.

Analysis of Cyanocobalamin Aptamers. Figure 3 shows the results of the sequencing of clones derived from the final round of the cyanocobalamin-binding selection. Two clones dominate the pool; one clone comprises 60% of the pool and the other 25%. These clones share a sequence at their 5′-ends (5′-CCAGCGGA-3′) that can form a duplex with part of the 5′-constant region of the pool (5′-UCCGACUGG-3′). RNAs of both sequences bind to cyanocobalamin agarose and are specifically eluted by 1 mM cyanocobalamin. Of the three remaining sequences, B12.25 and B12.29 also bind to cyanocobalamin agarose and are specifically eluted by cyanocobalamin, but B12.18 interacts only weakly with cyanocobalamin agarose.

A representative clone (B12.9) from the major class of aptamers was selected for further study. Elution of B12.9 RNA from cyanocobalamin agarose was found to be highly specific, in that it is eluted (>95% of bound RNA in 3 column vol) by 1 mM cyanocobalamin and 1 mM cobinamide dicyanide, but not eluted (<1% in 3 column vol) by 1 mM adenosylcobalamin, 1 mM quinoxalinecarboxylic acid, 1 mM purine riboside, or 1 mM cobalt(III) hexamine chloride. These compounds were used because of their structural similarity (or lack thereof) to chemical groups in cyanocobalamin. Binding of the aptamer to cyanocobalamin agaorse is dependent on the presence of high concentrations of Li^+. The aptamer binds only weakly in concentrations of LiCl less than 1 M and does not bind at all when 1 M NaCl or KCl is substituted for LiCl, even in the presence of $MgCl_2$. Mg^{2+} is not required for binding in the presence of Li^+, and high concentrations of Mg^{2+} cannot substitute for Li^+.

Cyanocobalamin Binding Site. To define the binding site for cyanocobalamin, 3′-truncated derivatives of the B12.9 aptamer were synthesized by cutting the double-stranded PCR DNA encoding the B12.9 clone (and the T7 RNA polymerase promoter) with either *Fsp*I, *Mnl*I, *or Hae*III restriction enzymes and generating run-off transcripts from these shortened DNA templates. Figure 3 shows the positions of these restriction sites. The truncated RNAs were tested for binding to cyanocobalamin agarose and for elution by cyanocobalamin. Only the *Hae*III-truncated RNA maintained the ability to bind to cyanocobalamin. Thus, the 3′-border of the binding site lies between the *Mnl*I and *Hae*III sites.

To further define the binding site, a second cyanocobalamin-binding selection was conducted using a pool of variants of the B12.9–*Hae*III fragment, including the former 5′-constant region, partially randomized so as to yield 30% mutation at each nucleotide position. The total length of the pool RNA was 111 bases (including two new primer binding sites), with 75 mutagenized bases, and a sequence complexity of approximately 5×10^{14} molecules. This second cyanocobalamin

```
                                FspI                MnlI     HaeIII

B12.9   GCCAGCGGACAAATCCGGTGCGCATAACCACCTCAG-TGCGAGCAACGATGGCCTTTCTACCCAAAGATTTT
B12.2   GCCAGCGGACAAATCCGGTGCGNATAACCACCTCAG-TGCGAGNAACGATGGCCTTTCTACCCAAAGATTTTT
B12.3   NNNAGCGGACAAATCCGGTGNNNATAACCACCTCAG-TGCGAGNNANGATGGCCTTTCTACCCAAAGATTTTT
B12.10  GCCAGCGGACAAATCCGGTNCGCATAACCACCTCAG-TGCGAGCAACGATGGCCTTTCTACCCAAAGATTTT
B12.13  GCCAGCGGACAAATCCGGTNCGCATAACCACCTCAG-TGCGAGCAACGATGGCCTTTCTNCCCAAAGATTTT
B12.21  GCCAGCGGACAAATCCGGTGCGCATAACCACCTCAG-TGCGAGCAACGATGGCCTTTCTACCCAAAGATTTT
B12.22  GCCAGCGGACAAATCCGGTGCGCATAACCACCTCAG-TGCGAGCAACGATGGCCTTTCTACCCAAAGATTTT
B12.27  GCCAGCGGACAAATCCGGTGCGCATAACCACCTCAG-TGCGAGCAACGATGGGCATTCTACCCAAAGATTTTT
B12.20  ---- CGGACAAATCCGGTGCGCATAACCACCTCAG-TGCGAGCAACGATGGCCTTTCTACCCAAAGATTTTT
B12.16  CCAGCGGCACAAATCCGGTGGGCATAACCACCTCAGCTGCGAGNAACGATGGCCTTTCTACCCAAAGATTTTT
B12.30  GCCN-CGGACAAATCCGGTGCGCATAACCACCTCAGGT-----CAACGATGGCCTTTCTACCCAAAGATTT
B12.6   ----------CAATCCGGTNNTAACNACCTCAGNNNAT-----------GGCCTTTC-ACNNAAAGATTTTT

B12.8    CCAGCGGAACGAAACAGGGTGTCCAGCTTGTGNNCGAGACAGGATTCATGGGAGCTATTCGCATNGNAAAATC
B12.11  CCAGCGGAACGAAACAGGGTGTCCAGCTTGTGGACGAGACAGG-TTCTTGGGANCTATTCGCATCGCAAAANN
B12.14  CCAGCGGAACGAAACAGGGTGTCCAGCTTGTGGACGAGACAGG-TTCATGGGAGCTATTCGCATCGCAAAATC
B12.19  CCAGCGGAACGAAACAGGGTGTCCAGCTTGTGGACGAGACAGG-TTCATGGGAGCTATTCGCATCGCAAAATC
B12.23  CCAGCGGAACGAAACAG-GTGTCCAGCTTGTGGACGAGACAGG-TTCATGGGAGCTATTCGCATCGCAAAATC

B12.18  CGAAAACGTTTAACCGAATCGTGGACCGTCCGAACCGCAGGCCTGTGAACGCGTATTCGCCTTCCGNGCG

B12.25  GAGAACCCCAAGTTGCTTCAAGGGCGTGACAAAGTCTAGCCTATATAGATCTAAGGTAAGGGGAATTCAAGCA

B12.29  GTGGAATCGCAAGTGGGCCTCGAGGGAACCACAACGAGCGCTGGGCGATAGTCCTTGAATAAGTGATATCG
```

FIGURE 3: Sequences of clones from the final (eighth) round of the cyanocobalamin-binding selection. The restriction sites used to map the binding site of the major clone are underlined in clone B12.9. The enzyme *Mnl*I cuts seven bases to the 3′-side of the sequence 5′-CCTC-3′. Large deletions may be artifacts of cloning, and molecules with deleted regions are not necessarily functional.

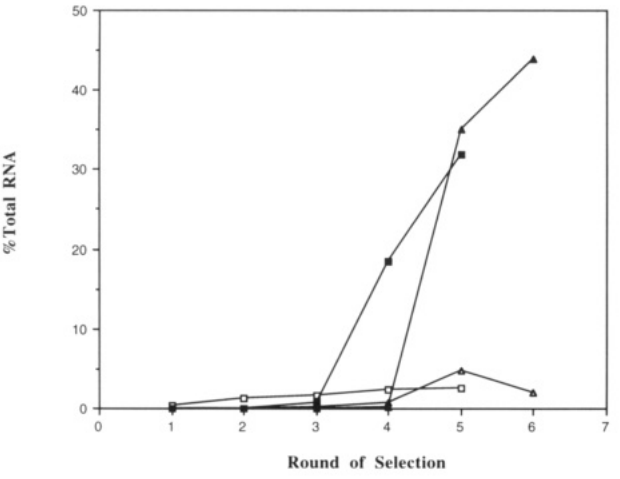

FIGURE 4: Summary of the selections from the mutagenized B12.9–*Hae*III pool. The curves represent the percent of the RNA loaded on the column that was specifically eluted by cyanocobalamin (■ for Li^+ selection and ▲ for Na^+/Mg^{2+} selection) or remained bound to the column after affinity elution (△ and □ for Li^+ and Na^+/Mg^{2+} selections, respectively) for each round of selection.

selection was conducted in the same manner as the original selection (Figure 4). After five rounds of selection and amplification, binding was restored to levels comparable to those of the original aptamer. Sequence analysis of 45 clones from the fifth round pool revealed a highly conserved 31-nucleotide motif (Figure 5). A number of positions in this motif were found to covary in a manner consistent with Watson–Crick base-pairing (Figures 5 and 6c). Comparative analysis of the sequence data suggests a pseudoknot as the secondary structure of the aptamer (Figure 6a). In this structure, bases G3–U5 (see Figure 6a for base numbering) form a duplex with A16–C18, and C7–A10 form a duplex with U23–G26. U11–C15 and U19–G22 make up the loops joining the two helical regions. One pair of covarying bases (C15,G22) also appears to provide us with a tertiary constraint on the structure.

Evidence of the existence of the first helix is given by the covariational data. Chemical modification experiments provide the only evidence for the formation of the second helix. Clones representing each class of covariants were tested for their ability to bind cyanocobalamin. The clone Dope.19, which has a G3 to A mutation but not the compensatory C18

```
                                        1   5   10   15   20   25   30          Binding
WT      GGAACACTATCCGACTGGCACCGCCAGCGGACAAATCCGGTGCGCATAACCACCTCAGTGCGAGCAACGATGGCC
DOPE.1  CGAATTGTATCCTAGGGGCATAGTCAGGAGACTAAACCGGTGCGCATAACCACCTCAGTGCGAGCAACTTGTCCGCG
DOPE.2  GGAACCGTATCCGACTTAGACCGCCAGAGG-CAGATCCGGTGCGCATAACCACGTCAGTGCGAGCAACCTTGGCC
DOPE.3  TGGACTTTGGAGGACTGGCACCGCCAGAGATCAAATCCGGTGCGCATAACCACCTCAGTGCGAGCAACTTTGACC
DOPE.4  GAAACACAATGCGACTGACACGGCCAAGAGACAATTCCGGTGCGCATAACGACCTCATTGCGAGCAACT-TGTCC      +
DOPE.5  GGATTTGTATCCGACAGAAACCGACGACCGAGAGATCCGGAGCGCATAACCTCCTCAGTGCGAGCAACGGTTGTCC     +
DOPE.6  GGAGCCCTATCTC-CGGGCACGGCCAGGGGACGAATCCGGTGCGCATAACCACCTTAGTGCGAGCAACT-T-GCCT
DOPE.7  GGAACCACCTG-GACCGCCA-GACAACAAGAGACTTCCGGTGCGCATAACCACCTTAGTGCGAGCAATGTTGGAT
DOPE.8  GCCAGTCTCCCTGACTGGCACTGCCAGCTTAGAATTCCGGTGCGCATAACCACCTTAGTGCGAGCAACTTGGCCC
DOPE.9  CGAACGCAGTCCGACACG-ACGGCCTG-GGA---TTCCGGTGCGCATAACCACCTCAGTGCGAGCAACGATCCTC
DOPE.10 GGTCCACTATCAG-TGAGAAC-GCCAAAGTTAAAATCCGGTGCGCATAACCACCTTAGTGCGAGCAACTACCTTC
DOPE.11 GTAACAGCATTAATCTGGCATCGCTGGAGAAAAATTCCGGTGCGCATAACCACCTCAGTGCGAGCATCGAT-CCC
DOPE.12 GGAACACGAACTG-CGAGCACAGCCAGCGTCGAAACAAGGTGCGCATAACGACCTCAACGCCGCCACCGATTGAC      +
DOPE.13 GGAGCACGATCTTATCGGCACCGCCAGAGGAAAAATCCGGTGCGCATAACCACCTCAGTGCGAGCAAGTTGGAC-
DOPE.14 CGAAAACTATNCGTCAGACACTGCTAGAGGACAATTCCNNNNCGCATAACCACCTCAGTGCGAGCAACAATATCC
DOPE.15 GGGGCACGCTCCTACTGGC-CNGC-GGAGG--AA-TCCGGTGCGCATAACCACCTGAGTGCGAGCAACTTTGGCC
DOPE.16 GGTACACTATCCTACTGGNGCNGC-AGAGGACAAAGCCGGTGCGNATAACCACCTTAGTGCGAGNAATTTTGGAC
DOPE.17 GGATTGGTTTCCCACTGACCCCACCAGCCTAAGGTTTCGGTGCGCATAACCACGTCAGTGCGAGCAATGTTGGCC
DOPE.18 GGAACACCATGCAAATGGCTGCGCTAA-GGAGAATTCCGGTGCGCATAACCACCTCAGTGCGAGCAACTCGGGTC
DOPE.19 TGCTTCCTATCAGGCTGGTACAGACATAGGGGAAAACCAGTGCGCATAACCACCTTAGTGCGAGCAATCTTGGCA      -
DOPE.20 GGAACACGCCGCGTCTAGCCACCCTAGGGGAAAACACCGGTGCGCATAACCACCTCAGTGCGAGCAAC-TTGGCC
DOPE.22 GTAAGGCCACCCGATTGGGACCGCTAGAGGACAAATCCGGTGCGCATAACCACCTCAGTGCGAGCAAC-TTGGGT
DOPE.23 GGCACTCACCCCGAGTGTCACTGCCATCGTGAGGATCCGGTGCGCATAACCACCTCAGTGCGAGCAACTTTGATC
DOPE.24 GGACCATTCTCCGTTTGGTACGGCGAACAGAGTAATCCGGTGCGCATAACCACCTTAGTGCGAGCAACGAGGTCC
DOPE.25 CGAAAGATATCTGATTGGCTCGGACGGCGGACCGAGCCGGTGCGCATAACCACCTTAGTGCGAGCAAAGATGGCC
DOPE.26 GGAAGACTAGCTGACTGGCACCGCCAGAGGATAAATCCGGTGCGCATAACCACCTCAGTGCGAGCAACTTTGATC
DOPE.27 GGAACCCGGTGCACTAAGCACCGCCACCGACCAATTCCGGTGCGCATAACGACGTCATTGCGAGCAACCTTGGCC      +
DOPE.28 AGAAAACTATCTGACGGGCGTCGGCTTCGAACCGATGCGGTGCGCATAACCACCTTAGTGCGAGCAAACATGGCC
DOPE.29 AGAACAGG-TCTGACTTGCACTCCCAGCGTAAGAGTGCGGTGCGCATAACCACCTTAGTGCGAGCAACATTGACC
DOPE.30 GGAACACTATCCTATTCTCACAGACTGAGGAGTTTTCCGGTGCGCATAACGACCTCATTGCGAGCAACTTTGGTC
DOPE.31 GGAGCACTATCCGTGTTGCATCGGTAGCCTAGAGGTCCGGTGCGCATAACCACCTCAGTGCGAGAACTTGGTC--
DOPE.32 GGCACGCTATCCGCCTAGGA----------AGAGGTTCGGTGCGCATAACCACCTCAGTGCGAGCAACTTTGGAA
DOPE.34 CGAACAGAATCCGAGTCGTACCGCCAGCGGTTTAAACCGGTGCGCATAACCACCTCAGTGCGAGCAAACTTGTTC
DOPE.35 TCTACACGGCCAGCCTGGCACCGTAAGCTGACCATTCCGGTGCGCATAACCACCTTAGTGCGAGCAATCTTGGGC
DOPE.37 GGAGCACTATTCGAGAGGCTCCGAGCGCCGTGAGTTCCGGTGCGCATAACCACCTCAGTGCGAGCAATTTTGGAT
DOPE.38 GAAACAACATTGNCTTGAA-CGGNNNGTGAGGACTTCCGGTGCGNATAACCACGTCAGTGCGAGNNACTTTGGTC
DOPE.39 GGAACCCTACCACACTATGACCGCCAGAGGAAATATCCGGTGCGCATAACCACCTCAGTGCGAGCAACTTTGGCC
DOPE.40 GTCGGCCTATCCGACAGGCACCGCGAGAGGACCATTATAGTGCGCATAACCACGTTCAGTGCGAGCAAAAATTTGG      +
DOPE.41 AGAATCATATCGGATTGGCACCGCCTGCGTGCAAATCCGGTGCGCATAACCACGACAGTGCGAGCAACGATGTCC
DOPE.42 AGAACCGAATCGGACTGTGATCGCCGGCGGTCACGTCCGGTGCGCATAACCACGTCAGTGCGAGCTACTATGGCC      +
DOPE.43 CGAAAACTATTTGGCGGGTACCTCCATGAGACCCG-CCGGTGCGCATAACCACGTCAGTGCGAGCAACGATGGCC
DOPE.44 GCAACAGTCTCCGACTTGCGAGGCCAGGGAACTATTAAGGTGCGCATAACGACCTCATGTGCGAGCAACTTTGGTC
DOPE.45 AGAACATTTTTCGATAANCCCCGCTAACGGACGAATCGGGTGCGCATAACCACCTCAGTGCGAGCAACGATGACC
DOPE.46 GGACCATGATCCGCCAGGCCCAACCT--GGAACNTTCCGNTNCGCATAACGNNCTCATGTGCGAGCAACTTGNNC-
DOPE.47 CGAATAGGACCCTACGGGAACGCCCAGAGGACAATTCCGGTGCGCATAACGACCTCATGTGCGAGCAACGTTGGAT      +
DOPE.48 GGACCACTAGCN--CTGGCCCGGNNGGCGGGCACTTCCGGTGCGCATAACGACCTCATGTGCGAGCAAC-TTGGTC
```

FIGURE 5: Results of sequencing of clones from the selection for Li^+-dependent cyanocobalamin aptamers from the mutagenized B12.9–*Hae*III pool. The wild-type (WT) sequence is the sequence of the B12.9–*Hae*III fragment, including the original 5′-primer binding site. Changes from wild-type are shaded. Changes from wild-type that covary in a nanner consistent with Watson–Crick base-pairing are in boxes. Clones from each class of covariants and clones that did not fit the structural model were tested for their ability to bind cyanocobalamin (see text). The results of these experiments are indicated by a + or – on the right side of the sequences.

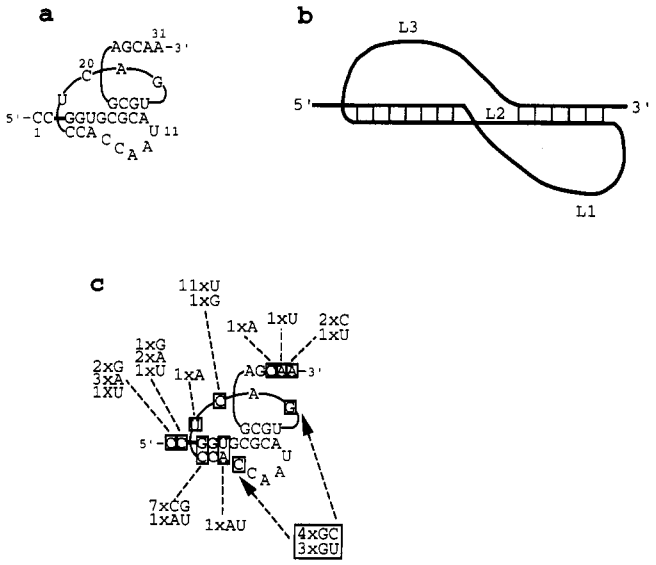

FIGURE 6: (a) Proposed secondary structure of the Li^+-dependent cyanocobalamin aptamer. Base numberings are shown next to the figure. The small (35-nt) aptamer had a 5′-GGAA-3′ tail on its 5′-end to facilitate transcription. (b) Conventional nomenclature of pseudoknots (Abrahams et al., 1990; ten Dam et al., 1992). (c) Varying positions in the highly conserved sequence region from the mutageized pool selection. Positions that covary are labeled with the number of times they change to each other base pair, and bases that vary independently are labeled with the number of times and the bases to which they change (total of 45 sequences). Clone Dope.19 was excluded because it was found not to bind cyanocobalamin. Base changes from Dope.12 also are not shown.

to T mutation (as seen in the clone Dope.40), does not bind to cyanocobalamin agarose, suggesting that a Watson–Crick base-pair at this position is essential for the formation of the proper aptamer structure. Since complete purification is not possible, a few nonfunctional clones would be expected in the final pool of RNA. All of the other clones tested, including Dope.40, bind to cyanocobalamin agarose and are specifically eluted by cyanocobalamin. All of the active sequences except clone Dope.12 are consistent with the model shown in Figure 6a. Dope.12 may be an alternative structural solution that utilizes some, but not all, of the structural elements of the other clones. It can still form the first duplex (5′-GGU...ACC-3′), but the sequences required for formation of the second helix (5′-CGCA...UGCG-3′), which are conserved in all other clones, are not entirely retained.

Based on the sequence data from the mutagenized pool selection, a smaller aptamer (35 nucleotides long) was made by run-off transcription of a synthetic DNA oligonucleotide (Figure 6a). This shortened version of the cyanocobalamin aptamer binds cyanocobalamin agarose and is eluted specifically by cyanocobalamin. It has the same salt requirements as the full-length aptamer (B12.9).

It is interesting to note that the 5′-conserved region from the two major clones of the original selection (see above) proved not to be a part of the binding domain. It is possible that the formation of a duplex with the constant region prevents the constant region from interfering with the formation of the functional aptamer structure. Examples of this type of inhibition of proper folding by external sequences and the restoration of folding by the addition of "anti-inhibitory" sequences have been observed in our laboratory (C. Wilson and J. W. Szostak, unpublished observations).

Solution Binding and Measurement of K_d's. The fact that the aptamers are specifically eluted from cyanocobalamin–agarose by cyanocobalamin is good evidence that they bind

Table 1: Dissociation Constants of Cyanocobalamin Aptamers

receptor molecule	buffer[a]	ligand	K_d (μM)
B12.9	1 M LiCl	$Cb(CN)_2$	8.8 ± 0.55
B12.9	1 M LiCl	CNCba	0.32 ± 0.09
35-mer aptamer	1 M LiCl	$Cb(CN)_2$	19.7 ± 8.7
35-mer aptamer	1 M LiCl	CNCba	0.087 ± 0.019
35-mer aptamer	0.5 M LiCl	$Cb(CN)_2$	711 ± 220
35-mer aptamer	0.25 M LiCl	$Cb(CN)_2$	>1000
35-mer aptamer	1 M NaCl/ 5 mM $MgCl_2$	$Cb(CN)_2$	>1000
tRNA	1 M LiCl	$Cb(CN)_2$	>1000

[a] Buffers contained 25 mM HEPES (pH 7.4) and the salts indicated in the table.

their ligand in solution, as well as on the column. Using equilibrium dialysis, we have measured the dissociation constants (K_d's) of the aptamers for cyanocobalamin and cobinamide dicyanide in solution (Table 1). The small 35-nucleotide aptamer binds to its ligands with affinities similar to those of the original full-length clone. Both bind to the cobinamide part of the molecule with good affinity (low micromolar K_d), but bind even more tightly (~100 nM K_d) to the entire cyanocobalamin molecule. These data indicate that the aptamer interacts with both the cobinamide and the dimethylbenzimidazole ribonucleotide portions of the molecule.

Chemical Modification Protection Experiments. In order to probe the tertiary interactions in the small aptamer, we carried out chemical modification experiments using dimethyl sulfate/$NaBH_4$ or diethyl pyrocarbonate modification followed by aniline-induced strand cleavage (Figures 7 and 8). These experiments examine the sensitivity of the N7 positions of guanosine and adenosine, respectively, to chemical modification. Five bases were found to become protected from modification upon binding to cyanocobalamin to the aptamer: A10, A12, A13, and A16 were protected from DEPC modification, and G26 was protected from DMS modification. These positions may be involved in contacting the ligand or in new, ligand-induced intramolecular contacts. Two other positions, G8 and G22, become hyperexposed upon the addition of ligand, indicating that the aptamer undergoes a conformational change when it binds cyanocobalamin. No ligand-induced changes were observed in 1 M NaCl/5 mM $MgCl_2$ buffer, consistent with the lack of ligand binding under these salt conditions.

Dimethyl sulfate methylation protection experiments using reverse transcriptase primer extension analysis (Figure 9), which probes the exposure of Watson–Crick hydrogen-bonding groups, demonstrate significant protection of several of the bases in the proposed second duplex region (C7–A10, U23–G26), relative to other positions in the molecule and to denaturing conditions (5 mM sodium cacodylate (pH 7.4), 92 °C, 2 min): C25, G24, U23, and G22 are all protected from DMS modification, as are G8, C7, G6, U5, and G4. In our hands, under the buffer conditions used (1 M LiCl and 40 mM sodium cacodylate (pH 7.4) or 1 M NaCl and 40 mM sodium cacodylate (pH 7.4), DMS modification was not base-specific. The reason for this lack of specificity is not clear, but could be due to the high salt concentrations used.

Change in Ion Dependence. We were interested in determining whether the salt requirements of the aptamer could be changed by *in vitro* selection. Using the same pool of variants used to define the binding site of the B12.9 clone, we were able to select for RNAs capable of binding cyanocobalamin in the presence of NaCl and 5 mM $MgCl_2$. A summary of this selection is shown in Figure 4. The restoration

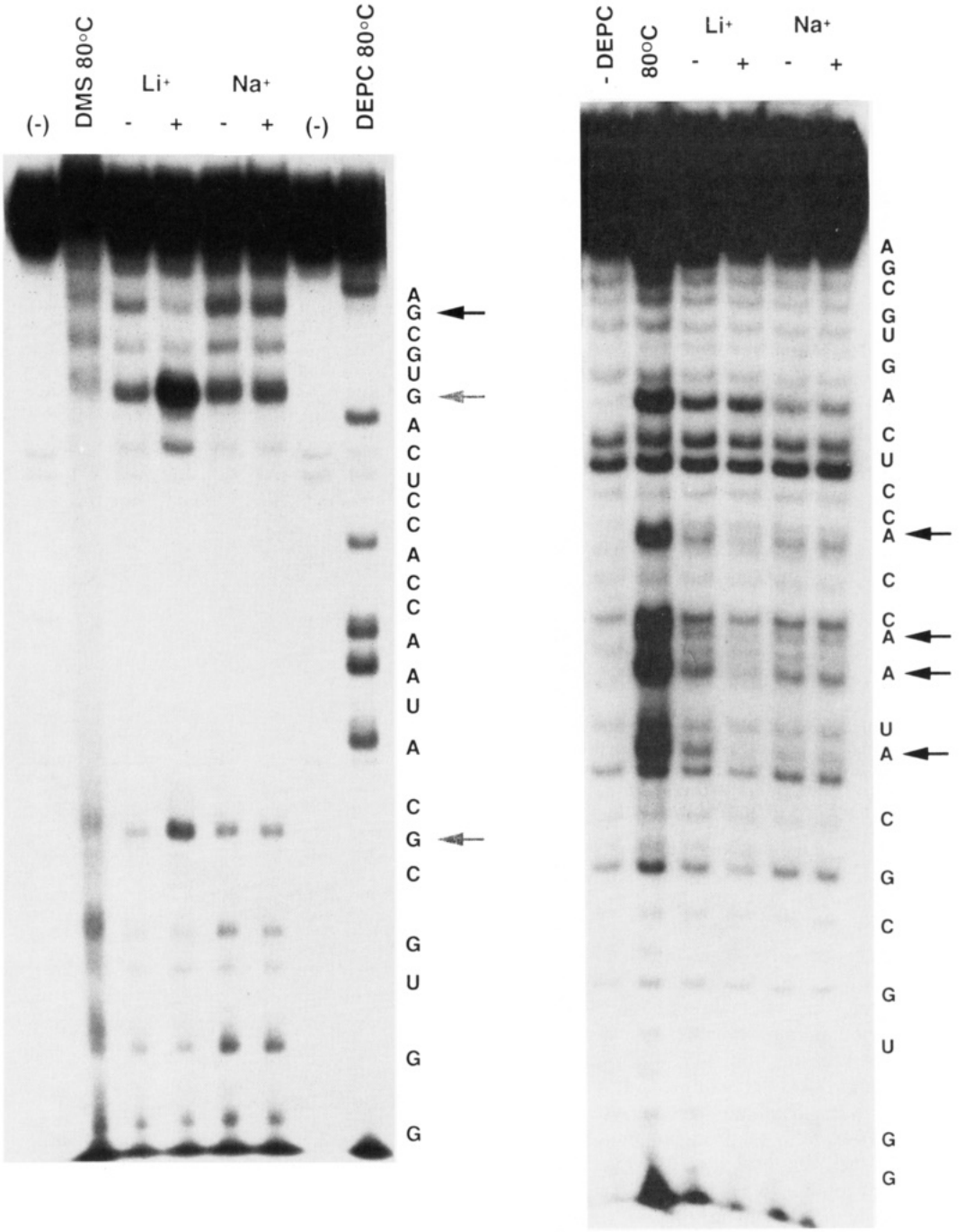

FIGURE 7: Autoradiograms of typical gels from chemical modification protection experiments with the small, 35-nt aptamer. (a, left) DMS protection experiment. Abbreviations: (–), no DMS treatment ($NaBH_4$ and aniline treatment were performed, however); DMS/DEPC 80 °C, denaturing DMS/DEPC control; Li^+/Na^+, experiments performed in 1 M LiCl or 1 M NaCl containing buffers. + and – refer to the presence or absence of 100 μM cyanocobalamin during chemical modification. (b, right) DEPC protection experiment. Abbreviations: –DEPC, no DEPC treatment (aniline treatment performed); 80 °C, DEPC-denaturing control; other abbreviations are the same as in a. Dark arrows show the positions of ligand-dependent protections. Light arrows indicate ligand-induced hyperexposure. DEPC- and DMS-treated RNAs are preferentially cleaved by aniline on the 5′-side of A and G, respectively. Aniline induces a number of modification-independent cleavages, which can be seen in the no treatment lanes.

of binding activity took one cycle longer than did the analogous selection using 1 M LiCl buffer, suggesting that, as expected, there are fewer solutions in the pool for the restoration of binding in altered salt conditions than there are for the restoration of binding in the original salt conditions. Sequence analysis of clones after six rounds of selection (Figure 10) supports this notion: out of 12 clones, one sequence was found four times and another was found twice. In the original mutagenized pool selection using LiCl-containing buffer, out of 45 sequences, no two clones were the same. Although it is not possible to make statistically significant arguments regarding base changes from only eight different sequences, there are five completely conserved changes away from the Li^+-dependent cyanocobalamin aptamer sequence. These aptamers share a sequence at their 5′-ends that is not found in the Li^+-dependent aptamer (i.e., 5′-CACGGCTAGCGGA-3′). The conservation in the 3′-end of the sequences observed in the Li^+-dependent aptamers, however, is not present in the Na^+/Mg^{2+}-dependent clones.

Aptamer NaCl.12 was chosen for further study because of its good binding characteristics. It was found that this clone requires Mg^{2+} for binding and does not bind to cyanocobalamin

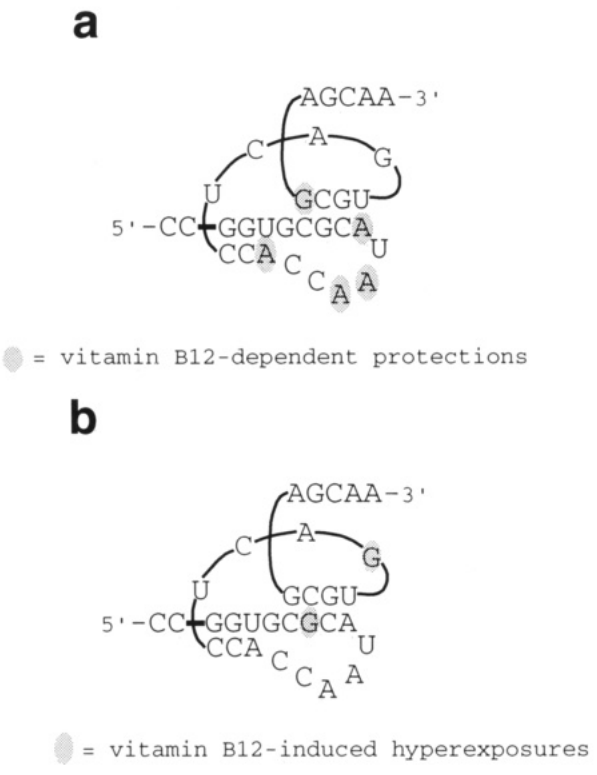

FIGURE 8: Summary of DMS and DEPC modification protection experiments. (a) Bases that are protected from chemical modification in the presence of 100 μM cyanocobalamin. (b) Ligand-induced hyperexposure of bases. Protections and hyperexposures shown were dependent on the presence of 1 M LiCl and were not observed in 1 M NaCl/5 mM $MgCl_2$ buffer.

agarose in 1 M LiCl, even in the presence of 5 mM $MgCl_2$. The aptamer does, however, bind in 1 M KCl with 5 mM $MgCl_2$ and in 0.5 M NaCl with 5 mM $MgCl_2$. The aptamer is eluted from cyanocobalamin–agarose only weakly by cobinamide dicyanide. It is also weakly eluted by adenosylcobalamin. Equilibrium dialysis measurements showed that the aptamer binds cobinamide dicyanide with a K_d of $\sim$1 mM. Because of this weak binding, no equilibrium dialysis competition experiments could be performed. It is clear that the aptamer binds cyanocobalamin, because it binds to cyanocobalamin agarose and is specifically eluted by cyanocobalamin. The aptamer remains bound to cyanocobalamin agarose for at least 20 column vol of washing, so that we can give a rough estimate of the upper limit of the K_d for cyanocobalamin on the column as 40 μM, based on the equation,

$$K_d \approx [(V_c - V_0)/(V_e - V_0)][L]$$

where V_c is the column volume, V_0 is the void volume, V_e is the isocratic elution volume (in the absence of free ligand), and [L] is the ligand concentration on the column (Arnold et al., 1986). Because of the possibility of interactions between the RNA and the column matrix, and the possibility that the actual accessible ligand concentration is lower than 1 mM, the actual K_d for cyanocobalamin in solution could be considerably higher or lower than this estimate.

DISCUSSION

The minimal solution for achieving specific recognition of cyanocobalamin by RNA appears to be significantly more complex than the solutions to other small molecule aptamer selections that have been reported to date (Ellington & Szostak, 1990, 1992; Connel et al., 1993; Sassanfar & Szostak, 1993).

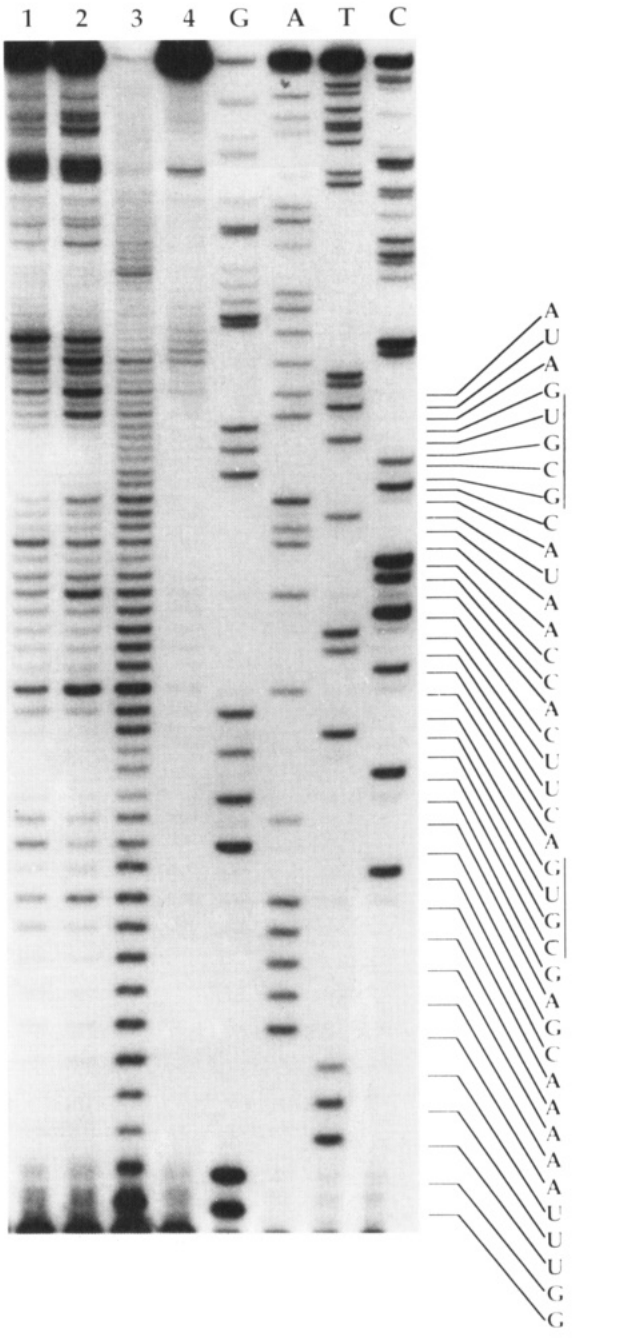

FIGURE 9: Autoradiogram of the gel of reverse transcription chemical modification protection experiments with Dope.40 RNA. The sequence of the RNA from its 3′-end to the 5′-end of the binding site is shown on the right side of the figure: lane 1, DMS-modified RNA (5-min modification); lane 2, DMS-modified RNA (15-min modification); lane 3, 92 °C denaturing control; lane 4, no treatment. G, A, T, and C are dideoxy sequencing lanes. Reverse transcription terminates one base to the 3′-side of the modified nucleotide. Bases that are protected from modification are indicated by a vertical bar to the right of the sequence.

The largest number of completely conserved bases in an aptamer that has been reported previously is eight for the ATP aptamer (Sassanfar & Szostak, 1993). Connel et al. report up to 13 bases that do not vary in the phylogeny from their initial, random pool selection for arginine aptamers (five different clones), but most of these bases are forced to be invariant by the fact that they form Watson–Crick base-pairs with the (constant) 3′-primer binding site; their actual identity is not required. The cyanocobalamin aptamer reported here has 14 absolutely invariant bases (and an additional six bases

FIGURE 10: Results of sequencing of clones from the sixth round of the Na^+/Mg^{2+}-dependent cyanocobalamin-binding selection. The wild-type (WT) sequence is that of the B12.9–*Hae*III fragment, on which the mutagenized pool was based. Changes from this WT sequence are shaded. Clones NaCl.1, NaCl.6, NaCl.7, and NaCl.14 are all the same sequence, as are clones NaCl.2 and NaCl.11. Several clones were tested for cyanocobalamin binding. The results of these experiments are shown to the right of the sequences.

that vary only once in the 45 sequences from the mutagenized pool selection). Furthermore, in previously reported selections, functional aptamers appeared at a frequency of about 1 in 10^9–10^{11} sequences (Ellington & Szostak, 1990, 1992; Connel et al., 1993; Sassanfar & Szostak, 1993). Cyanocobalamin aptamers are much rarer: only ~1 in 10^{14} sequences is able to bind cyanocobalamin specifically and tightly.

The minimal aptamer binds cyanocobalamin surprisingly tightly. The dissociation constant of approximately 90 nM is the lowest reported to date for any aptamer specific for a small molecule ligand. This dissociation constant compares favorably with those of many protein–ligand complexes. We have not yet been able to determine which groups on the ligand are in contact with the RNA. The fact that the aptamer binds more tightly to cyanocobalamin than to cobinamide dicyanide is good evidence that contacts exist between the aptamer and the dimethylbenzimidazole ribonucleotide portion of the ligand, as well as between the aptamer and the cobinamide moiety. It is also possible that a group in the RNA, for instance the N7 of an adenosine or a guanosine, could displace the dimethylbenzimidazole nucleotide as the lower axial ligand for the Co(III). The fact that the aptamer does not bind to adenosylcobalamin suggests that the aptamer may need to have the top face of the ligand accessible in order to bind.

On the basis of the covariational data from the mutagenized pool selection, a model of the secondary structure of the aptamer has been developed. The data are consistent with a pseudoknot structure in which the 5′-loop (L1) is shortened to zero and the joining region (L2) is lengthened (see Figure 6b; Abrahams et al., 1990; ten Dam et al., 1992). One of the covarying positions (C15,G22) imposes a tertiary constraint on the molecule, which suggests that it must be folded in half, with the helices at near-right angles to each other. The possibility of base-pairing between C7–A10 and U23–G26 suggests that these regions may form a helix, but because these positions are completely invariant there is no covariational data to support this hypothesis. The complete conservation of all of the bases in the putative second helix demonstrates that these bases are either required for binding or are involved in structurally important tertiary interactions (because of the relatively low level of doping, base triple covariations would be too rare to have been seen in less than ~300 sequences). The Watson–Crick faces of U23–C25 and C7 and G8 are all protected from DMS modification, which could be due to the formation of a duplex between these two regions.

Another interesting feature of the original aptamer is the Li^+ dependence of its structure and B12 binding. The chemical modification protection experiments show that the structure of the aptamer is different in 1 M NaCl/5 mM $MgCl_2$ *vs* 1 M LiCl. Li^+ has a very high affinity for phosphate groups, and thus it is likely that Li^+ allows the phosphate backbone to adopt unusual conformations by strongly shielding the negative charges and allowing for close phosphate–phosphate contacts and, hence, more compact structures than would otherwise be possible (Liebmann et al., 1982; Cotton & Wilkinson, 1988).[2] We suggest that the use of high LiCl concentrations will allow for the formation of a large variety of novel nucleic acid structures and, thus, may be useful in the search for nucleic acid-based aptamers, catalysts, and other reagents.

It was possible to change the salt requirements of the original aptamer from Li^+ to Mg^{2+} plus Na^+ or K^+ using *in vitro* evolution. The changes in the sequence that correspond to the change from Li^+ dependence to Na^+/K^+ plus Mg^{2+} dependence are intriguing. How the base changes have allowed the RNA to adopt a cyanocobalamin-binding structure in the absence of Li^+ is not yet clear, but it seems likely that the changes involve movement of the phosphate backbone to allow Mg^{2+} to bridge close phosphate–phosphate contacts. Structural studies of the two different classes of aptamers should prove interesting. Lehman and Joyce (1993) previously showed that it was possible to change the divalent cation requirement of the *Tetrahymena* group I intron from Mg^{2+} to Ca^{2+} using *in vitro* evolution. Both of these studies demonstrate the utility of *in vitro* evolution for changing the properties of functional RNAs.

The NaCl.12 aptamer interacts with cyanocobalamin in a different manner than the Li^+-dependent aptamer. Equilibrium dialysis and cyanocobalamin–agarose elution experiments indicate that NaCl.12 RNA binds only weakly to the cobinamide portion of the molecule. It also binds weakly to adenosylcobalamin, in contrast to the Li^+-dependent binder, which does not detectably bind adenosylcobalamin. It is possible that the NaCl selection isolated molecules whose abilities to bind the dimethylbenzimidazole ribonucleotide portion of the molecule, with which the original Li^+-dependent aptamer has been shown to interact, have been optimized. Apparently, solutions that could bind cobinamide in the absence of Li^+ were too distant in sequence space to have been in the mutagenized pool. This suggests that the reason the 3′-end of the Li^+ aptamer is not conserved in the Na^+/Mg^{2+} selection may be that the 3′-region is involved in interacting with cobinamide. The reason that adenosylcobalamin only weakly elutes the NaCl.12 RNA is not clear. It is possible

[2] This idea was first suggested to us by Brian Seed.

that the axial adenosyl group still has some unfavorable steric interactions with the RNA. In any event, this aptamer should be a good starting point for the selection of adenosylcobalamin aptamers.

In conclusion, we have shown that RNA is capable of binding cyanocobalamin tightly and in a highly specific manner. This finding demonstrates that, in as much as binding is the better half of catalysis, cobalamin-dependent ribozymes may have existed in the past. The aptamers we have selected may also serve as a starting point for the *in vitro* evolution of cobalamin-dependent ribozymes.

A number of aptamers isolated in this work have not yet been characterized. It will be interesting to see how their structures and functions compare with those molecules that are already under study. We are currently conducting further structural studies on the aptamers reported in this article. In so doing, we hope to learn more, at an atomic level, about how RNA is able to mediate specific and tight recognition of small molecules. Such data should aid in the design of improved selections for aptamers and RNA-based catalysts.

ACKNOWLEDGMENT

David Bartel synthesized the original random sequence RNA pool. The authors thank David Bartel, Charles Lauhon, and Charles Wilson for discussions and advice. We also thank the aforementioned people and Alicia Hager for comments on this manuscript.

REFERENCES

Abrahams, J. P., van den Berg, M., van Batenberg, E., & Pleij, C. (1990) *Nucleic Acids Res. 18*, 3035–3044.

Allen, R. H., & Majerus, P. W. (1972) *J. Biol. Chem. 247*, 7695–7701.

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

Bartel, D. P., & Szostak, J. W. (1993) *Science 261*, 1411–1418.

Bartel, D., Zapp, M. L., Green, M. R., & Szostak, J. W. (1991) *Cell 67*, 529–536.

Beaudry, A. A., & Joyce, G. F. (1992) *Science 257*, 635–641.

Benner, S. A., Ellington, A. D., & Traver, A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7054–7058.

Bock, L. C., Griffin, L. C., Latham, J. A., Vermaas, E. H., & Toole, J. J. (1992) *Nature 355*, 564–566.

Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry Part III: The Behavior of Biological Macromolecules*, pp 1303–1308, W. H. Freeman and Co., New York.

Connel, G. J., Illangeskare, M., & Yarus, M. (1993) *Biochemistry 32*, 5497–5502.

Cotton, F. A., & Wilkinson, G. (1988) *Advanced Inorganic Chemistry*, p 124, John Wiley and Sons, New York.

Ellington, A. D., & Szostak, J. W. (1990) *Nature 346*, 818–822.

Ellington, A. D., & Szostak, J. W. (1992) *Nature 355*, 850–852.

Eschenmoser, A. (1988) *Angew. Chem., Int. Ed. Engl. 27*, 5–39.

Famulok, M., & Szostak, J. W. (1992) *J. Am. Chem. Soc. 114*, 3990–3991.

Georgopapdakou, N. H., & Scott, A. I. (1977) *J. Theor. Biol. 69*, 381–384.

Green, R., & Szostak, J. W. (1992) *Science 258*, 1910–1915.

Guerrier-Takado, C., Gardiner, K., Marsh, T., Pace, N., & Altman, S. (1983) *Cell 35*, 849–857.

Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., & Cech, T. R. (1982) *Cell 31*, 147–157.

Lehman, N., & Joyce, G. F. (1993) *Nature 1361*, 182–185.

Liebmann, P., Loew, G., McLean, A. D., & Pack, G. R. (1982) *J. Am. Chem. Soc. 104*, 691–697.

Milligan, J. F., & Uhlenbeck, O. C. (1989) *Methods Enzymol. 180*, 51–64.

Moazed, D., Stern, S., & Noller, H. F. (1986) *J. Mol. Biol. 187*, 399–416.

Pan, T., & Uhlenbeck, O. C. (1992) *Biochemistry 16*, 3887–3895.

Peattie, D. A. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1760–1764.

Peattie, D. A., & Gilbert, W. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 4679–4682.

Reichard, P. (1993) *Science 260*, 1773–1777.

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

Sassanfar, M. & Szostak, J. W. (1993) *Nature 364*, 550–552.

Stubbe, J. (1990) *J. Biol. Chem. 265*, 5329–5332.

Szostak, J. W. (1993) *Trends Biochem. Sci. 17*, 89–93.

ten Dam, E., Pleij, K., & Draper, D. (1992) *Biochemistry 311*, 1665–11676.

Tsai, D. E., Kenan, D. J., & Keene, J. D. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8864–8868.

Tuerk, C., MacDougal, S., & Gold, L. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6988–6992.

Walsh, C. T. (1979) *Enzymatic Reaction Mechanisms*, pp 640–665 and 849–851, W. H. Freeman and Co., New York.