Isolation of Hammerhead Ribozymes with Altered Core Sequences by *in Vitro* Selection†

Narendra K. Vaish, Paul A. Heaton, and Fritz Eckstein*

Max-Planck-Institut für experimentelle Medizin, Hermann-Rein-Strasse 3, D-37075 Göttingen, Germany

Received December 23, 1996; Revised Manuscript Received March 17, 1997⊗

ABSTRACT: The hammerhead ribozyme has an invariant nucleotide sequence in the core region. In order to search for alternative sequences which can support the cleavage after the triplet GUC, the core region of 10 nucleotides was randomized and subjected to *in vitro* selection by repeated cycles of transcription, reverse transcription, and PCR. Active sequences were isolated after each transcription by denaturing PAGE, and after nine cycles of selection, two sequences dominated the pool. Both sequences conformed broadly to the consensus core region except that in one sequence a single A^9U mutation was observed while in the other two mutations at A^9U and U^7A were seen. The catalytic efficiencies of these ribozymes were 6.4 and 14.1 μM^{-1} min^{-1}, respectively, as compared to 163 μM^{-1} min^{-1} for the consensus sequence. Interestingly, the consensus was not found in any of the selected sequences. This discrimination against the consensus sequence was attributed to the specificity of the enzymes used in the selection procedure.

Hammerhead ribozymes are small self-cleaving RNAs that are found in certain virus and satellite RNAs that replicate via a rolling-circle mechanism [for a review, see Symons (1992)]. The ribozyme's two-dimensional structure, depicted in Figure 1, was determined by sequence homology and consists of three base paired helices linked by a central core of conserved nucleotides. Further information on the sequence requirements of the hammerhead was determined by systematic mutation of the nucleotides in the core, which indicated that the closing base pair of stem II must be $R^{10.1}$—$Y^{11.1}$ (where R and Y indicate purine and pyrimidine nucleotides, respectively) (Ruffner et al., 1990; Tuschl & Eckstein, 1993); variations are tolerated at position 7 in the core (Ruffner et al., 1990); and finally, the cleavage triplet, after which the ribozyme cleaves, has the general sequence NUH (where N is any nucleotide and H is either A, C, or U) (Ruffner et al., 1990; Shimayama et al., 1995; Zoumadakis & Tabler, 1995). Recently, the three-dimensional X-ray crystal structure has been determined (Pley et al., 1994; Scott et al., 1995), which has revealed a wealth of information on the secondary interactions of the nucleotides in the catalytic core. On the basis of the X-ray structures, the catalytic core has been divided into two regions, which form two distinct structural motifs: domain I, comprising nucleotides C^3−A^6, contains a "uridine-turn"; and domain II, comprising nucleotides U^7−A^9 and G^{12}−A^{14}, contains a GA tandem mismatch. Since the uridine-turn motif has the general sequence requirement UNR and other sequences have been found to adopt a structure similar to this motif (Jucker & Pardi, 1995), then it is not unreasonable to ask whether other sequences are possible in the hammerhead catalytic core. These alternative sequences could not, of course, be detected by single point mutations since two or more

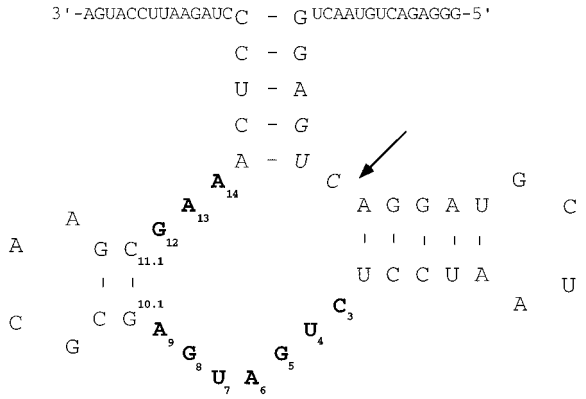

FIGURE 1: Schematic representation of the hammerhead ribozyme. Nucleotides in boldface indicate the conserved catalytic core nucleotides that were randomized; italics indicate the cleavage triplet, and the arrow indicates the cleavage site. Ribozyme numbering is in accordance with Hertel et al. (1992).

complementary mutations may be required to enable an active catalytic core to form. Thus, an alternative method, such as *in vitro* selection, is required to search all possible core sequences.

In vitro selection is a technique where RNA or DNA sequences which exhibit certain properties are isolated from a large number of random sequences by repeated cycles of selection and amplification. The technique, which was first reported in 1990 (Ellington & Szostak, 1990; Tuerk & Gold, 1990), was initially used to select RNA motifs which were able to bind other molecules, such as proteins and small organic molecules [for reviews, see Abelson (1996); Famulok and Szostak (1993)], but has now developed sufficiently to select RNA and DNA sequences which have novel catalytic properties and to improve the catalytic activity of known ribozymes [for reviews, see Abelson (1996), Chapman and Szostak (1994), Famulok and Szostak (1993), and Kumar and Ellington (1995)]. The procedure described here to

† This work was supported by the Deutsche Forschungsgemeinschaft and by a fellowship from the Alexander von Humbolt-Stiftung to N.K.V.

* Corresponding author. Telephone: +49 551 3899 274. Fax: +49 551 3899 388. E-mail: eckstein@mail.mpiem.gwdg.de.

⊗ Abstract published in *Advance ACS Abstracts,* May 1, 1997.

[1] Abbreviations: PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; RT, reverse transcription.

S0006-2960(96)03134-0 CCC: $14.00

select hammerhead ribozymes uses polyacrylamide gel electrophoresis (PAGE)[1] to isolate the cleavage products from *cis* cleaving ribozymes transcribed from a dsDNA pool. With each successive selection cycle, the time allowed for cleavage was reduced to progressively enrich the dsDNA pool with the sequences of faster cleaving ribozymes. Similar procedures have been used previously to select other ribozymes (Nakamaye & Eckstein, 1994; Pan & Uhlenbeck, 1992; Thomson et al., 1996; Williams et al., 1995). The *cis* cleaving ribozymes used in this study incorporate two randomized nucleotide regions in place of the catalytic core nucleotides, C^3-A^9 and $G^{12}-A^{14}$. This allows all possible hammerhead core sequences to be explored.

MATERIALS AND METHODS

Oligodeoxyribonucleotides and oligoribonucleotides were synthesised and purified as previously described (Nakamaye & Eckstein, 1994; Tuschl & Eckstein, 1993). The sequences of the primers and templates used are as follows: primer 1, 5′-CTGATGC**AAGCTT**AATACGACTCACTATA-GGGA*GTC*AGGATGCTA-3′; restoration primer, 5′-CT-TAATACGACTCACTATAGGGAGACTGTAACT-GGA*GTC*AGGATGCTA-3′; primer 2, 5′-TCATG**GAAT-TCTAGA**GGAGTNNNGCTTGCGCNNNN-NNNAGGATTAGCATCCT-3′; RT-primer, 5′-TCATG-**GAATTCTAGA**GGAGT-3′; and PCR-primer, 5′-CTGAT-GC**AAGCTT**AATACGACTCACTATAGG-3′, where the restriction sites for cloning, *Hin*dIII (5′-end) and *Eco*RI and *Xba*I (3′-end), are shown in boldface; the T7 promoter is underlined; the ribozyme cleavage site is in italics; and randomized bases are denoted by N. Ribonucleoside triphosphates, 2′-deoxyribonucleoside triphosphates, and GTPγS were purchased from Boehringer Mannheim. [α-^{32}P]ATP, [α-^{32}P]CTP, and [α-^{32}P]UTP (specific activity $\approx$3000 Ci/mmol), [γ-^{32}P]ATP (specific activity $\approx$5000 Ci/mmol), and [α-^{35}S]dATP (specific activity $\approx$1000 Ci/mmol) were obtained from Amersham Buchler GmbH. Sequenase 2.0, MMLV reverse transcriptase, Taq DNA polymerase with accompanying 10$\times$ buffer, T4 RNA ligase, T4 DNA ligase, and T4 polynucleotide kinase were also from Amersham Buchler GmbH. *Eco*RI, *Hin*dIII, and *Xba*I restriction endonucleases were purchased from New England Biolobs. T7 RNA polymerase was purified from the overproducer *Escherichia coli* BL21/pAR1219 kindly provided by F. W. Studier, Brookhaven National Laboratories, Upton, NY. Human placental ribonuclease inhibitor (RNaseguard) was obtained from Pharmacia. Microcon-10 microconcentrators were purchased from Amicon. Qiaquick spin PCR purification kit and plasmid DNA purification columns (P-20) were obtained from Qiagen. Agarose gels were performed using Metaphor intermediate melting agarose supplied by Biozym, Oldendorf, Germany. JETsorb gel extraction kit was supplied by Genomed GmbH, Bad Oeynhausen, Germany. Sequencing was performed using a Sequenase Quick-Denature plasmid sequencing kit supplied by Amersham Buchler GmbH. Polyacrylamide gels were analyzed using a Fuji BAS2000 Bio-Imaging system. LIXA plates and 2$\times$ YT medium were prepared as previously described (Sambrook et al., 1989).

An overview of the selection procedure is depicted in Figure 2.

Step 1: Preparation of Pool 0 dsDNA Template. Primers 1 and 2 (final concentration 2 μM each) were annealed by heating to 90 °C for 5 min and cooled to room temperature. The following reagents were added to the final concentrations indicated: Tris-HCl (pH 8.0), 65 mM; NaCl, 50 mM; $MgCl_2$, 5 mM; DTT, 5 mM; dNTPs, 375 μM each; and Sequenase 2.0, 0.13 unit/μL. The mixture, typically 500 μL, was incubated at 37 °C for 1 h, before isolating the dsDNA product using a Microcon-10 microconcentrator, used according to the manufacturer's instructions.

Step 2: T7 RNA Transcription and Intramolecular Ribozyme Cleavage. Typically, T7 RNA polymerase transcriptions contained the following reagents at the final concentrations indicated (Milligan et al., 1987): dsDNA template, 1 μM; Tris-HCl (pH 8.0), 40 mM; $MgCl_2$, 20 mM; spermidine, 1 mM; Triton X-100, 0.01%; DTT, 5 mM; NTPs, 2 mM each (where GTP was replaced by GTPγS); T7 RNA polymerase, 50 units/μL; and either [α-^{32}P]ATP, [α-^{32}P]CTP, or [α-^{32}P]UTP. For transcriptions from pool 0, the reaction volume was large (4 $\times$ 250 μL); in later reactions, this was reduced to 150 μL. The transcription reactions were incubated at 37 °C for decreasing periods of time from 12 h (first cycle) to 30 min (ninth cycle), after which the reaction mixtures were then applied to a 12% mercury−PAGE gel (Igloi, 1988). Since the transcription conditions are well suited for the cleavage reaction, active ribozyme sequences underwent cleavage during transcription. The band corresponding to the ribozyme cleavage product was located using a length marker and excised from the gel. The RNA was eluted from the gel piece into NaOAc (1 M, pH 5) and precipitated by addition of ethanol. The cleavage product was collected and dissolved in water (50 μL).

Step 3: Reverse Transcription of Cleavage Product. The RNA cleavage product, in 20 μL, was added to the RT-primer (4 μM final concentration) and heated at 90 °C for 2 min and then cooled to room temperature. The following reagents were added to the concentrations indicated: Tris-HCl (pH 8.3), 20 mM; KCl, 100 mM; $MgCl_2$, 3 mM; dNTPs, 0.5 mM each; RNase inhibitor, 1 unit/μL; and MMLV reverse transcriptase, 2 units/μL, in a total volume of 100 μL. These conditions are similar to those used for PCR. The reaction mixture was incubated at 42 °C for 1 h and then heated at 72 °C for 10 min to inactivate the reverse transcriptase. After cooling to room temperature, NaOH (final concentration 0.2 M) was added and the mixture incubated at 37 °C for 1 h, which destroys any RNA cleavage product remaining in the mixture. The cDNA product was precipitated (−80 °C, 2 h) by addition of NH_4OAc and ethanol (Sambrook et al., 1989). The cDNA was collected and taken up in water (25 μL).

Step 4: Preparation of dsDNA for Amplification by PCR. The entire reverse transcription product and the restoration primer (1.2 equiv) were mixed, heated to 90 °C, and allowed to cool to room temperature. Polymerization was performed using Sequenase 2.0, under the conditions described in step 1, usually in a total volume of 100 μL. The extended cDNA product was amplified by PCR without isolation.

Step 5: PCR Amplification of the cDNA Product. The dsDNA from the Sequenase polymerization with the restoration primer was amplified by PCR as follows (final concentrations given): RT- and PCR-primers (0.5 μM) were added to the dsDNA in addition to Tris-HCl (pH 8.3), 20 mM; KCl, 50 mM; $MgCl_2$, 1.5 mM; dNTPs, 200 μM; and Taq DNA polymerase, 0.025 unit/μL, typically in a total volume of 1 mL. This was then aliquoted into 10 portions and each

overlaid with silicone oil (20 μL). The aliquots were heated initially to 95 °C, 30 s, and then the temperature was cycled: 95 °C, 30 s; 55 °C, 2 min; 72 °C, 1 min. The number of PCR cycles was varied from 15 to 20 depending upon the initial concentration of cDNA. The aliquots were recombined, washed with chloroform (2 × 500 μL), and centrifuged, and the aqueous layer was collected. The product was isolated by precipitation with NaOAc and ethanol (−80 °C, 2 h). The dsDNA was purified by an agarose gel (3.5%), where the band corresponding to the full-length product was cut from the gel and the dsDNA eluted from the gel piece with JETsorb, according to the manufacturer's instructions. This purified dsDNA was reamplified by PCR, using the conditions described above, and isolated by precipitation (NaOAc−ethanol), and the product was taken up in water (50 μL), ready for use in further selection cycles or cloning. Cycling was continued until a strong cleavage band was observed after a 30 min incubation of the transcription mixture.

Cloning and Sequencing. Pool 5 and 9 dsDNA products were digested with *Hin*dIII and *Eco*RI as follows: dsDNA (1 μg) and *Eco*RI and *Hin*dIII restriction enzymes (10 units each) were combined in *Eco*RI buffer (1×, supplied by manufacturer) in a total volume of 20 μL and incubated at 37 °C for 20 h. The double-digested products were recovered with a QIAquick-spin PCR purification kit used according to the manufacturer's instructions. pUC19 plasmid DNA (5 μg) was double-digested and isolated in a similar way. Typically, the isolated, double-digested products were taken up in 50 μL. The selected dsDNA was ligated into pUC19 as follows: double-digested pUC19 (5 μL) and selected dsDNA (15 μL) and T4 DNA ligase (1 unit) were combined in ligase buffer (1×, supplied by manufacturer) in a total volume of 25 μL and incubated at 16 °C for 16 h. The crude ligation reaction mixture was then used to transform competent *E. coli* cells (TG1 strain). The transformed cells were smeared onto LIXA plates and grown at 37 °C for approximately 8 h. White colonies were selected and used to inoculate 2× YT medium (4 mL), from which plasmid DNA was isolated using Qiagen plasmid DNA purification columns (P-20), according to manufacturer's instructions. Active ribozyme-producing sequences were established by runoff transcriptions from *Xba*I-linearized plasmids. Sequencing was performed using the reverse sequence-primer provided in the Sequenase sequencing kit.

Cis and Trans Ribozyme Cleavage. Trans cleaving ribozymes, with the general sequence 5′-GGGUCCUN-NNNNNNGCGCAAGCGAAACUCC-3′, were prepared; the nonspecified nucleotides correspond to CUGAUGU in HH-U[9], CUGAAGU in HH-A[7]U[9], UUGAAAA in HH-U[3]A[7]A[8], and CUGAUGA in the hammerhead ribozyme consensus sequence. The substrate strand, 5′-GGGAGUCAGGAU-3′, was 5′-end-labeled using T4 polynucleotide kinase, according to the manufacturer's instructions. Kinetic parameters were determined under multiple turnover conditions as previously described (Tuschl & Eckstein, 1993). Typically, a mixture containing ribozyme, 1−25 nM; substrate, 50−500 nM; Tris-HCl (pH 7.5), 50 mM; and $MgCl_2$, 10 mM, in a total volume of 50 μL was incubated at 25 °C, and aliquots (8 μL) were taken periodically (1−10 min). The uncleaved substrate and the reaction products were separated by 20% PAGE; the gels were imaged, and the extent of cleavage was quantified. Rates of *cis* cleavage were determined by transcription of dsDNA obtained by PCR of the isolated plasmids using the RT- and PCR-primers as previously described. The transcription conditions were as described above, except that concentrations of dsDNA template and T7 RNA polymerase were reduced to final concentrations of 0.25 μM and 10 units/μL, respectively, in a total volume of 50 μL. Aliquots (8 μL) were taken periodically, and the transcription products were separated by PAGE (12%). The concentrations of cleaved and full-length transcript were determined, and the rate of *cis* cleavage was determined as previously described (Long & Uhlenbeck, 1994).

Influence of Enzymes on Selection Results. Two dsDNA templates were prepared with Sequenase 2.0 as previously described using a modified primer 1, where the GTC ribozyme cleavage triplet was replaced by a noncleaving GTG triplet, and a modified primer 2, where the randomized nucleotides were substituted for either the catalytic core of ribozyme HH-U[9] or the consensus sequence. These dsDNA templates were then mixed in equal proportions and noncleaving ribozymes transcribed from the combined templates using T7 RNA polymerase. The transcription products were separated by PAGE (12%), and the full-length product was then isolated and amplified by RT-PCR. The dsDNA products were loaded onto an agarose gel (3.5%), from which the full-length dsDNA as well as a shorter dsDNA product were isolated. These products were reamplified separately and cloned and sequenced as previously described.

RESULTS

Establishing the Selection Procedure. An initial attempt at ribozyme selection was performed using the procedure described by Nakamaye and Eckstein (1994). Briefly, this involved the preparation of potential *cis* cleaving ribozymes, containing a randomized core region, by transcription of an initial random pool of dsDNA (pool 0), incubation at 37 °C, and separation of the components by PAGE. In this case, the longer ribozyme cleavage product contains the selected catalytic core sequences which can be converted into a complete *cis* ribozyme by RT-PCR. After four cycles of selection, the only product observed after transcription was of equivalent length to the expected cleavage product; no full-length transcription product was present. At this point, the pool of dsDNA was cloned and the plasmid DNA isolated. All of the runoff transcripts from the linearized plasmids appeared to be cleaved completely after a 1 h incubation. However, subsequent sequencing of the plasmids revealed that all the clones contained ribozymes with deletions in the randomized core region and stem−loop II; the number of deletions was equivalent to the number of nucleotides expected to be lost during cleavage. This domination of deletion products is consistent with the earlier observations of Nakamaye and Eckstein (1994).

In an attempt to hinder the formation of deletion products, the selection procedure was modified so that the part of the ribozyme lost during cleavage was restored by means of the restoration primer in a separate Sequenase-mediated polymerization step instead of during the first cycle of PCR. In addition, transcriptions were performed in the presence of GTPγS instead of GTP. As a result, all transcription products, where transcription is initiated with GTPγS, contained a phosphorothioate label at the 5′-end. The components of the transcription mixture were then separated

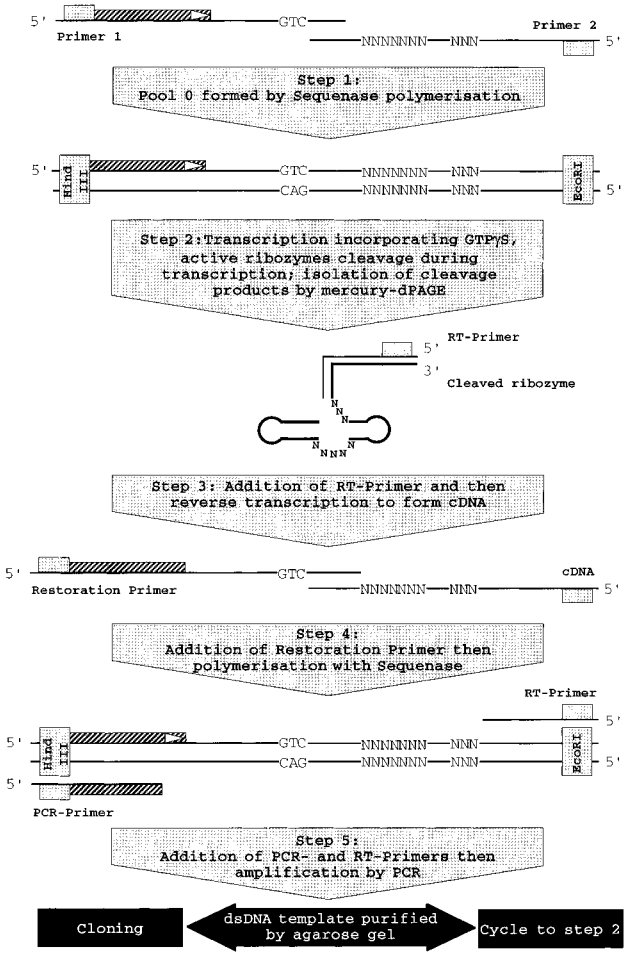

FIGURE 2: Outline of the established selection protocol. This differs from that of Nakamaye and Eckstein (1994) in two ways: first, GTPγS rather than GTP is used during transcription; and second, the T7 promoter is restored in a separate step rather than during reverse transcription.

by mercury−PAGE (Igloi, 1988). Again, the bands corresponding to cleavage products were excised from the gel and amplified ready for further cycles of selection. Analysis of the dsDNA PCR products indicated that the formation of deletion products had been hindered and that pool 3 contained few deletion products. Thus, the introduction of GTPγS and mercury−PAGE step had an effect on reducing the accumulation of deletion products, but this was not totally effective.

To reduce the accumulation of deletion products still further, the dsDNA PCR product was purified by an agarose gel (3.5%) after each cycle of selection. The purification step effectively removed deletion products from each pool of dsDNA. The transcription incubation time was progressively reduced from 6 h at the fifth cycle to 30 min, at the ninth cycle at which point cycling was stopped. The complete selection procedure is depicted in Figure 2.

Cloning and Sequencing of Selected Pools. Pool 5 and pool 9, the cDNA from the fifth (6 h) and ninth (30 min) transcriptions, respectively, were ligated into pUC 19 and cloned using TG 1 cells and the plasmids isolated. The runoff transcripts, from the linearized plasmids, were then used to select clones for sequencing. A selection of clones was chosen, whose transcripts showed self-cleavage activity after a 1 h transcription, and these clones were subsequently sequenced. All of the clones selected contained ribozyme

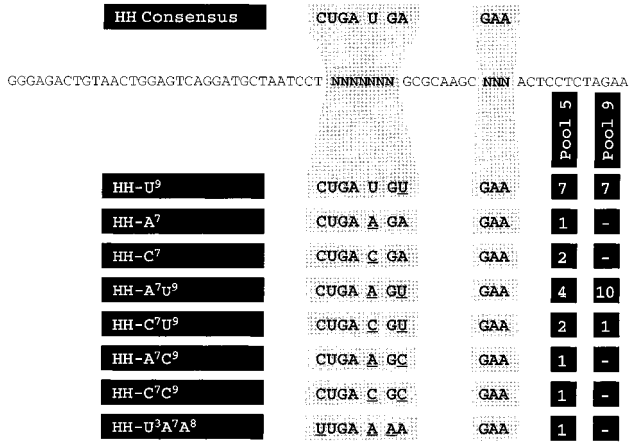

FIGURE 3: Results of sequencing selected clones from pools 5 and 9. The sequence of the ribozyme transcript is depicted, where N denotes the randomized positions. For each unique sequence determined, only nucleotides in the randomized region are shown. The underlined nucleotides indicate positions of variation from the consensus sequence indicated but which was not selected. The number of times a given sequence was found in either pool 5 or pool 9 is indicated.

sequences which conformed broadly to the hammerhead ribozyme consensus sequence (Symons, 1992) but contained a varying number of mutations (Figure 3). All of these mutations were found in the lower single-stranded region connecting stems I and II; the GAA region connecting stems II and III was conserved in all clones sequenced. Surprisingly, none of the clones sequenced were found to contain the unmutated consensus sequence.

As would be expected, pool 5 contained a wider variety of mutated ribozymes than the later pool 9 (Figure 3). The sequences found can be divided into three groups: those containing a single mutation at positions 7 or 9; those containing a double mutation at positions 7 and 9; and a single sequence which contained a triple mutation at positions 3, 7, and 8. The only hammerhead sequence with the single mutation at position 9 was HH-U^9, and this was the most common sequence found (7 times from 19 clones sequenced). The remaining single mutations were at position 7, in HH-A^7 (one clone) and HH-C^7 (two clones), which is a position where variations are tolerated (Ruffner et al., 1990). From the double-mutant hammerheads, HH-U^9A^7 was the most common (four clones). Double mutants with an A to C mutation at position 9 did not appear to be very active from the runoff transcripts. It was surprising to find the triple-mutant HH-$A^8A^7U^3$ which contains mutations at conserved positions 3 and 8 was still sufficiently active to proceed though to pool 5. Pool 9 contained only three unique sequences, which comprised the two most prevalent sequences of pool 5, HH-U^9 and HH-U^9A^7, with roughly equal frequency, and an additional double mutant, HH-U^9C^7 (Figure 3).

Kinetic Characterization of Ribozymes. The two predominant mutant ribozymes from pools 5 and 9 (HH-U^9 and HH-U^9A^7) and the triple mutant from pool 5 (HH-$A^8A^7U^3$) were synthesized as *trans* cleaving ribozymes, along with the consensus sequence, in order to determine their kinetic rate constants (Table 1). The single-mutant ribozyme HH-U^9 and the double-mutant HH-U^9A^7 were found to have similar catalytic efficiencies, $k_{cat} = 0.39$ and 0.31 min^{-1} and $K_M = 61$ and 22 nM, respectively, under typical assay

Table 1: Kinetic Parameters Determined for the Consensus and Selected Ribozymes

	k_{cat} (min^{-1})	K_m (nM)	k_{cat}/K_m (μM^{-1} min^{-1})
HH consensus[a]	6.70	41	163.4
HH-U^9 [a]	0.39	61	6.4
HH-A^7U^9 [a]	0.31	22	14.1
HH-$U^3A^7A^8$ [b]	0.10	34	2.9

[a] 10 mM $MgCl_2$. [b] 50 mM $MgCl_2$.

conditions (Tris-HCl, 50 mM, pH 7.5; $MgCl_2$, 10 mM). The triple-mutant HH-$A^8A^7U^3$ was inactive under these conditions, but on increasing the Mg^{2+} concentration (50 mM), its activity improved with $k_{cat} = 0.10$ min^{-1} and $K_M = 34$ nM. Kinetic constants $k_{cat} = 6.7$ min^{-1} and $K_M = 41$ nM were determined for the hammerhead consensus sequence under typical assay conditions. *Cis* cleavage of ribozyme transcripts was performed as previously described (Long & Uhlenbeck, 1994), and this clearly indicated that the consensus sequence ribozyme cleaved faster than HH-U^9 (data not shown).

Influence of Enzyme Specificity on Selection Results. To investigate the effect of enzyme specificities on the selection process, two noncleaving ribozymes, containing a GUG cleavage triplet with a catalytic core which corresponded to either HH-U^9 or the consensus sequence, were prepared by transcription. The full-length transcripts were then amplified by RT-PCR. Analysis of the products by agarose gel indicated that a shorter product had been formed in addition to the expected full-length dsDNA. Each of these components was isolated, reamplified, cloned, and sequenced. Almost all of the 25 clones analyzed from the full-length dsDNA had a sequence which corresponded to the ribozyme HH-U^9; only one corresponded to the consensus sequence, which initially had comprised half of the sequences in the pool. Sequencing of the shorter dsDNA product indicated that it was the result of deletion of stem−loop II and part of the seven nucleotide loop in the catalytic core of the ribozyme.

DISCUSSION

Investigations to determine alternative sequence requirements for the hammerhead ribozyme using an *in vitro* selection procedure have been reported previously (Long & Uhlenbeck, 1994; Nakamaye & Eckstein, 1994; Thomson et al., 1996). The majority of these selections used *cis* cleaving ribozymes transcribed from randomized dsDNA templates, where cleaved ribozymes were isolated by PAGE and their cDNA was amplified by PCR. In addition, a novel selection scheme has been reported using an *trans* acting ribozyme cleaving a bound substrate (Ishizaka et al., 1995). So far, studies with *cis* cleaving ribozymes have been of limited value, since only a small number of randomizations were involved, which were placed either in the core region or in stem−loop II, and a small number of selection cycles used.

The first report of an *in vitro* selection was based on an AUA-cleaving hammerhead ribozyme derived from barley yellow dwarf virus (BYDV) (Nakamaye & Eckstein, 1994). This ribozyme differs from the consensus sequence in two respects: the G-C base pair in stem II adjacent to the core is replaced by $A^{10.1}C^{11.1}$, and U^7 is replaced by C. These three positions were randomized and two cycles of selection made. Sequencing of the selected ribozymes revealed that the preferred nucleotides at the randomized positions were those expected from the consensus sequence and enhanced the cleavage activity of the BYDV ribozyme at the AUA site. In a subsequent study by Long and Uhlenbeck (1994), active ribozymes were selected from a pool of hammerhead ribozymes, where stem−loop II had been replaced with six random nucleotides but which still retained the consensus sequence in the core. After three cycles of selection, the most active ribozyme comprised a G-C base pair adjacent to the core and a tetranucleotide loop, consistent with the consensus sequence. In a more recent study (Thomson et al., 1996), the three nucleotides forming the single-stranded region between stems II and III were replaced by four random nucleotides. After two cycles of selection, an HGAA sequence (H is either A, U, or C) was selected at the randomized positions. It was argued that this sequence predominated as the sequence could allow a typical hammerhead core to form by incorporating the extra nucleotide as a bulge in stem II.

The only other attempt to select a hammerhead ribozyme from a random pool was the selection of a *trans* cleaving ribozyme based on a "minizyme" a hammerhead ribozyme lacking stem−loop II (Ishizaka et al., 1995). Here, a randomized stretch of 14 nucleotides replaced the core region and stem−loop II and was flanked by the 2 constant regions which formed stems I and III. After eight rounds of selection to cleave a support-bound substrate, ribozymes were sequenced which had a core region consistent with the consensus sequence and a stem II consisting of a single G-C base pair, terminated by a three nucleotide loop. Surprisingly, the selected ribozymes contained a total of 15 nucleotides between the 2 flanking constant regions, 1 more than the initial random pool. The authors attributed the extra nucleotide to an error in transcription of the dsDNA to form the initial random pool; no supporting evidence was given for this.

The work presented here, describes the first successful selection of a *cis* cleaving hammerhead ribozyme from a random pool, where the entire core region has been randomized, and offers an improved selection procedure to that described by Nakamaye and Eckstein (1994). Initially our attempts at selection followed this previously reported procedure, and unsurprisingly we also observed the domination of shorter transcription products during repeated cycles of selection. Sequencing the dsDNA products from the selection indicated that deletions had occurred in the core and stem−loop II region so that the transcription products from this dsDNA were of comparable length to the expected cleavage product. It appeared that these products were an artifact of the RT-PCR step; such artifacts have been reported previously in both RT (Buell et al., 1978) and PCR (Cariello et al., 1991; Green & Sargan, 1991). These reports describe the removal of DNA and RNA regions which are able to form self-structures, such as hairpins. In addition, the tendency to delete sequences appears to be a characteristic of the enzymes used. For example, when a human gene fragment was amplified by Taq DNA polymerase, deletion artifacts were observed; sometimes they were the sole product of amplification (Cariello et al., 1991). In contrast, when Sequenase 2.0 was used, no deletion products could be detected.

Three steps were introduced in order to prevent the domination of deletion products. One potential source was thought to be a consequence of preferential amplification of deleted cDNA during PCR. Deletion products retain the nucleotides at the 5′-end, which have been removed in the cleavage products. These extra nucleotides would give the cDNA from the deletion products an increased overlap with the restoration primer in the first cycle of PCR than the cDNA from the cleavage products and thus have an advantage in the amplification step. To circumvent this perceived problem, the restoration of the cleaved nucleotides from the ribozyme cleavage product was performed in a separate Sequenase-mediated polymerization prior to PCR amplification. Thus, the 5′- and 3′-ends of the cDNA from both the deletion and the cleavage products were identical in the PCR amplification step, giving the cleavage products an equal chance of amplification.

A second potential source was comigration of deletion products with cleavage products on PAGE. As a result, a method was established to selectively remove the deletion products from the transcription mixture. The method involved performing the transcription in the presence of GTPγS. Under these conditions, all of the transcription products, that is, full-length transcripts, deletion products, and premature termination products, are formed with a phosphorothioate at the 5′-end, since transcription is initiated with GTPγS. In contrast, the cleavage products and other hydrolysis products lack this label. Mercury−PAGE (Igloi, 1988) can now be used to retard the mobility of thio-containing oligonucleotides, while allowing unlabeled oligonucleotides to move through the gel normally, enabling deletion and cleavage products of the same length to be separated. However, all samples of GTPγS contain small amounts of hydrolysis products, GDP and GMP, which are also able to initiate transcription. Thus, deletion products primed with either GMP or GDP can still comigrate with the ribozyme cleavage products, since both lack a phosphorothioate, so the procedure merely limits the amount of deletion product carried through to the following cycle.

In order to completely eliminate the deletion products from the selection, it was necessary, in addition to the above measures, to purify the dsDNA product from the PCR, by excising the desired band from an agarose gel. Thus, not only did the RNA cut from the transcription gel have to be of the correct length, but also the dsDNA had to be of the correct length, which was only possible for the undeleted ribozymes. With these modifications included, the selection protocol was very reliable in selecting for cleaving ribozymes, and this technique could therefore be applied to other areas where similar problems occur.

After nine cycles of selection, two sequences were found to dominate the pool, neither of which corresponded to the consensus sequence. This was surprising, since determining the kinetic parameters for the *trans* versions of the selected and the consensus hammerhead ribozymes clearly shows the consensus hammerhead to be the more efficient cleaver. The relative rates obtained for the two selected ribozymes, which contain mutations at positions 7 and 9, compared with the consensus sequence are consistent with those previously reported (Ruffner et al., 1990). Initially, it was thought that when the ribozymes are *cis* cleavers, the selected ribozymes cleave more favorably to the consensus hammerhead sequence, whose cleavage was hindered in some way. To investigate this possibility, the *cis* cleavage of the various ribozymes was monitored during transcription using a method developed by Long and Uhlenbeck (1994). This clearly indicated that the hammerhead consensus sequence was far more active than the selected mutant, HH-U^9 (data not shown).

All of the ribozymes sequenced from the selection, either from pool 5 or from pool 9, contain mutations which occur at either position 7 and/or position 9 in the core. These nucleotides form part of domain II, previously described in the crystal structures (Pley et al., 1994; Scott et al., 1995), which consists of a tandem GA mismatch and a non-Watson−Crick A-U base pair. From the mutants identified in pool 5, HH-C^7 and HH-A^7 have been shown to have activities comparable to the consensus sequence, 20% and 50%, respectively (Ruffner et al., 1990), and would have been anticipated to survive further selection cycles and dominate in pool 9. Surprisingly, the less active mutant HH-U^9, which retains only 6% activity (Ruffner et al., 1990), and the double mutant HH-U^{9}A^7, which has a comparable activity to HH-U^9, dominate the pool, excluding completely the consensus sequence.

In order to investigate the possibility that some factor in the selection protocol was preventing the domination of the consensus sequence, a dsDNA pool was prepared which comprised the sequences of two *cis* cleaving ribozymes with catalytic cores corresponding to the selected ribozyme HH-U^9 and the consensus sequence. The cleavage triplets of these ribozymes were mutated from GUC to noncleavable GUG so that each ribozyme's ability to cleave was not a factor in the selection. The initial pool contained equal proportions of the dsDNA for these ribozymes, and one cycle of selection was performed using the full-length transcript excised from the gel. The dsDNA obtained was then cloned and sequenced. From the 25 clones sequenced, 24 were ribozymes with catalytic cores from HH-U^9, the ribozyme obtained from the selection, while only 1 conformed to the consensus sequence (Figure 3). This demonstrates quite conclusively that the selection procedure contains additional selection pressures, other than the rate at which the ribozyme cleaves. Since it would be unlikely that there would be a great difference in the ability of the dsDNA of the two ribozymes to be transcribed, the most likely source of these additional selection pressures is the RT-PCR step. Deletion, which has been a recurring problem during the selection process, offers a route by which cleaving ribozymes could be lost from the selection pool, and, if it were sequence-specific, it would impose an additional pressure on the selection process. Thus, it is plausible that the consensus sequence, with a GA tandem mismatch adjacent to stem−loop II, is particularly prone to deletion. Disruption of the mismatch, as in HH-U^9, appears to be sufficient to allow some RT-PCR, but the tendency to delete has not been completely removed. It appears, therefore, that the selection procedure disfavors an A at position 9 favoring a U, which is the most active of the mutations at this position, or to a lesser degree a C. These mutations at position 9 disrupt the GA tandem mismatch in domain II. Of course, the tandem mismatch can be disrupted by varying any of the other three nucleotides forming the motif, but these would tend to form mutants which have a lower activity to HH-U^9 (Ruffner et al., 1990). Thus, the ribozymes predominating during the

selection are the most active ribozymes possible with a disrupted GA tandem mismatch.

The results described here have important implications for selection of both aptamers and catalytic RNAs. We have shown that certain, unpredictable sequences may be excluded during the selection process not only on the basis of the rate of ribozyme cleavage but also on the ability of the RNA and its corresponding cDNA to be efficient substrates for the enzymes involved in the selection procedure. Thus, a selected RNA may not be the most optimal for a particular function but may represent a compromise between RNA function and the substrate preferences of the enzymes used. Within the limitations described above, our selection of a GUC-cleaving hammerhead ribozyme indicates that the consensus sequence, derived from viruses and virusoids, is probably the most optimal sequence. It also confirms the findings of Ruffner and Uhlenbeck (1990), in that mutations at position 7 have little effect on catalytic efficiency and that the most favorable nucleotide at position 9, after adenosine in the consensus sequence, is U. In addition, alternative "uridine-turns", which comprise domain I, such as UNR, GNRA, and UUCG (Jucker & Pardi, 1995) cannot support efficient catalysis, and their incorporation cannot be compensated for by mutation in other parts of the motif.

ACKNOWLEDGMENT

We are indebted to Drs. J. B. Thomson and K. A. Birikh for their continual support and advice throughout the course of this work and U. Kutzke and A. Zeuch for their skillful technical assistance. We also thank Prof. T. Cech for helpful discussions.

REFERENCES

Abelson, J. N., Ed. (1996) *Methods Enzymol. 267*.
Buell, G. N., Wickens, M. P., Payvar, F., & Schimke, R. T. (1978) *J. Biol. Chem. 253*, 2471−2482.
Cariello, N. F., Thilly, W. G., Swenberg, J. A., & Skopek, T. R. (1991) *Gene 99*, 105−108.
Chapman, K. B., & Szostak, J. W. (1994) *Curr. Opin. Struct. Biol. 4*, 618−622.
Ellington, A. D., & Szostak, J. W. (1990) *Nature 346*, 818−822.
Famulok, M., & Szostak, J. W. (1993) *Nucleic Acids Mol. Biol. 10*, 271−284.
Green, I. R., & Sargan, D. R. (1991) *Gene 109*, 203−210.
Hertel, K. J., Pardi, A., Uhlenbeck, O. C., Koizumi, M., Ohtsuka, E., Uesugi, S., Cedergren, R., Eckstein, F., Gerlach, W. L., Hodgson, R., & Symons, R. H. (1992) *Nucleic Acids Res. 20*, 3252.
Igloi, G. (1988) *Biochemistry 27*, 3842−3849.
Ishizaka, M., Ohishima, Y., & Tani, T. (1995) *Biochem. Biophys. Res. Commun. 214*, 403−409.
Jucker, F. M., & Pardi, A. (1995) *RNA 1*, 219−222.
Kumar, P. K. R., & Ellington, A. D. (1995) *FASEB J. 9*, 1183−1195.
Long, D. M., & Uhlenbeck, O. C. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 6977−6981.
Milligan, J. F., Groebe, D. R., Witherell, G. W., & Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 8783−8798.
Nakamaye, K. L., & Eckstein, F. (1994) *Biochemistry 33*, 1271−1277.
Pan, T., & Uhlenbeck, O. C. (1992) *Biochemistry 31*, 3887−3895.
Pley, H. W., Flaherty, K. M., & McKay, D. B. (1994) *Nature 372*, 68−74.
Ruffner, D. E., Stormo, G. D., & Uhlenbeck, O. C. (1990) *Biochemistry 29*, 10695−10702.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Scott, W. G., Finch, J. T., & Klug, A. (1995) *Cell 81*, 991−1002.
Shimayama, T., Nishikawa, S., & Taira, K. (1995) *Biochemistry 34*, 3649−3654.
Symons, R. H. (1992) *Annu. Rev. Biochem. 61*, 641−671.
Thomson, J. B., Sigurdsson, S. T., Zeuch, A., & Eckstein, F. (1996) *Nucleic Acids Res. 24*, 4401−4406.
Tuerk, C., & Gold, L. (1990) *Science 249*, 505−510.
Tuschl, T., & Eckstein, F. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 6991−6994.
Williams, K. P., Ciafré, S., & Tocchini-Valentini, G. P. (1995) *EMBO J. 14*, 4551−4557.
Zoumadakis, M., & Tabler, M. (1995) *Nucleic Acids Res. 23*, 1192−1196.

BI963134R