Enzymatic combinatorial nucleoside deletion scanning mutagenesis of functional RNA

Katarzyna Wawrzyniak-Turek and Claudia Höbartner*

Research Group Nucleic Acid Chemistry, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

Table of Contents

Index of Supporting Figures

Oligonucleotides

All RNA and DNA sequences are given in 5´ to 3´ direction:

a) *RNA libraries*

Red nucleotides correspond to the regions that contain the substitutions with the C3 spacer unit

Hammerhead ribozyme (HHR)
AUCCAGCUGAUGAGUCCCAAAUAGGACGAAACAUGUUUCUGCAUGUCCUGGAU
Hairpin II of U1 snRNA
GAGCCAUUGCACUUCGGCUUGCUAGACUUAC

b) *DNA library*

Green nucleotides correspond to the regions that contain the substitutions with the 2'-OH-C3 spacer unit

9HR17 deoxyribozyme
CCGTCGCCATCTCCAGCTATATGTGCTGGACTGAGAGGGGTAGTTTCGCAGTGAGGTGTAGG
TTCCCGTATTATCC

c) *DNA primers*

HHR GACATGCAGAAACA
U1 snRNA GTAAGTCTAGCAAG
9HR17 GTGGATAATACGGGAAp
For primer extension optimization on RNA GTAAGTCTCAGGG
For primer extension optimization on DNA ATGAGAGGTAAGTCTCATG and
ATGAGAGGTAAGTCTCATGTACT

d) *Substrate and adaptor sequences for 9HR17 deoxyribozyme library*

Adaptor strand (RNA) pppGGAAGAGAUGGCGACGG
Deoxyribozyme substrate GGATAATACG(rC)TTCACTGCG

e) *DNA templates for in vitro transcription*

Hammerhead ribozyme
ATCCAGGACATGCAGAAACATGTTTCGTCCTATTTGGGACTCATCAGCTGGATCCTATAGTG
AGTCGTATTACAG
Hairpin II of U1 snRNA
GTAAGTCTAGCAAGCCGAAGTGCAATGGCTCTATAGTGAGTCGTATTACAG

f) *Sequence for primer extension optimization*

RNA template
GGGAAGGGAAGAAACTGCGGGTTCTCCCGGCTTCCCTGAGACTTAC

95:5 mixture of rG:1 phosphoramidites was used during the solid-phase synthesis at positions depicted in red

DNA template
GTGGXAGCXXXTGCXXTAXCXGXCCCXXAAAGTACATGAGACTT
Phosphoramidite 1 was used during the solid-phase synthesis at positions depicted in red

The unmodified DNA oligonucleotides were purchased from Sigma Aldrich and purified by PAGE (10, 15 or 20 % of polyacrylamide). The RNA and DNA libraries as well as ribonucleotide-containing DNA substrate for 9HR17 deoxyribozyme were prepared by solid-phase synthesis. The unmodified RNA oligonucleotides for hammerhead ribozyme and hairpin II of U1 snRNA were prepared by in vitro transcription using T7 RNA polymerase and synthetic templates.

Solid-phase synthesis of oligonucleotides

Solid-phase syntheses were performed on a Pharmacia Gene Assembler Plus using standard phosphoramidite chemistry (ribonucleotides were 2′-*O*-TOM protected).[5] Reagents and polystyrene custom primer supports (0.7 µmol) were purchased form ChemGenes, ProligoReagents and GE Healthcare. Phosphoramidite solutions (100 mM in CH_3CN), S-Benzylthiotetrazole (BTT) solution and acetonitrile were dried over activated molecular sieves for overnight. Following synthesis conditions were applied: S-Benzylthiotetrazole (250 mM in CH_3CN) was used as activator; coupling time was 4 min; 1:1 (v/v) A/B mixture was applied for one minute for capping (A, Ac_2O/sym-collidine/CH_3CN, 2:3:5; B, 0.5 M 4-(dimethylamino)pyridine in CH_3CN); 10 mM I_2 in CH_3CN/sym-collidine/water (10:1:5) was used for oxidation; detritylation was achieved with 3% dichloroacetic acid in 1,2-dichloroethane. Following the solid-phase synthesis oligonucleotides were cleaved from the solid support, deprotected, analyzed by anion-exchange HPLC and purified by PAGE.

Preparation of the RNA libraries

The libraries for hammerhead ribozyme and hairpin II of U1 snRNA were synthesized using standard RNA phosphoramidites for the primer binding sites and phosphoramidites mixtures of 95:5 (rN:1) for the mutagenized (randomized) regions. Cleavage from the solid support and deprotection were accomplished with 1:1 mixture of $MeNH_2$ in water (40 wt %) and in ethanol (33 wt %) at 37°C for 6 h. The 2′-*O*-TOM groups were removed with 1 M TBAF in THF (37°C, overnight), followed by desalting of the crude product on a Sephadex G10 column (3x5ml HiTrap column, GE Healthcare) and purification of the desired RNA by PAGE (15% polyacrylamide). The library for the hairpin II of U1 snRNA was 5′-^{32}P-trace-labeled by T4 polynucleotide kinase (Fermentas) according to the manufacturer's instruction.

Selection of the active library variants

In order to isolate the active mutants from the unseparated libraries, the RNA and DNA pools were incubated with their cognate substrates. Following the cleavage or ligation reaction catalysed by (deoxy-)ribozymes or the RNA recognition by the U1A protein, the active fractions were separated from inactive variants by gel electrophoresis at denaturing or native conditions. Each separation experiment was repeated at least two times.

Ribozyme

Hammerhead ribozyme library (0.5 nmol) was dissolved in 17 µL of annealing buffer (40 mM Tris.HCl pH 8.0, 150 mM NaCl and 1 mM EDTA), heated up to 95°C for 2 min and cooled down to room temperature for 15 min. Cleavage reaction was initiated by addition of 4 µL 5x reaction buffer (1M NaCl, 0.5 M Tris.HCl pH 8.0) and 2 µL of 100 mM $MgCl_2$. Reaction proceeded at 37°C for 10 min, 30 min or 2h and was quenched by addition of gel loading buffer containing 50 mM EDTA and 80% formamide[1]. The samples were then applied on 15% polyacrylamide gel (polyacrylamide, 1 x TBE, 7 M urea, 0.4 mm, 20x30 cm) and run for 2 h at 35 W. Oligonucleotides were visualized by UV shadowing. The band corresponding to the active fraction was extracted from the gel and analysed by primer extension.

RNA protein-interaction

Mutants of the hairpin II of U1 snRNA that could be recognized by U1A protein were separated by gel mobility shift assays as described previously[2]. Briefly, reaction mixtures containing 5′-^{32}P-labeled RNA library (10 pmol) and protein (0.1-1 equivalent) in 10 mM Tris.HCl pH 7.4, 250 mM NaCl, 1 mM EDTA, 0.5% Triton X-100 and 1mg/ml tRNA in a total volume of 10 µL were equilibrated for 1h. The glycerol to a final concentration of 5% was added and the reactions were separated on 8% polyacrylamide gel in a buffer containing 100 mM Tris-borate (pH 8.3), 1mM EDTA and 0.1% Triton X-100 for 30 min at 200 V in an ice-bath. Bromophenol blue in a glycerol was loaded on a side to monitor the gel progress. Gels were visualized on a Molecular Dynamics Storm PhosphorImager. For each separation condition five identical samples were prepared that were run on one gel and cut together to ensure enough material for analysis by primer extension (50 pmol before separation).

9HR17 Deoxyribozyme

The deoxyribozyme substrate (166 pmol) and the ^{32}P-labeled DNA library (50 pmol) ligated to the adaptor strand[6] were annealed in the presence of 5 mM HEPES pH 7.5, 15 mM NaCl and 0.1 mM EDTA by heating at 95°C for 2 min and cooling down to room temperature for 15 min. The ligation reaction was initiated by addition of 5x reaction buffer (250 mM HEPES pH 7.5, 750 mM NaCl, 10 mM KCl) and 20 mM $MnCl_2$. The reaction was performed at 37°C for 1.5h. The ligation products were separated by denaturing PAGE (10%, 0.7 mm, 35 W, 1.5 h) and the active (ligated) fractions were extracted and analysed by primer extension.

Primer extension

RNA unseparated libraries and RNA active variants were analysed by primer extension reactions performed by reverse transcriptases (SuperScript III from Invitrogen and M-MuLV from NEB were used). DNA primers were 5′end labeled with γ^{32}P-ATP using T4 polynucleotide kinase. Sequencing ladders were obtained by a Sanger's method.

Two µL of labeled DNA primer and 1 µL of each dNTP (10 mM) were added to 8.5 µL of RNA template from separation reaction and allowed to anneal at 65°C for 5 min and at 4°C for 1 min. After addition of 4 µL of 5x first strand buffer (250 mM of Tris.HCl pH 8.3, 375 mM KCl and 15 mM $MgCl_2$), 1 µL DTT (0.1 M) and 100 units of Super Script III reverse transcriptase the reaction mixture was incubated at 50°C for 2 min. For difficult templates the reaction temperature can be increased up to 60°C. Reactions were stopped by addition of 1 µL of 4N NaOH and incubation at

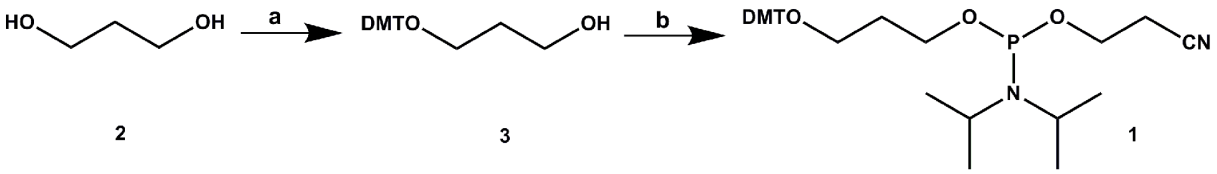
95°C for 5 min. Radiolabeled cDNA strands were recovered by ethanol precipitation and re-suspended in gel loading buffer. Sequencing ladders were produced by addition of 1 µL of ddNTP (10 mM) and 1 µL of corresponding dNTP (1 mM) to the reaction mixture of in vitro transcribed RNA samples (5-10 pmol) prior to annealing. Sequencing reactions were run for 30 min. Primer extension products were separated on 15% polyacrylamide gel at 35 W. The gel was dried at 80°C for 30 min and exposed to a ^{32}P-sensitive storage screen for overnight before scanning by a Storm PhosphorImager.

When M-MuLV reverse transcriptase was applied for the primer extension reaction the final reaction volume prior to annealing was 17.5 µL, 10x M-MuLV reverse transcriptase reaction buffer (50 mM Tris.HCl pH 8.3, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT) was used and the reaction was conducted at 42°C with 100 units of the enzyme.

Data analysis and calculation of interference effects

The primer extension products were visualized using a Storm PhosphorImager and quantified using ImageQuant software (area analysis tool). All bands resulting from stops of the enzymatic reactions were assigned to a corresponding nucleotide positions with the help of the sequencing ladders. In case of RNA samples (i.e. in reverse transcription reactions), the bands in the lane of the RNA libraries are always downshifted by 1 in reference to the bands in the sequencing ladders; (the final nucleotide is incorporated opposite to the last templating nucleotide before the C3 spacer). For DNA samples the bands from the primer extension reactions on libraries correspond directly to the bands in the sequencing lanes; (the last nucleotide is incorporated directly opposite to the C3 spacer, i.e. non-templated). Data sets were normalized for loading variations and primer extension efficiency by dividing all intensities by the intensity of the band where the random stops of the polymerases are expected to be the same in all fractions (e.g. a position at which the deletion of the nucleobase did not have any effect on activity). Finally, the interference effect at every single position was calculated by dividing the intensity of a particular band in an unseparated library by the intensity of the corresponding band in the active fraction. Interference effect values are represented as bar graphs. Positions where the interference effect was higher than 1.5 are sensitive to the replacement with Δ, and therefore likely required for the function. Positions with interference values below 1.5 are expected to accept mutations or deletions.

Chemical synthesis of phosphoramidite 1



Scheme S1. Synthesis of phosphoramidite **1**. Reagents and conditions: a) DMT-Cl, pyridine, RT, 4h, 55%; b) CEP-Cl, Me_2NEt, CH_2Cl_2, RT, 40 min, 75%.

3-(4,4'-Dimethoxytrityloxy)-1-propanol[3] (3). To a solution of propan-1,3-diol (**2**) (7.6 g, 100 mmol) in pyridine (40 ml) 4,4'-dimethoxytrityl chloride (DMT-Cl, 1.7 g, 5 mmol) was added under argon and the mixture was stirred for 4h at room temperature. Ether was then added (50 ml) and the solution was washed with water, dried over anhydrous Na_2SO_4, filtered and concentrated. The crude product was purified by column chromatography (2% NEt_3, hexane:DCM 1:1) to give pale yellow product **3** (1.1 g,

2.9 mmol, 55%). R_f = 0.7 (DCM:EtOH 5:2); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44-7.41 (m, 2H), 7.34-7.26 (m, 7H), 6.85-6.81 (m, 4H), 3.80 – 3.75 (m, 8H), 3.28 (t, 2H), 2.21 (bt, 1H), 1.85 (quintet, 2H); ESI-MS calc. for $C_{24}H_{26}O_4$ $[M+Na]^+$ 401.46, found $[M+Na]^+$ 401.3.

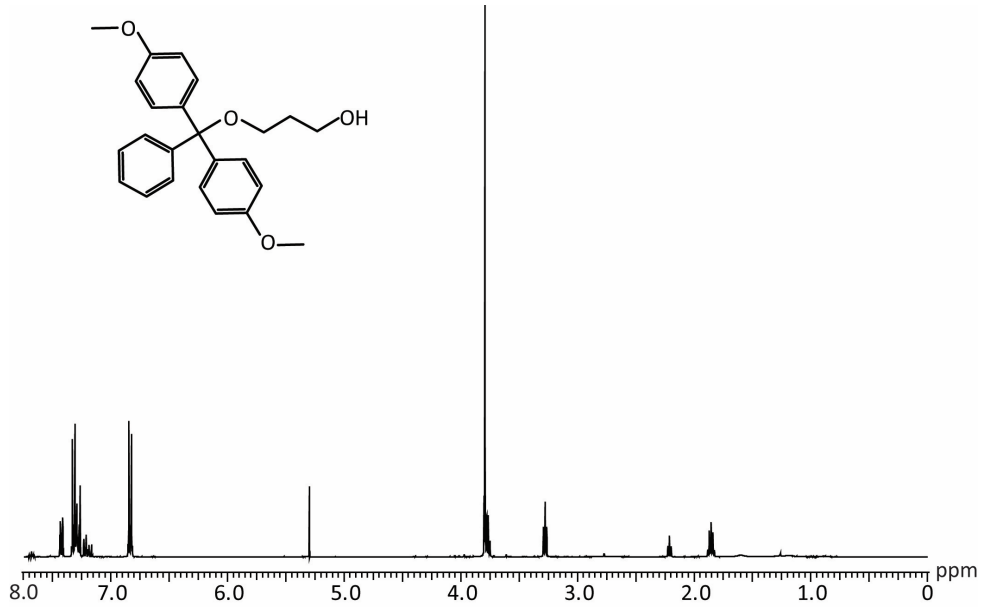


Figure S2. ^{1}H-NMR (400 MHz, $CDCl_3$) of 3-(4,4'-Dimethoxytrityloxy)-1-propanol (compound 3)

3-(4,4'-Dimethoxytrityloxy)propyl-1-[(2-cyanoethyl)-(*N*,*N*-diisopropyl)]-phosphoramidite[4] (**1**). To a mixture of compound **3** (0.66 mmol, 250 mg) and *N*,*N*-dimethylethylamine (0.72 mL, 6.6 mmol) in DCM (8 mL) CEP-Cl (190 mg, 0.8 mmol) was added under argon and the mixture was stirred for 40 min at room temperature. The reaction was quenched with 0.5 mL methanol, diluted with DCM (40 mL) and extracted with saturated $NaHCO_3$ (20 mL) and water (20 mL). The organic fraction was then dried over anhydrous Na_2SO_4, filtered, concentrated and purified by column chromatography (2% NEt_3, Et_2O:hexane 1:1) to give compound **1** as a colorless oil (0.5 mmol, 285 mg, 75%). R_f = 0.95 (DCM:EtOH 5:2); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44-7.41 (m, 2H), 7.33-7.25 (m, 7H), 6.83-6.80 (m, 4H), 3.84-3.67 (m, 10H), 3.60-3.50 (m, 2H), 3.16 (t, 2H), 2.61-2.49 (m, 2H), 1.92 (quintet, 2H), 1.14 (dd, J = 4Hz, J = 20 Hz, 12H); ^{31}P NMR (121 MHz, $CDCl_3$) δ 147; ESI-MS calc. for $C_{33}H_{43}N_2O_5P$ $[M+Na]^+$ 601.68, found $[M+Na]^+$ 601.4.

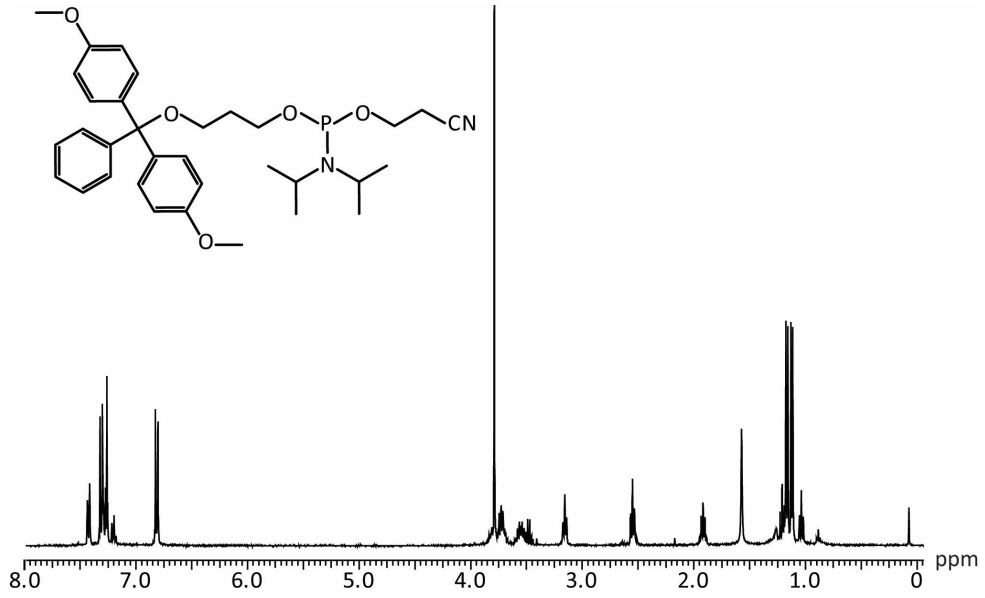


Figure S3. ^{1}H-NMR (400 MHz, $CDCl_3$) of 3-(4,4'-Dimethoxytrityloxy)propyl-1-[(2-cyanoethyl)-(*N*,*N*-diisopropyl)]-phosphoramidite (compound **1**)

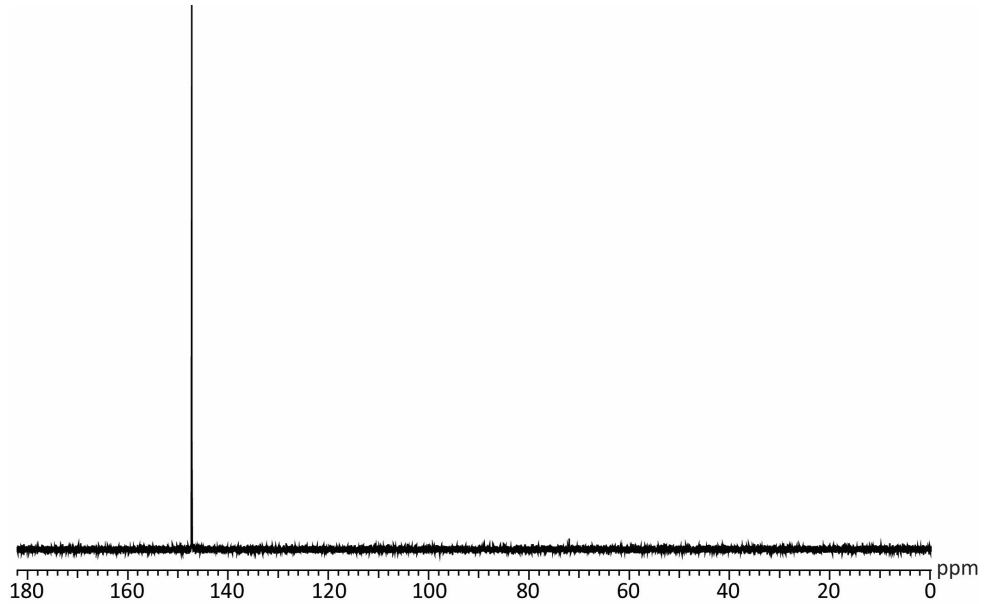


Figure S4. ^{31}P-NMR (121 MHz, $CDCl_3$) of 3-(4,4'-Dimethoxytrityloxy)propyl-1-[(2-cyanoethyl)-(*N*,*N*-diisopropyl)]-phosphoramidite (compound **1**)

Coupling efficiency of the phosphoramidite 1 in competition with RNA phosphoramidites

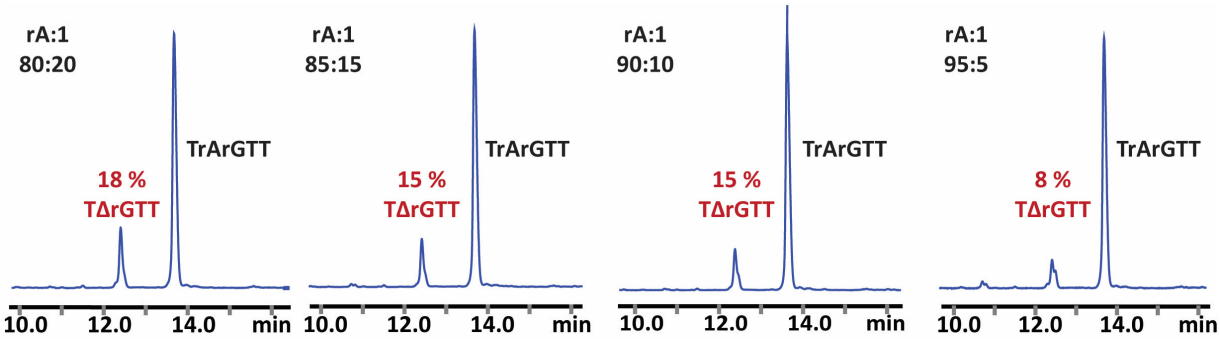


Figure S5. HPLC analysis of the oligonucleotides prepared to determine coupling efficiency of phosphoramidite **1** in RNA. Pentamer oligonucleotides were synthesized using mixtures of the rA and **1** phosphoramidites at the second position in a ratio of 80:20, 85:15, 90:10 and 95:5 at a final concentration of 100 mM. Dionex DNAPac PA200, 0-24% B in 6 CV. Eluent A: 25 mM Tris.HCl pH 8.0, 6M urea; eluent B: 25 mM Tris.HCl pH 8.0, 0.5 M $NaClO_4$, 6 M urea; 80°C, detection at 260 nm.

Optimization of the primer extension reaction

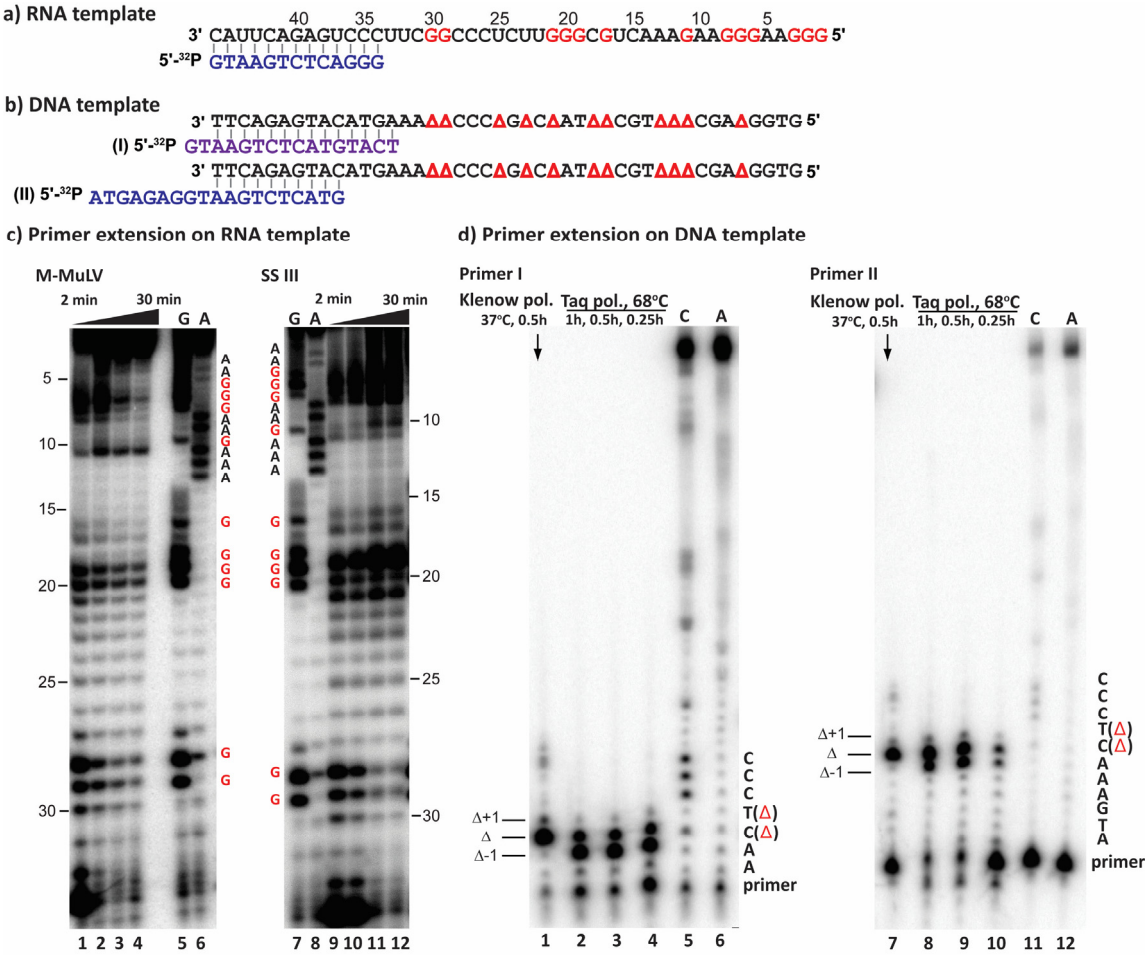


Figure S6. Optimization of the primer extension conditions on RNA and DNA templates containing the C3 spacers (Δ). a) RNA template used for primer extension reaction was synthesized using standard

phosphoramidite solutions at positions depicted in black and a mixture of rG and **1** phosphoramidites in a ratio of 95:5 at positions depicted in red, which resulted in a library-like RNA template where each G position is replaced by a C3 linker statistically once per RNA strand. b) DNA template was synthesized using phosphoramidite **1** (100 mM) at positions depicted in red. Primer extension was performed with 2 different primers (I and II). c) Gels represent primer extension pattern from reactions catalysed by M-MuLV reverse transcriptase (M-MuLV) and Super Script III reverse transcriptase (SS III) performed on RNA template that contains C3 spacers (a). Each primer extension reaction was performed for 2, 10, 20 and 30 minutes. The bands that correspond to the stops at Δ positions become less prominent when the reaction time is longer. The bands in the primer extension reactions performed on a template that contains C3 linkers are downshifted by 1 in reference to the sequencing ladders obtained from the unmodified template. This can be explained by the primer extension termination at the last position before the C3 spacer (Δ). Lanes: 1, 2, 3, 4: primer extension catalysed by M-MuLV RT for 2, 10, 20 and 30 min; 5, 6, 7, 8: sequencing ladders; 9, 10, 11, 12: primer extension catalysed by SS III RT for 2, 10, 20 and 30 min. d) Gels represent primer extension pattern from reactions performed on DNA template (b) with 2 primers (I and II) catalysed by Klenow fragment (exo-) and One Taq DNA polymerase. No full length product is observed in the reactions performed on templates that contain C3 spacers due to the primer extension termination at first Δ position (in agreement with expectations). When the Klenow fragment is used the most prominent band appears directly at Δ position (the last nucleotide in the primer extension reaction is incorporated opposite to the acyclic linker). When One Taq polymerase is used there are two prominent bands corresponding to one C3 spacer (band at Δ and at Δ-1) independent on the reaction time. Reproducible results were obtained with each primer. Lanes: 1 and 7: primer extension catalysed by Klenow fragment (exo-) at 37°C for 30 min; 2, 3, 4 and 8, 9, 10: primer extension catalysed by One Taq polymerase for 60, 30 and 15 min at 68°C; 5, 6 and 11, 12: sequencing ladders. Template used for sequencing: GTGGGAGCCTTTGCTATAGCTGTCCCTCAAAGTACATGAGACTT.

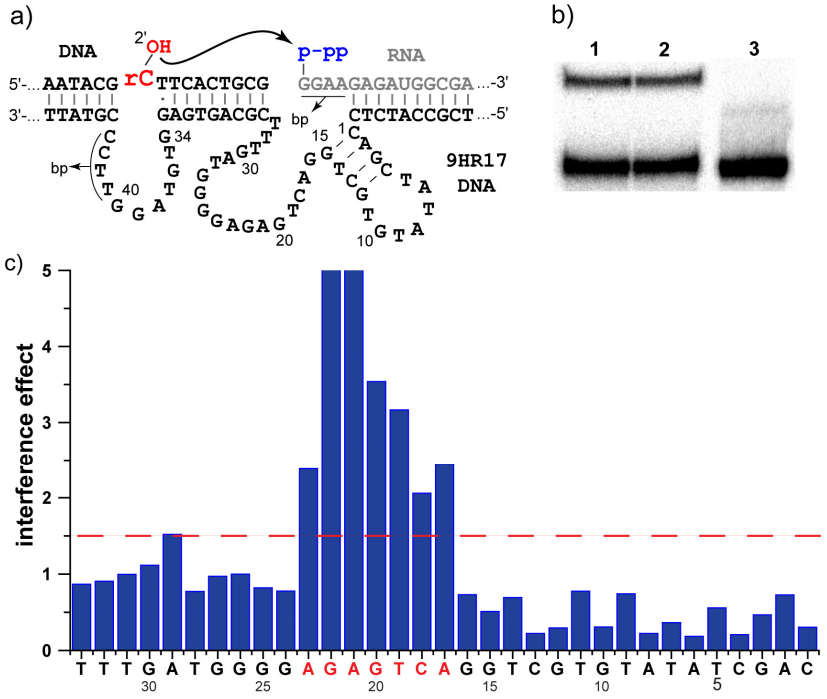


Figure S7. a) 9HR17 deoxyribozyme; b) separation of the active and inactive library variants of 9HR17 deoxyribozyme. Ligation yield was 20%. Lane 1 and 2: Separation reaction performed in the presence of 20 mM Mn^{2+} at pH 7.5, Lane 3: unseparated library. c) bar graph representing the analysis of primer extension results for loop A, nt 1-33. (Compare ref 5, Figure 4c; nucleotides in red (17-23 are most critical for catalysis and cannot be deleted).

HPLC analysis of the synthesized RNA libraries

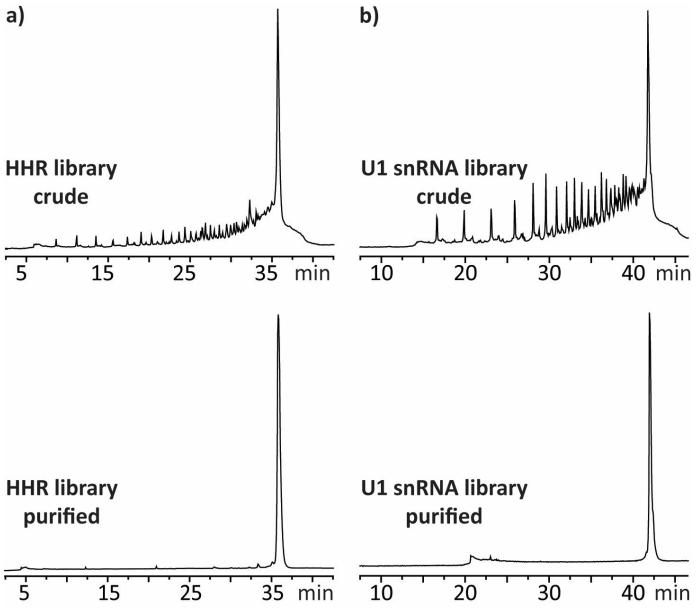


Figure S8. HPLC analysis of RNA and DNA libraries synthesized using solid-phase synthesis. a) Hammerhead ribozyme, 0-50% B in 12 CV; b) Hairpin II of U1A RNA, 0-52 % B in 13 CV; c) 7YK35 deoxyribozyme, 0-80% B in 20 CV; d) 10MD5AC deoxyribozyme, 0-60% B, 15 CV. Dionex DNAPac PA200. Eluent A: 25 mM Tris.HCl pH 8.0, 6 M urea; eluent B: 25 mM Tris.HCl pH 8.0, 0.5 M $NaClO_4$, 6 M urea; 80°C, detection at 260 nm.

Hammerhead ribozyme

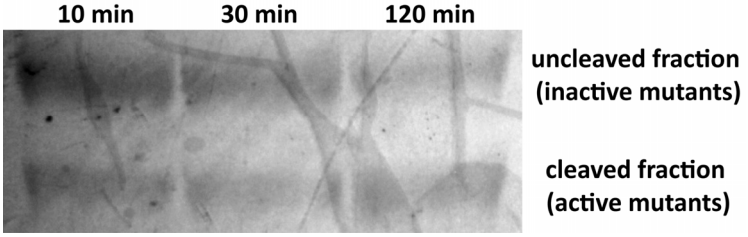


Figure S9. Separation of the active hammerhead ribozyme mutants. Cleavage reactions were performed in the presence of 10 mM $MgCl_2$ for 10 min, 30 min and 2 h. Cleaved molecules represent the active fraction of the library.

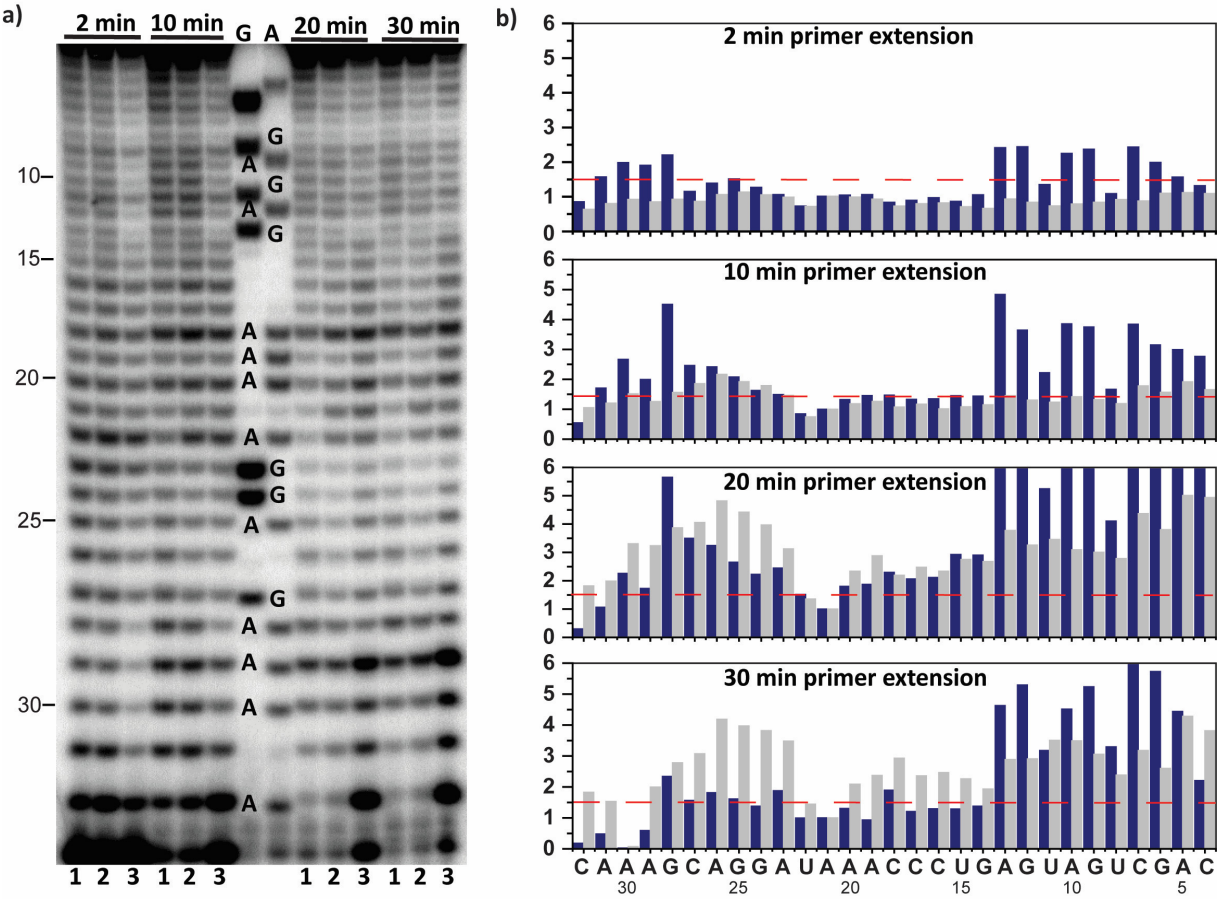


Figure S10. Primer extension analysis of active hammerhead ribozyme variants. a) Gel represents the primer extension pattern of the unseparated library (1) and catalytically active fraction after 2 h (2) and 10 min (3) cleavage reactions. Primer elongation for each sample was performed with M-MuLV reverse transcriptase for 2, 10, 20 and 30 min. b) Interference effect values calculated separately for each set of primer extension reactions. Blue bars represent the results from the primer elongation performed on the active fraction following 10 min separation reaction while grey bars correspond to the samples after 2 h separation reaction. (Extension of the figure 2 from the main paper).

Hairpin II of U1 snRNA

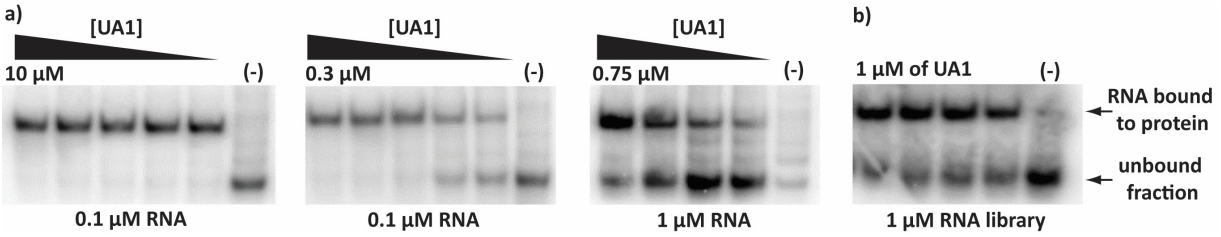


Figure S11. Representative gel mobility shift analysis of U1A protein binding to hairpin II of U1 snRNA. a) Binding experiments performed with in vitro transcribed non-modified U1 snRNA (0.1 µM and 1 µM). The highest protein concentration used for the assays was 10 µM and a 1:1 serial dilution was performed. The highest protein concentration on each gel is depicted in the upper left corner. b) Example of the separation experiment. The upper bands represent the active fraction of the library (RNA molecules

bound to the protein). The unbound fraction corresponds to inactive molecules. (-) refers to the control samples where no protein was present.

References for supporting information

1. P. Hendry and M. J. McCall, *Nucleic Acids Res*, 1995, **23**, 3928.
2. Y. Benitex and A. M. Baranger, *BMC biochemistry*, 2007, **8**, 22.
3. M. Takeshita, C. N. Chang, F. Johnson, S. Will and A. P. Grollman, *J. Biol. Chem.*, 1987, **262**, 10171.
4. M.Y.-X. Ma, L. S. Reid, S. C. Climie, W. C. Lin, R. Kuperman, M. Sumner-Smith, R. W. Barnett *Biochemistry*, 1993, **32**, 1751.
5. F. Wachowius, F. Javadi-Zarnaghi and C. Höbartner, *Angew. Chem. Int. Ed.*, 2010, **49**, 8504
6. B. Samanta and C. Höbartner, *Angew. Chem. Int. Ed.*, 2013, **52**, 2995.