Nucleic Acid Chemistry

DOI: 10.1002/anie.201208103

Combinatorial Nucleoside-Deletion-Scanning Mutagenesis of Functional DNA**

*Biswajit Samanta and Claudia Höbartner**

Artificial functional DNA molecules, such as deoxyribozymes and aptamers, find increasing applications in research areas ranging from biochemistry and molecular engineering to experimental medicine.[1] Such functional single-stranded oligonucleotides are generally identified by in vitro selection from random-sequence DNA pools.[2] The size of the initially randomized region is mostly arbitrarily chosen, and spans from as few as 20 nucleotides (nt) to more than 200 nt. Although DNA sequences with desired activities may initially be discovered from pools with longer random regions, short sequences are usually preferred for practical applications. Computational analyses including sequence-alignment and structure-prediction methods typically need large datasets to reliably identify functional motifs. These approaches are often limited by the absence of partial randomization and reselection data for reported functional sequences. Advances in this direction are expected based on increased utilization of next-generation sequencing techniques in nucleic acid in vitro selection and screening approaches.[3] However, for the analysis and optimization of known functional nucleic acids, an experimentally simple and reliable method for minimization and characterization of active sequences is currently not available.

Herein we present combinatorial nucleoside-deletion-scanning (NDS) mutagenesis as a new technique to efficiently distinguish essential from non-essential nucleosides in the catalytic region of deoxyribozymes, using a single synthetic DNA library. We expect that this analysis will provide minimal constructs for crystallization optimization, advance mechanistic studies of DNA catalysis, and reveal preferred regions for deoxyribozyme engineering. The method is based on statistical "deletion" of individual nucleosides in a synthetic DNA library, by replacement of original deoxyribonucleosides with non-nucleosidic spacer units. Conceptually, nucleoside-deletion-scanning mutagenesis is reminiscent of alanine-scanning mutagenesis in proteins, which is used to determine the catalytic or functional role of specific residues.[4] By replacing individual amino acids with alanine, the side chain beyond the β carbon atom is eliminated, but the main-chain conformation is retained. In distant analogy, herein we eliminate nucleosides in DNA by replacing one nucleoside per molecule with the acyclic spacer Δ (Figure 1).

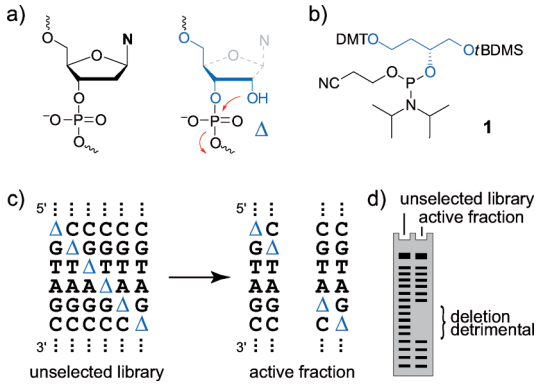


Figure 1. Concept of combinatorial nucleoside-deletion-scanning (NDS) mutagenesis. a) Comparison of a standard DNA nucleoside with the cleavable non-nucleosidic spacer Δ (blue; the gray part outlines the formally "deleted" portion of the nucleoside). Red arrows indicate the preferred pathway for cleavage of the phosphodiester backbone upon alkaline hydrolysis. b) Structure of phosphoramidite **1**, used for incorporation of Δ. c) Schematic representation of NDS; parent nucleosides are statistically replaced by Δ. After separation of inactive library members, the active fraction contains only those mutants, in which nucleoside deletion does not interfere with catalysis. d) Analysis of the interference pattern by denaturing polyacrylamide gel electrophoresis (PAGE) reveals nucleosides that are sensitive to replacement by Δ. DMT = 4,4′-dimethoxytrityl, *t*BDMS = *tert*-butyldimethylsilyl, N = nucleobase.

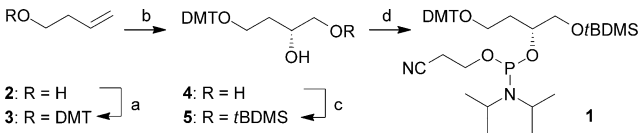
This spacer moiety is a 1,2,4-butantriol derivative that resembles an incomplete ribose residue, while the connectivity (i.e., six bonds between two phosphorous atoms) as well as the polarity of a standard 3′,5′-connected phosphodiester backbone are retained. The primary hydroxy group of Δ, which formally corresponds to the 2′-OH of a ribonucleoside, enables specific cleavage of the DNA backbone upon alkaline hydrolysis and thereby serves as a chemical tag to encode the nucleoside deletion in the combinatorial DNA library.[5] Owing to the open-chain structure of Δ, the flexibility of the backbone is increased. Thereby possible interference effects from the 2′-OH group in a cyclic abasic site analogue are avoided.

The acyclic spacer Δ was easily introduced into DNA by solid-phase synthesis using phosphoramidite **1**, which was

[*] B. Samanta, Dr. C. Höbartner
Research Group Nucleic Acid Chemistry, Max Planck Institute for Biophysical Chemistry
Am Fassberg 11, 37077 Göttingen (Germany)
E-mail: claudia.hoebartner@mpibpc.mpg.de
Homepage: http://www.mpibpc.mpg.de/hoebartner

[**] We are grateful to the Max Planck Society and the MPI for biophysical chemistry for the generous support of our research. Members of the research group Nucleic Acid Chemistry are gratefully acknowledged for stimulating discussions. We thank Uwe Plessmann and Jürgen Bienert from the chemistry and mass spectrometry facilities at MPIbpc for recording NMR and MS spectra.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201208103.



Scheme 1. Synthesis of phosphoramidite **1** for incorporation of Δ. Reagents and conditions: a) DMT-Cl, pyridine, room temperature, 6 h, 99%; b) AD-mix-β, *t*BuOH, H_2O, 0°C, 24 h, 73%; c) *t*BDMS-Cl, imidazole, CH_2Cl_2, room temperature, 16 h, 74%; d) (2-cyanoethyl) *N*,*N*-diisopropyl chlorophosphoramidite, Me_2NEt, CH_2Cl_2, RT, 45 min, 75%.

generated in four steps from 3-buten-1-ol (**2**, Scheme 1). First, the primary hydroxy group was protected as a dimethoxytrityl ether to give **3**. Sharpless asymmetric dihydroxylation of the alkene using AD-mix-β afforded the vicinal diol **4** in the desired *R* configuration.[6] Silylation of the primary hydroxy group of **4**, followed by phosphitylation of the secondary alcohol **5**, yielded the phosphoramidite **1**.

Upon mixing each of the four standard DNA nucleoside phosphoramidites with **1**, and using these mixtures in solid-phase synthesis, combinatorial NDS libraries of deoxyribozymes or any other functional DNA can be generated. The next steps resemble the analysis performed in other interference methods recently reported for characterization of functional DNA, including CoMA and dNAIM.[7] First, active and inactive library members are separated by electrophoresis or affinity chromatography, for example, based on the reaction catalyzed by a DNA enzyme of interest, or binding of a ligand in the case of aptamers (Figure 1c). The readout of the nucleoside deletion is then initiated by alkaline hydrolysis, which leads to cleavage of the DNA strands at the deleted nucleoside positions (i.e. at sites where Δ was incorporated), and is followed by PAGE analysis and quantification of the interference effect (Figure 1d). Missing hydrolysis bands in the active fraction suggest that nucleosides at these positions are essential for DNA function and must be retained. On the other hand, hydrolysis bands that are present in the active fraction reveal nucleosides that are dispensable for DNA activity. Shortened DNA sequences are then generated and functional assays are used to confirm the combinatorial screening results. Initially positive variations are then combined to look for synergistic improvements of activity, to optimize the minimal functional sequence.

For our proof-of-concept demonstration of combinatorial NDS mutagenesis of functional DNA, we chose two deoxyribozymes that catalyze the synthesis of covalently branched nucleic acids. These DNA enzymes activate the 2′-hydroxy group of an internal branch-site nucleotide in the scaffold strand substrate to form a new phosphodiester bond to the 5′-terminus of the adaptor strand substrate (Figure 2a).[8] Scaffold and adaptor strand can each be either DNA or RNA, resulting in four possible combinations (note: a DNA scaffold strand must have at least one internal ribonucleotide that serves as a branch site). Deoxyribozymes have been identified for each of these four combinations to generate branched nucleic acids for various applications.[9] The most prominent combination is 2′,5′-branched RNA, which resembles the core structure of lariat RNA, the product generated

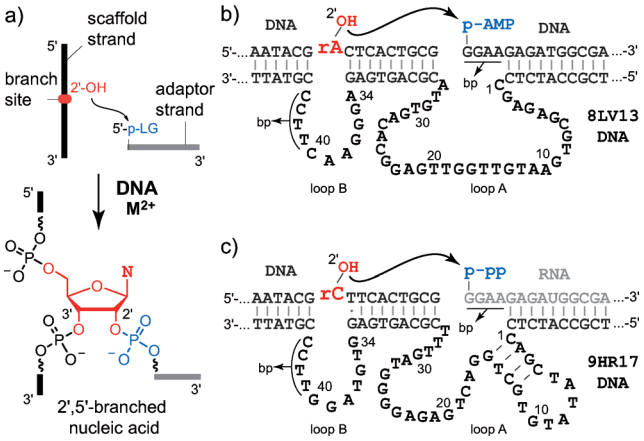


Figure 2. a) DNA-catalyzed synthesis of 2′,5′-branched nucleic acids. b) 8LV13 deoxyribozyme for synthesis of 2′,5′-branched DNA.[8] c) 9HR17 deoxyribozyme for connecting an RNA adaptor strand to a DNA scaffold strand.[12] The first four nucleotides of each adaptor strand form a base-paired (bp) region with the marked nucleotides of loop B to generate a three-helix junction structure. LG = leaving group.

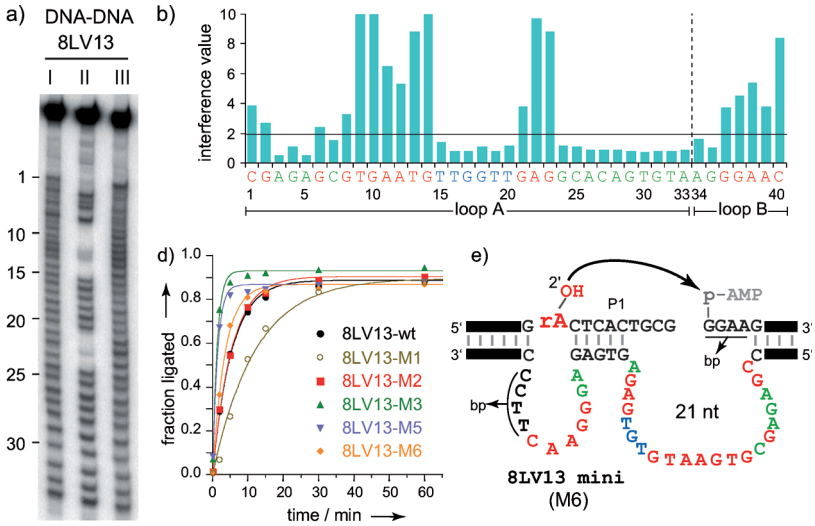
after the first step of RNA splicing.[10] RNA-DNA-branching deoxyriboyzmes have been used to attach DNA constraints to ribozymes for the regulation of RNA function, and the DNA-catalyzed synthesis of 2′,5′-branched DNA has been considered as construction element for DNA nanotechnology.[11] Herein, we study two deoxyribozymes that connect different adaptor strands to the branch-site ribonucleotide of DNA scaffold strands. The 8LV13 deoxyribozyme uses 5′-adenylated DNA as the adaptor substrate (Figure 2b),[8] and the 9HR17 deoxyribozyme uses 5′-triphosphorylated RNA to attach an RNA adaptor to the DNA scaffold (Figure 2c).[12] Both deoxyribozymes were selected in a structural context that allows for the formation of a three-helix junction structure and the formation of two single-stranded loop regions. Loop A contains 33 nt, and loop B has 7 nt, which together account for the 40 randomized nucleotides in the DNA library used for in vitro selection.

We demonstrate that combinatorial NDS mutagenesis can easily distinguish which of these 40 nt are required for the activity of the deoxyribozyme, and which ones play no essential role. The first step entailed the synthesis of the respective DNA libraries, containing statistically distributed Δ residues. To adjust the number of nucleoside deletions to about one Δ per molecule, we studied the incorporation efficiency of phosphoramidite **1** in competition with each of the four standard DNA phosphoramidites (dN). Pentameric model oligonucleotides were synthesized with different dN:**1** ratios. After deprotection, the products were analyzed by HPLC, and the product ratio was determined by integrating the peak area of the UV absorbance trace.[13] The desired incorporation ratio was achieved by using 30% **1** at a phosphoramidite concentration of 100 mM, and a coupling time of four minutes. Using these conditions, deoxyribozyme libraries were synthesized, for which the phosphoramidite mixtures were employed only for loops A and B (the binding arms were synthesized with standard DNA phosphoramidites only). After deprotection and purification, the DNA libraries

were ligated to the respective adaptor strand substrates, and then ^{32}P-labeled on the 3′-end. The DNA-catalyzed ligation reaction was performed with the respective scaffold strand substrates. Active and non-ligated fractions were isolated by electrophoresis (after various incubation times), and then individually subjected to alkaline hydrolysis. The optimal hydrolysis conditions (0.4 M NaOH, 95 °C, 15 min) were determined using a model substrate.[13] Owing to the higher flexibility of the primary hydroxy group in Δ, harsher hydrolysis conditions were necessary compared to CoMA and dNAIM.[7] The nucleoside deletion mutagenesis results for the 8LV13 deoxyribozyme are depicted in Figure 3, and the results for 9HR17 are summarized in Figure 4. The interference results are depicted as bar graphs, with the nucleotide sequence on the *x*-axis and the interference value (ratio of band intensities for each nucleotide in the full library, lane I, and active fraction, lane II) as *y*-axis. Large interference values indicate detrimental effects, while low values (less than 2) mean that the deletion mutation is tolerated.

A distinct pattern was observed in the hydrolysis lane of the active 8LV13 library fraction (Figure 3a, lane II), clearly indicating several missing bands (most prominently from nt 8–14 and nt 21–23 in loop A), suggesting that these regions are essential for catalysis. In the 3′-region of loop A, from nt 24–33, all hydrolysis bands were present, indicating that these nucleotides are not directly required, and suggesting that this region could be shortened or potentially removed entirely. Other stretches of non-essential nucleosides were found from nt 15–20, and from nt 3–5. In loop B, all nucleosides except nt 34 and 35 showed large interference effects. To examine the effect of removing non-essential nucleotides from the DNA sequence, several shorter versions of the 8LV13 sequence were then prepared, and their activity was determined in kinetic ligation assays (Figure 3c,d). We found that loop A could be shortened by deleting nt 24–32 (derivative M2) without affecting the ligation activity; that is, M2 was as active as the wt DNA. Interesting effects were observed for the region of nt 15–20, which could not be entirely deleted, nor replaced by an arbitrary sequence of the same length (see Table S2 in the Supporting Information), even though the NDS results suggested that individually replacing any one of these six nucleotides with Δ had no effect on the catalytic activity. The repetitive sequence element TTGGTT represents an interesting feature of the 8LV13 deoxyribozyme, which could successfully be replaced by TGT, that is, deleting one nucleotide of each of the three nucleotide pairs was tolerated. These results demonstrate that combinatorial NDS mutagenesis can identify critical nucleosides upon careful analysis of the sequence context. The shortened 8LV13 variant M3 (with TGT) gave an almost fivefold increase in k_{obs}. The accelerating effect of this sequence change was also observed in a second example, as depicted in the summary of active 8LV13 mutants in Figure S18.[13] Moreover, the DNA binding arm that hybridizes to the scaffold strand in the 3′ direction of the branch-site (P1) could be shortened from 9 to 5 nt by arbitrary mutation or deletion of four nucleotides towards loop A. The shortest 8LV13 variant M6 with a 5 bp stem and a 21-nt loop A is still about twofold faster than the wt DNA which has a 9 bp-stem and a 33 nt loop A. The k_{obs} of this minimized version of 8LV13 also compares highly favorably with another related and previously reported DNA-branching deoxyribozyme (15HA9, which has a 5 bp binding arm and contains 15 nt in loop A).[11c]

Combinatorial NDS mutagenesis can also provide information on nucleoside requirements under different reaction conditions, for example, using different metal-ion cofactors for DNA-catalyzed reactions, as exemplified herein for the 9HR17 deoxyribozyme. This DNA enzyme was originally selected to catalyze the ligation of an RNA adaptor to a DNA scaffold in the presence of 40 mM Mg^{2+} at pH 9.0, but the details of metal-ion requirements were not analyzed.[12] For several other deoxyribozymes that catalyze RNA cleavage,[14] RNA ligation,[15] or the formation of nucleopeptide linkages,[12] faster reaction rates have been reported with Mn^{2+} rather than Mg^{2+}. We found that 9HR17 performs 16-fold faster ligation of DNA and RNA substrates when 20 mM Mn^{2+} is pro-



c)

No	Sequence	k_{obs} [min^{-1}]	mut/wt
	1 5 10 15 20 25 30 ---P1---- 35 40		
8LV13 wt	CGAGAGCGTGAATGTTGGTTGAGGCACAGTGTACGCAGTGAGAGGGAAC	0.17	1.0
M1	CGAGAGCGTGAATGTTGGTTGAGGCACAGTGTACGCAGTGAGA GGAAC	0.09	0.5
M2	CGAGAGCGTGAATGTTGGTTGAG ACGCAGTGAGAGGGAAC	0.16	1.0
M3	CGAGAGCGTGAATGTGT GAG ACGCAGTGAGAGGGAAC	0.80	4.7
M4	CGTT GCGTGAATGTGT GAG ACGCAGTGAGAGGGAAC	0.25	1.5
M5	CGTT GCGTGAATGTGT GAG ATTTTGTGAGAGGGAAC	0.73	4.3
M6	CGAGAGCGTGAATGTGT GAG AGTGAGAGGGAAC	0.29	1.7

Figure 3. Combinatorial NDS mutagenesis for 8LV13-catalyzed DNA–DNA ligation. a) Hydrolysis gel for 33 nt-loop A (full gel in Supporting Information) Hydrolysis samples: I unselected library, II active fraction after ligation, 20 mM Mn^{2+}, pH 7.5, 37 °C, 1.5 h. III unligated fraction. b) Interference values for each nucleotide; green: nucleotides with interference values less than 2, except nt 15–20 (see text for details), red: values over 2. c) Selection of mutants and their k_{obs} values. d) Kinetic plots for selected 8LV13 mutants. e) Minimized 8LV13 DNA (M6). wt = wild type.

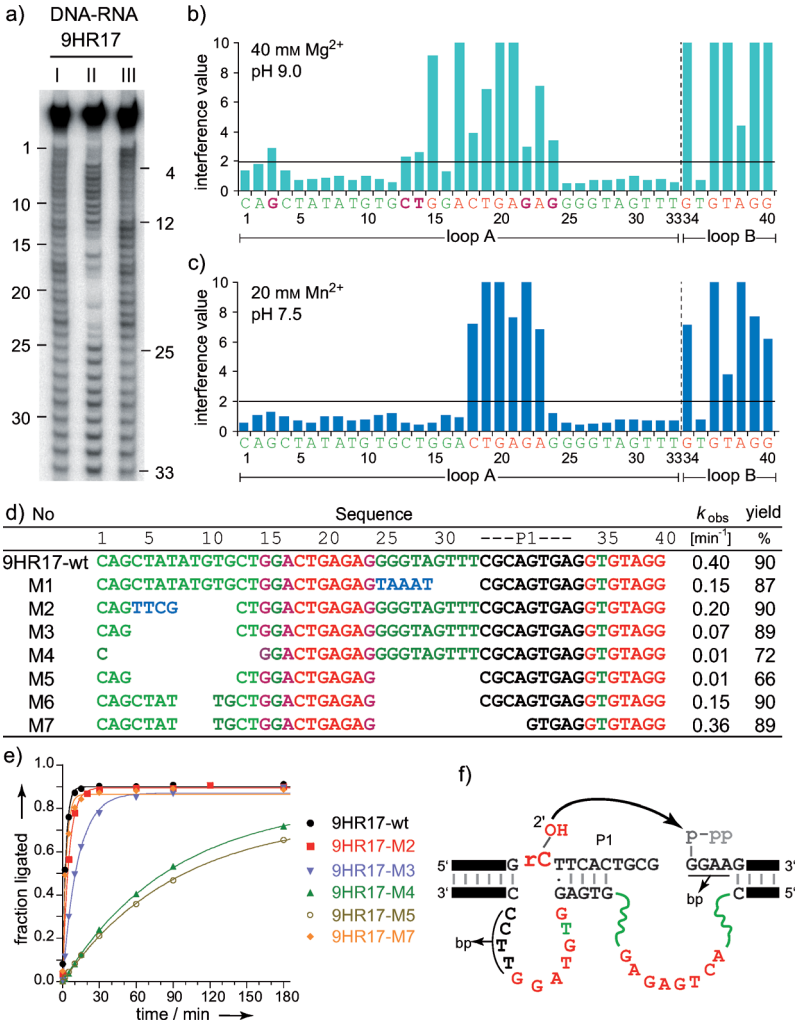


Figure 4. Combinatorial NDS mutagenesis of 9HR17-catalyzed DNA–RNA ligation. a) Hydrolysis gel for 33 nt-loop A (full gel in Supporting Information). Hydrolysis samples: I unselected library, II active fraction after ligation, 40 mM Mg^{2+}, pH 9.0, 37 °C, 1.5 h. III unligated fraction. b,c) Interference values for ligation with Mg^{2+} (b) and Mn^{2+} (c); nucleotides with interference values less than 2 are green, those with values over 2 are red; purple indicates differences in presence of Mg^{2+} and Mn^{2+}. d) Selection of truncated mutants and their k_{obs} values and final ligation yields (after 3 h reaction time in 20 mM Mn^{2+}). e) Kinetic plots for selected 9HR17 mutants. f) Schematic diagram of minimized 9HR17 with 5 nt in P1 and essential nucleotides of loop A highlighted. The green lines indicate several options for productively positioning the core nucleotides.[13]

vided at pH 7.5, compared to the original conditions. Therefore, we analyzed the NDS interference pattern of 9HR17 for ligation with each of these divalent metal ions. At both ends of loop A, at least 9 nt were not essential for catalysis with either Mg^{2+} or Mn^{2+} (Figure 4b,c). A central region of six nucleotides (nt 18–23) did not tolerate the deletion when the ligation was performed in the presence of Mn^{2+}. The critical region was three nucleotides longer in the presence of Mg^{2+} (purple nucleotides in Figure 4d). Since only a few nucleotides were essential in the center of loop A, only one of the flanking regions could be deleted while maintaining the proper positioning of the central nucleotides. Moreover, our results indicate that the predicted stem-loop in loop A (Figure 2c) is not relevant for the active conformation of the deoxyribozyme. These conclusions were confirmed by kinetic assays of the ligation activity using several 9HR17 derivatives (Figure 4d,e) in which either the 5′ or the 3′ part of loop A was truncated. For several mutants we found that shortening the loop sequence was better tolerated in presence of Mn^{2+} than Mg^{2+}, as can be deduced from the ratios of k_{obs} values listed in Table S4.[13] As in the case of 8 LV13 variants, we found that the minimized version of 9HR17 (M6) did not require the 9 bp interaction with the RNA substrate in P1. Indeed, by shortening this region to 5 bp in the context of the minimized loop (M7), the k_{obs} improved about twofold and compared well with the k_{obs} of wt 9HR17 (Figure 4d,e) demonstrating that this shortened DNA is a highly efficient deoxyribozyme.

In summary, we have developed combinatorial NDS mutagenesis as a simple and effective method to distinguish essential from non-essential nucleosides in the catalytic loop regions of deoxyribozymes. This new combinatorial approach is less laborious compared to systematic replacement of nucleotides by an abasic site analogue or an alkyl linker in individual DNA strands, which is practicable only for a small number of nucleotides, such as recently reported for the RNA-cleaving 10–23 and 8–17 deoxyribozymes.[16] For the DNA catalysts analyzed in this study, we found that a significantly smaller number of nucleotides than originally present in the random pool is sufficient to support efficient catalysis. This finding is of general interest for future in vitro selection experiments and may stimulate further investigations on the correlation between nucleic acid catalysis and random sequence length.[17]

Combinatorial NDS mutagenesis represents a novel, resource- and time-efficient approach for deoxyribozyme characterization and engineering and will be easily adaptable to other functional DNAs, including ligand-binding DNA aptamers. The most obvious practical advantage of NDS compared to CoMA[7a] lies in the faster data collection, since only a single synthetic library is required, and several functional DNAs can easily be analyzed in parallel and under different reaction conditions. The results of the NDS interference analysis narrow down the number of nucleotides for structural analyses. Shorter sequences are less prone to misfolding, which increases the probability for crystallization. For NMR spectroscopy, minimized systems are favorable for improved resonance separation and reduced cost for isotope-labeling. Furthermore, by identifying nucleotides that cannot be replaced by Δ, NDS directly maps crucial residues that are promising to reveal mechanistic insights on DNA catalysis upon further analyses, for example, via dNAIM,[7b] in which a large number of combinatorial

libraries are synthesized using precious modified ribonucleotides. In a practical sense, the NDS results directly suggest permissive sites for insertion of regulatory elements or replacement of non-essential nucleotides with external functional units. The minimized deoxyribozymes offer ample opportunities for further engineering, for example in combination with aptamers to build aptazymes for the development of sensors and as construction elements in DNA nanotechnology.

Received: October 8, 2012
Revised: January 3, 2013
Published online: February 1, 2013

Keywords: deoxyribozymes · DNA catalyst · interference analysis · mutagenesis · solid-phase synthesis

[1] a) I. Willner, B. Shlyahovsky, M. Zayats, B. Willner, *Chem. Soc. Rev.* **2008**, *37*, 1153; b) M. Famulok, G. Mayer, *Acc. Chem. Res.* **2011**, *44*, 1349; c) B. Strehlitz, C. Reinemann, S. Linkorn, R. Stoltenburg, *Bioanal. Rev.* **2012**, *4*, 1; d) S. K. Silverman, *Angew. Chem.* **2010**, *122*, 7336; *Angew. Chem. Int. Ed.* **2010**, *49*, 7180; e) M. Ye, J. Hu, M. Peng, J. Liu, H. Liu, X. Zhao, W. Tan, *Int. J. Mol. Sci.* **2012**, *13*, 3341.

[2] a) K. Schlosser, Y. Li, *Chem. Biol.* **2009**, *16*, 311; b) G. F. Joyce, *Annu. Rev. Biochem.* **2004**, *73*, 791; c) G. F. Joyce, *Angew. Chem.* **2007**, *119*, 6540; *Angew. Chem. Int. Ed.* **2007**, *46*, 6420.

[3] a) J. Hoinka, E. Zotenko, A. Friedman, Z. E. Sauna, T. M. Przytycka, *Bioinformatics* **2012**, *28*, i215; b) T. Schütze, B. Wilhelm, N. Greiner, H. Braun, F. Peter, M. Morl, V. A. Erdmann, H. Lehrach, Z. Konthur, M. Menger, P. F. Arndt, J. Glokler, *PLoS One* **2011**, *6*, e29604.

[4] a) B. C. Cunningham, J. A. Wells, *Science* **1989**, *244*, 1081; b) K. L. Morrison, G. A. Weiss, *Curr. Opin. Chem. Biol.* **2001**, *5*, 302.

[5] See Supporting Information for a statement on the choice of the butantriol-derived spacer Δ.

[6] a) K. B. Sharpless, W. Amberg, Y. L. Bennani, G. A. Crispino, J. Hartung, K. S. Jeong, H. L. Kwong, K. Morikawa, Z. M. Wang, *J. Org. Chem.* **1992**, *57*, 2768; b) H. C. Kolb, M. S. VanNieuwenhze, K. B. Sharpless, *Chem. Rev.* **1994**, *94*, 2483.

[7] a) F. Wachowius, F. Javadi-Zarnaghi, C. Höbartner, *Angew. Chem.* **2010**, *122*, 8682; *Angew. Chem. Int. Ed.* **2010**, *49*, 8504; b) F. Wachowius, C. Höbartner, *J. Am. Chem. Soc.* **2011**, *133*, 14888.

[8] C. S. Lee, T. P. Mui, S. K. Silverman, *Nucleic Acids Res.* **2011**, *39*, 269.

[9] C. Höbartner, P. I. Pradeepkumar in *New strategies in chemical synthesis and catalysis* (Ed.: B. Pignataro), Wiley-VCH, Weinheim, **2012**, p. 129.

[10] a) R. L. Coppins, S. K. Silverman, *Nat. Struct. Mol. Biol.* **2004**, *11*, 270; b) S. K. Silverman, *Acc. Chem. Res.* **2009**, *42*, 1521.

[11] a) E. Zelin, S. K. Silverman, *ChemBioChem* **2007**, *8*, 1907; b) E. Zelin, S. K. Silverman, *Chem. Commun.* **2009**, 767; c) T. P. Mui, S. K. Silverman, *Org. Lett.* **2008**, *10*, 4417.

[12] P. I. Pradeepkumar, C. Höbartner, D. A. Baum, S. K. Silverman, *Angew. Chem.* **2008**, *120*, 1777; *Angew. Chem. Int. Ed.* **2008**, *47*, 1753.

[13] See the Supporting Information for further details.

[14] a) R. P. G. Cruz, J. B. Withers, Y. Li, *Chem. Biol.* **2004**, *11*, 57; b) A. R. Feldman, E. K. Leung, A. J. Bennet, D. Sen, *ChemBioChem* **2006**, *7*, 98.

[15] R. L. Coppins, S. K. Silverman, *J. Am. Chem. Soc.* **2005**, *127*, 2900.

[16] B. Wang, L. Cao, W. Chiuman, Y. Li, Z. Xi, *Biochemistry* **2010**, *49*, 7553.

[17] An experimental study on that topic was published while this manuscript was under review: T. E. Velez, J. Singh, Y. Xiao, E. C. Allen, O. Y. Wong, M. Chandra, S. C. Kwon, S. K. Silverman, *ACS Comb. Sci.* **2012**, *14*, 680.

