DNA Enzymes

DOI: 10.1002/anie.201003940

Combinatorial Mutation Interference Analysis Reveals Functional Nucleotides Required for DNA Catalysis**

*Falk Wachowius, Fatemeh Javadi-Zarnaghi, and Claudia Höbartner**

DNA catalysts, also known as deoxyribozymes or DNA enzymes, are synthetic single-stranded DNA molecules that can catalyze chemical transformations with high selectivity. Since the first report of a DNA-catalyzed cleavage of an RNA phosphodiester linkage,[1] deoxyribozymes for a variety of reactions have been identified by in vitro selection.[2] Practical applications of DNA catalysts include their use as analytical tools, computational devices, and therapeutic agents, and as reagents for synthesis.[3] The DNA-catalyzed ligation of RNA is an experimentally attractive alternative to protein-catalyzed RNA ligation.[4] A powerful application of DNA catalysts is the synthesis of 2′,5′-branched RNA by activating a specific internal 2′-hydroxy group of one RNA substrate (L-RNA) for the nucleophilic attack to the 5′-triphosphate of the second RNA substrate (R-RNA). The prototype of this class of RNA ligases is the 7S11 deoxyribozyme[5] (Figure 1 a), which provides access to the 2′,5′-branched core structures of lariat RNAs, important RNA-splicing intermediates[6] that are difficult to obtain by other chemical methods.[7] Deoxyribozymes can also serve as useful tools for the linear ligation of two RNA fragments. The 9DB1 deoxyribozyme[8] (Figure 1 b) catalyzes the formation of a native 3′-5′-phosphodiester bond between two RNA substrates, using the 3′-OH of the L-RNA as a nucleophile to react with the 5′-triphosphate of the R-RNA.

The chemical mechanism of DNA-catalyzed RNA ligation is not known. Deoxyribozymes bind their RNA substrates by means of Watson–Crick base-pairing; the binding arms are connected by one or more single-stranded regions (Figure 1) that form the active sites for catalysis. For the single-stranded loops A and B of 7S11, limited mutagenesis data have provided preliminary information about functionally important nucleotides.[5b] In contrast, the nucleotides required for activity of the 9DB1 DNA have not yet been defined.

Understanding the functionality and sequence requirements of DNA catalysts is important from the mechanistic point of view and is expected to promote the engineering of deoxyribozymes for practical applications. In the absence of any three-dimensional structure of a DNA catalyst in an active conformation,[9] the identification of nucleotide functional groups that are essential for deoxyribozyme activity is fundamental to understanding the mechanisms of DNA catalysis. Traditional characterization methods are based on systematic deletion or substitution of individual nucleotides and meticulous kinetic analyses of many separate deoxyribozyme mutants.[10] Innovative alternatives to this rather labor-intensive approach should rapidly provide comprehensive and reliable data sets that permit the assessment of individual nucleotide contributions to catalytic activity. The results from such a comprehensive mutation analysis are expected to provide new insights into the molecular basis of DNA-catalyzed reactions.

We report herein the development of a combinatorial approach to mutation interference analysis that serves as a general tool for the characterization of functional single-stranded DNA, in particular for the identification of catalytically important nucleotides in deoxyribozymes. Combinatorial mutation interference analysis (CoMA) enables the simultaneous assessment of the catalytic ability of all possible single mutants of a deoxyribozyme (e.g., 120 single mutants for a deoxyribozyme with a 40-nucleotide catalytic region). For this endeavor, deoxyribozyme mutants are prepared in four combinatorial libraries by solid-phase synthesis. To

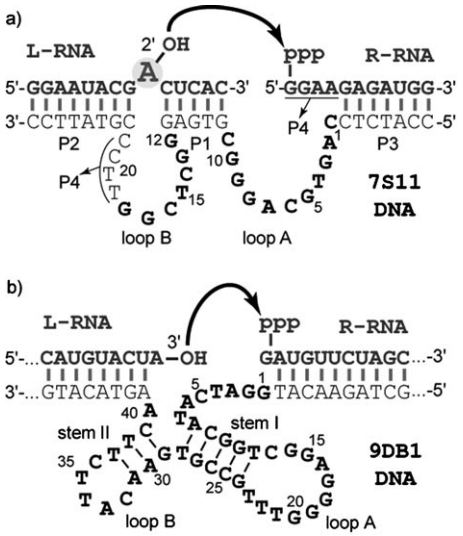

Figure 1. DNA-catalyzed RNA ligation. a) 7S11 enables the synthesis of 2′,5′-branched RNA.[5] b) 9DB1 joins two RNA substrates by means of a linear 3′-5′-phosphodiester linkage.[8] Secondary-structure prediction (mfold[11]) suggests two stem–loop domains in 9DB1. Both deoxyribozymes require Mg^{2+} as a divalent metal-ion cofactor.

[*] F. Wachowius, F. Javadi-Zarnaghi, Dr. C. Höbartner
Research group Nucleic Acid Chemistry
Max Planck Institute for Biophysical Chemistry
Am Fassberg 11, 37077 Göttingen (Germany)
Fax: (+49) 551-201-1680
E-mail: claudia.hoebartner@mpibpc.mpg.de

[**] Financial Support from the Max Planck Society is gratefully acknowledged. We thank Prof. S. K. Silverman, Prof. R. Micura, and Prof. P. I. Pradeepkumar for helpful comments and discussions.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201003940.

View this journal online at wileyonlinelibrary.com

© 2010 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

encode the nucleobase mutations in the deoxyribozyme libraries, the 2′-OH group of the ribonucleotides is used as a chemical tag. Each library therefore contains one of the four standard ribonucleotides statistically distributed in the catalytic core of the deoxyribozyme. The CoMA workflow consists of four steps as depicted in Figure 2: A) solid-phase synthesis of four 2′-OH-encoded combinatorial mutation libraries, B) separation of active and inactive library members, C) specific backbone cleavage at mutation sites by alkaline hydrolysis, and D) analysis of interference patterns by denaturing polyacrylamide gel electrophoresis (PAGE).

The solid-phase synthesis in step A uses standard DNA synthesis conditions, with the exception that mixtures of phosphoramidite solutions are employed (Figure 2A). The four 2′-*O*-triisopropylsilyloxymethyl (TOM)-protected ribonucleotide phosphoramidites (rN) are individually mixed with each of the four deoxyribonucleotide phosphoramidites (dN) in a ratio that would ideally result in one ribonucleotide mutation per DNA molecule on average. The efficiency of rN incorporation for different rN/dN ratios was examined for all 16 combinations by synthesizing pentamer oligonucleotides containing a single ribonucleotide mutation. The amount of rN incorporation was analyzed by anion-exchange HPLC based on the separation of parent (all DNA) from mutant (2′-*O*-TOM-rN-containing) pentamers. The experimental rN incorporation ratio was in the range of 40–70% of the rN content in the rN/dN phosphoramidite mixtures.[12]

In step B, active RNA ligase deoxyribozyme mutants are separated from inactive derivatives (Figure 2B) by means of DNA-catalyzed RNA ligation in a bimolecular format. In this setup, the R-RNA substrate is covalently linked to the 3′-^{32}P-labeled mutant DNA enzyme library, and the active DNA–RNA conjugates become attached to the L-RNA. Both active and inactive fractions are readily separated by denaturing PAGE. In step C, the unseparated mutant library as well as the active and the inactive fractions of the separated library are individually hydrolyzed under alkaline conditions (Figure 2C). In step D, the cleavage products are separated

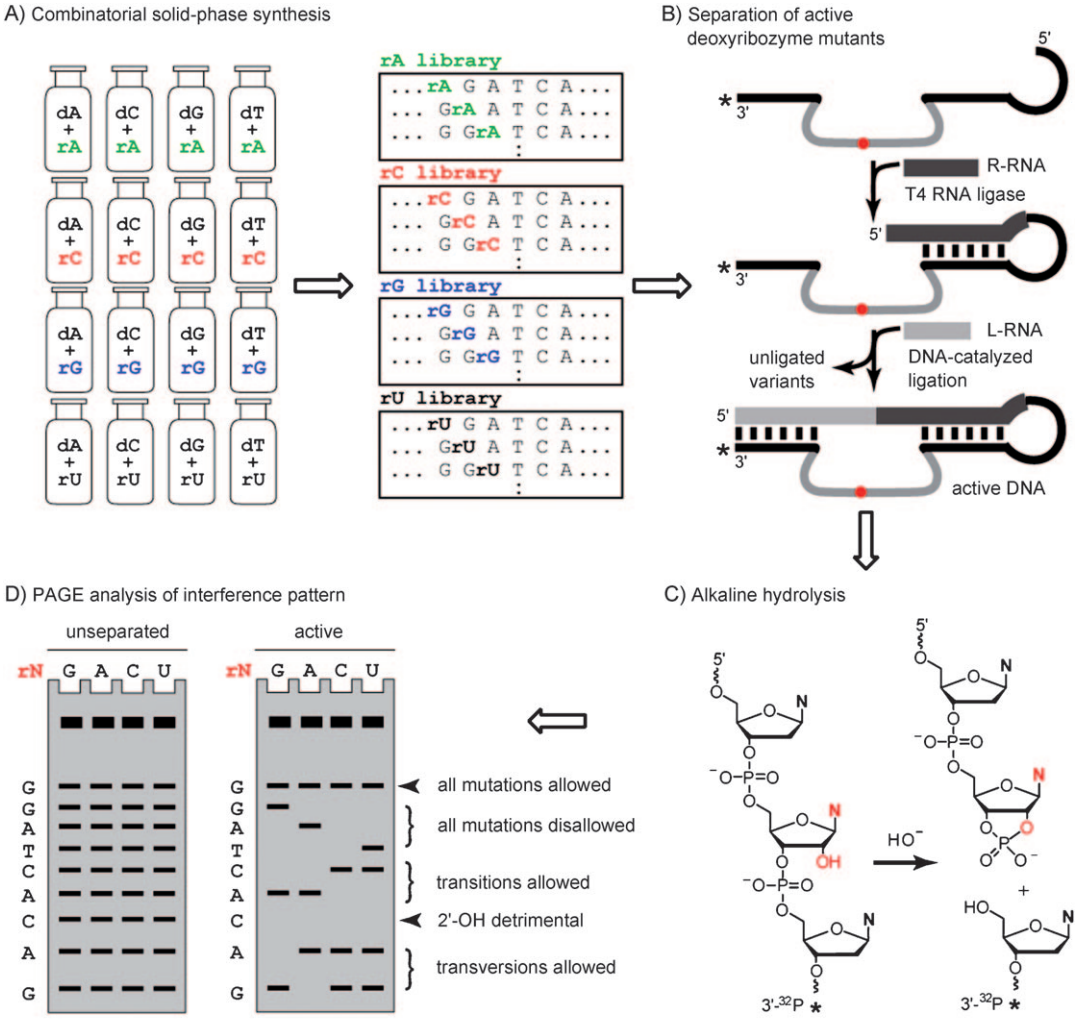

Figure 2. Workflow of steps A–D for combinatorial mutation interference analysis (CoMA). A) Mixtures of DNA and RNA phosphoramidites are used for solid-phase synthesis of mutant deoxyribozyme libraries. B) The DNA-catalyzed reaction is performed in a bimolecular format to separate active from inactive library members. C) The 2′-OH group enables specific backbone cleavage at mutated positions. D) PAGE analysis for missing bands reveals mutation effects. Transitions are G↔A, C↔U, transversions are G↔U, A↔C, and G↔C, A↔U.

 © 2010 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

on a denaturing polyacrylamide gel, and the resulting interference pattern is analyzed using a PhosphorImager.

Inspection of the hydrolysis lanes for the presence or absence of bands in all four libraries allows the assessment of mutation effects (see Figure 2D for a schematic representation). For a given nucleotide position, the presence of hydrolysis products in all four mutant libraries indicates that nucleobase mutations at this position are well tolerated and that the 2′-OH tag does not interfere with catalysis. In contrast, the absence of the hydrolysis product in all four mutant libraries indicates that the 2′-OH substitution at this nucleotide position is not tolerated. This suggests sensitivity to local conformational changes at the ribose moiety, in which case the effect of nucleobase mutations cannot be determined. The presence of a hydrolysis band only in the "parent" library of a given position (i.e., the absence of bands in the other three libraries) indicates that any type of standard nucleobase mutation at this position is detrimental to catalytic activity. At nucleotide positions at which transitions and/or transversions are allowed, hydrolysis products appear in more than one library. Numerical values for the interference effects of nucleotide mutations are obtained by quantification of individual band intensities (see below).

To demonstrate the concept of CoMA, we investigated the catalytic loops of the 7S11 deoxyribozyme. To assess whether individual ribonucleotides (i.e., the presence of the 2′-OH tag) would inhibit 7S11-catalyzed RNA ligation, we first synthesized a library of 7S11 variants in which only the nine 2′-deoxyguanosines in loops A and B were statistically replaced by the corresponding guanosine ribonucleotide. The rG-containing 7S11 library was covalently attached to the R-RNA substrate using T4 RNA ligase. The ligation products were radioactively labeled at their 3′-end by templated addition of ^{32}P-dATP with Klenow DNA polymerase. The 7S11-catalyzed ligation reaction was run in the presence of 40 mM $MgCl_2$, and the ligated and unligated fractions were isolated by denaturing PAGE. Both fractions were individually hydrolyzed with 10 mM NaOH at 95 °C for 10 min, and the cleavage products were separated on a sequencing gel. The analysis of the hydrolysis pattern revealed the presence of all nine 2′-OH-tagged DNA molecules in both fractions, indicating that no single guanosine ribonucleotide inhibited the catalytic ability of 7S11.[12] Similar results were obtained with libraries that contained rC, rA, and rU nucleotides at their parent deoxyribozyme positions.[12]

Based on these observations, we targeted the synthesis of four ribonucleotide mutant libraries of 7S11 for a comprehensive analysis of mutation interference effects at all loop nucleotide positions. The loop regions (including P4, see Figure 1A) were synthesized with rN/dN mixtures, whereas the substrate binding arms P1–P3 were synthesized with standard DNA phosphoramidites. The hydrolysis pattern of the four active 7S11 fractions in comparison to the four unseparated libraries is shown in Figure 3.[13] Individual band intensities were quantified, and interference values were calculated by dividing the band intensity of every nucleotide position in the unseparated library by the band intensity in the active fraction.[12] This resulted in an interference value of 1 if the modification (i.e., nucleobase mutation and 2′-OH group) had no effect on the catalytic activity. Values >1 indicated positions where the mutation inhibited activity, whereas values <1 (not actually observed) would have represented positions where mutations enhance the activity. The interference data were grouped for transitions (G↔A, C↔U; blue), transversions 1 (G↔U, A↔C; magenta), and transversions 2 (G↔C, A↔U; turquoise) for all loop nucleotide positions and represented as bar graphs. Interference values between 0.5 and 2 are considered insignificant and are marked in green (see the representation of 7S11 loop sequences, Figure 3 bottom). Medium interference values between 2 and 5 are colored pink, and strong interference values >5 are marked in red. No significant interference of the 2′-OH group by itself was observed at any 7S11 loop position (gray bar in Figure 3).

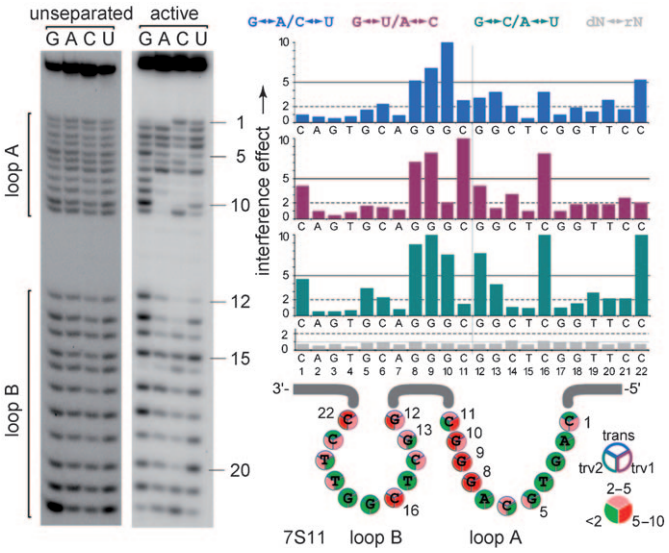

Figure 3. CoMA of 7S11. Gels show the alkaline-hydrolysis pattern of unseparated libraries and of the active fractions of 7S11 mutants. The DNA is ^{32}P-labeled at the 3′-end; therefore, nucleotide numbering from 5′ to 3′ runs from top to bottom on the gel. Interference values for transitions and both kinds of transversions are shown as bar graphs for loops A and B (including P4 as part of loop B). The circular color representation of the loop sequences at the bottom of the figure summarizes strong (red), weak (pink), and negligible (green) interference effects for transitions (trans; top third), transversions 1 (trv1; right third), and transversions 2 (trv2; left third) at all nucleotide positions.

The analysis of the 7S11 CoMA data identified three conserved guanosine nucleotides G8, G9, and G10 in loop A that cannot be changed to any other nucleotide without severely affecting catalytic activity. The interference values at these positions are >5, except for G10U with an interference value of 2.1. This smaller effect for the G10U mutation could indicate that the O6/N1 lactam functionality at position G10 contributes to catalysis, and that this function can partially be maintained by O4/N3 of uridine. The very strong effects at the other two guanosines for all mutations could be interpreted as more than one functional group of the nucleobase being essential for activity. A strong interference effect in loop A was also observed for C11A. Medium effects were detected for changing C1 into purines and for the C5G and C11U mutations. At position 6, the C6A mutation was well

 © 2010 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

tolerated, whereas diminished activity was observed when C6 was changed to U or G. This could indicate a functional importance of the exocyclic N4 amino group that can be imitated by N6 of adenosine. In loop B, nucleotides G12, C16, and C22 are most sensitive to alteration. Mutations at positions G13 and C14 show medium or small effects, and at T14, G17, and G18 any substitution is tolerated. Interestingly, single nucleotide mutations have also rather small effects in loop B at nucleotides 19–21, which are involved in base-pairing with the R-RNA substrate, apparently demonstrating that single mismatches at all but the first position of P4 are acceptable. For selected mutants, the 7S11 CoMA results have been independently confirmed by analyses of the ligation activity (as reported earlier[5b] and/or confirmed by our data[12]). Overall, these experiments demonstrate that CoMA provides comprehensive and reliable mutation data for the nucleotides in the catalytic region of 7S11.[14]

CoMA was then applied to investigate the 40 nt catalytic region of the 9DB1 deoxyribozyme that catalyzes the linear ligation of two RNA substrates. No mutagenesis data of any kind has been reported for 9DB1. According to the CoMA workflow outlined in Figure 2, four 9DB1 mutant libraries were synthesized using rN/dN mixtures at the 40 contiguous positions of the catalytic region. After separation of active and inactive deoxyribozyme mutants, the RNA–DNA hybrids were subjected to alkaline hydrolysis and the interference patterns were analyzed. The hydrolysis lanes of the active deoxyribozyme fractions, the interference values for transitions and transversions, and a schematic summary of the results are depicted in Figure 4.[13] The 2′-OH tag was tolerated at almost all 40 positions, with only three nucleotide positions showing OH-interference values slightly larger than 2. Strikingly, three consecutive guanosine nucleotides, G17–G19, in loop A were essential for catalytic activity (in red), reminiscent of the functionally important triple guanosine motif in loop A of 7S11. In addition, C13, G14, and A16 were highly sensitive to mutations. Nucleotides G15 and G20 showed small interference effects, and T21–T23 could be changed to any other nucleotide (green). Another exciting result was the clustering of green positions, that is, interference values less than 2, in stem II and loop B. This finding indicated that all mutations are allowed between positions 30 and 38 and suggested that this stem–loop is dispensable. In contrast, most of the nucleotides involved in stem I showed large interference effects, suggesting that mismatches are not tolerated in this stem. The predicted loop-closing wobble base-pair T12:G24 is most likely not formed because T12 can be changed to any other nucleotide but G24 is essential. Moreover, several nucleotides in the 5′-single-stranded part of the 9DB1 core could not be mutated without strongly affecting ligation activity. In particular, G1, A3, C5, and T7 seem to be involved in formation of the active conformation and/or participate in catalysis.

To validate the 9DB1 CoMA data, we synthesized several 9DB1 mutants and analyzed the DNA-catalyzed ligation rate in the trimolecular format (i.e., R-RNA is not covalently attached to the 9DB1 derivative). As expected from the interference data, stem II and loop B could be removed without considerably reducing the ligation rate (Figure 5). Mutations that showed high interference values in the combinatorial libraries were also detrimental when assayed in individual 9DB1 mutants. Tested examples include the G18A and G24A mutants that were about 800-fold and 1600-fold slower than 9DB1-mini (Figure 5). Further investigation of the predicted stem I by testing compensatory mutations revealed pronounced sensitivity to stem length and base-pair identity. Stabilization of the stem by changing A8:T28 to C8:G28 resulted in a 150-fold slower reaction. Changing individual base pairs resulted in 20–800-fold slower ligation

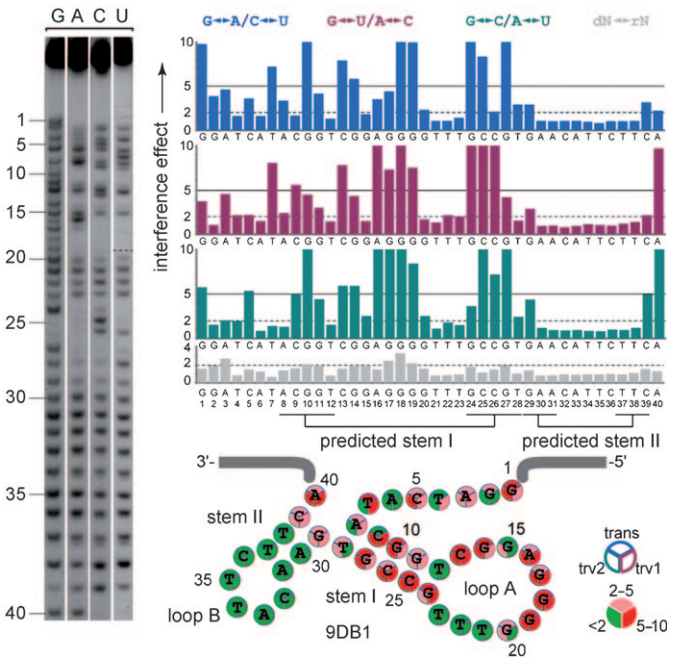

Figure 4. CoMA of 9DB1. The gel lanes depict the hydrolysis pattern of the catalytically active fractions of 9DB1 mutant libraries. The full gel including control lanes is shown in the Supporting Information. Interference values and secondary-structure representation follow the same scheme as that used in Figure 3.

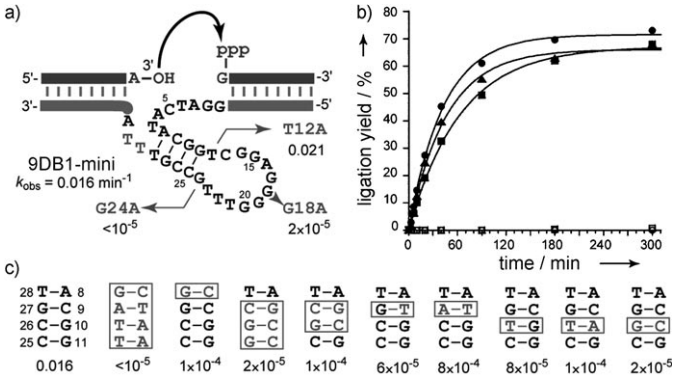

Figure 5. Analysis of the ligation activity of selected 9DB1 deoxyribozyme mutants. a) Minimized 9DB1 deoxyribozyme based on CoMA results. Selected loop mutations and k_{obs} values are indicated. b) Kinetic plots for original 9DB1 (●, k_{obs} = 0.023 min^{-1}), minimized 9DB1 (■), and loop mutations T12A (▲), G18A (□), and G24A (○). c) Selection of examined 9DB1 mini stem-mutants. Original stem base-pairs are in black, mutations in gray. k_{obs} values (min^{-1}) are shown below each stem.

© 2010 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

rates (Figure 5 and the Supporting Information). These results suggest that the helix orientation and the hydrogen-bonding donor/acceptor pattern in the grooves are important for interactions with other nucleotides.

The kinetic analyses confirmed that CoMA successfully identified core nucleotides that define the deoxyribozyme active site and are essential for catalytic activity. For both DNA enzymes studied, CoMA revealed a series of three consecutive guanosine nucleotides that are crucial for DNA-catalyzed RNA ligation of a ribose hydroxy group to a 5′-triphosphate. A comparison with other known deoxyribozymes that catalyze the formation of 2′,5′-branched nucleic acids[15] suggests that at least two consecutive guanosines might be generally needed to assist DNA-catalyzed RNA ligation of the 2′- or 3′-hydroxy groups to 5′-triphosphates. Based on the presented results for 7S11 and 9DB1, it seems likely that this requirement is independent of the branched or linear topology of the ligation product. More detailed investigations are certainly needed to identify the interaction partners of the functional nucleotides, which will lead to a mechanistic framework for DNA-catalyzed RNA ligation.

In this study, we have shown that combinatorial mutation interference analysis (CoMA) is a highly efficient method to identify catalytically essential nucleotides in deoxyribozymes. In addition, the CoMA results provide information on nonessential nucleotides and thereby facilitates minimization of the catalytic core regions. The ability to cleave mutant DNA libraries under alkaline conditions exclusively at 2′-OH-tagged positions makes it possible to map mutation-sensitive nucleotides. We found that the deoxyribose-to-ribose substitution is functionally silent for the large majority of nucleotide positions in the RNA-ligating deoxyribozymes investigated in this study. The application of solid-phase synthesis for the preparation of deoxyribozyme libraries enables the assessment of all possible mutants in one set of experiments using four distinct libraries. This cannot be achieved by enzymatic methods using template-dependent polymerase enzymes, which can only incorporate Watson–Crick-complementary nucleotides. In addition, CoMA is unbiased by the need to choose a subset of specific mutants for investigation, in contrast to the conventional mutagenesis approach. This is an especially important practical advantage of CoMA for longer functional nucleotide regions that are characteristic of many DNA catalysts and aptamers. The application of the 2′-OH group as an effective chemical tag will also allow the analysis of individual functional-group contributions to DNA catalysis. This application will be conceptually related to nucleotide analogue interference mapping (NAIM), which has been developed for assaying ribozyme catalytic mechanisms.[16] NAIM is based on the enzymatic incorporation of phosphorothioate-tagged nucleotide analogues by in vitro transcription.[17] In contrast, our approach for deoxyribozymes can employ incorporation of ribonucleotide analogues by solid-phase synthesis. Comprehensive results from mutagenesis and modification analyses will guide more detailed mechanistic investigations of DNA catalysts and may inspire the rational design of minimal functional units for the construction of more complex DNA architectures.

Received: June 29, 2010
Published online: September 24, 2010

Keywords: catalytic DNA · deoxyribozymes · mutation analysis · RNA ligation · solid-phase synthesis

[1] R. R. Breaker, G. F. Joyce, *Chem. Biol.* **1994**, *1*, 223.
[2] a) G. M. Emilsson, R. R. Breaker, *Cell. Mol. Life Sci.* **2002**, *59*, 596; b) A. Jäschke, *Curr. Opin. Struct. Biol.* **2001**, *11*, 321; c) A. Peracchi, *ChemBioChem* **2005**, *6*, 1316; d) C. Höbartner, S. K. Silverman, *Biopolymers* **2007**, *87*, 279; e) K. Schlosser, Y. Li, *Chem. Biol.* **2009**, *16*, 311.
[3] a) S. K. Silverman, *Chem. Commun.* **2008**, 3467; b) Y. Lu, J. W. Liu, *Curr. Opin. Biotechnol.* **2006**, *17*, 580; c) D. A. Baum, S. K. Silverman, *Cell. Mol. Life Sci.* **2008**, *65*, 2156; d) I. Willner, B. Shlyahovsky, M. Zayats, B. Willner, *Chem. Soc. Rev.* **2008**, *37*, 1153; e) K. Schlosser, Y. Li, *ChemBioChem* **2010**, *11*, 866.
[4] a) S. K. Silverman, *Acc. Chem. Res.* **2009**, *42*, 1521; b) S. K. Silverman, D. A. Baum, *Methods Enzymol.* **2009**, *469*, 95.
[5] a) R. L. Coppins, S. K. Silverman, *Nat. Struct. Mol. Biol.* **2004**, *11*, 270; b) R. L. Coppins, S. K. Silverman, *J. Am. Chem. Soc.* **2005**, *127*, 2900.
[6] R. A. Padgett, M. M. Konarska, P. J. Grabowski, S. F. Hardy, P. A. Sharp, *Science* **1984**, *225*, 898.
[7] a) C. B. Reese, Q. Song, *Nucleic Acids Res.* **1999**, *27*, 2672; b) D. Mitra, M. J. Damha, *J. Org. Chem.* **2007**, *72*, 9491.
[8] W. E. Purtha, R. L. Coppins, M. K. Smalley, S. K. Silverman, *J. Am. Chem. Soc.* **2005**, *127*, 13124.
[9] J. Nowakowski, P. J. Shim, G. S. Prasad, C. D. Stout, G. F. Joyce, *Nat. Struct. Biol.* **1999**, *6*, 151.
[10] M. Zuker, *Nucleic Acids Res.* **2003**, *31*, 3406.
[11] a) A. Peracchi, M. Bonaccio, M. Clerici, *J. Mol. Biol.* **2005**, *352*, 783 ; b) Z. Zaborowska, J. P. Furste, V. A. Erdmann, J. Kurreck, *J. Biol. Chem.* **2002**, *277*, 40617; c) Z. Zaborowska, S. Schubert, J. Kurreck, V. A. Erdmann, *FEBS Lett.* **2005**, *579*, 554.
[12] See the Supporting Information for details.
[13] The full gel with hydrolyzed unseparated and inactive fractions as well as non-hydrolyzed control lanes is shown in the Supporting Information.
[14] CoMA uses the four standard RNA nucleotides, including rU. We did not observe interference effects of rU at parent dT positions. In cases where rU interference is observed, it will be necessary to disentangle the effects of the 5-methyl and the 2′-OH groups, which can be achieved by using rT instead of rU for library synthesis.
[15] a) E. Zelin, Y. Wang, S. K. Silverman, *Biochemistry* **2006**, *45*, 2767; b) E. Zelin, S. K. Silverman, *ChemBioChem* **2007**, *8*, 1907; c) T. P. Mui, S. K. Silverman, *Org. Lett.* **2008**, *10*, 4417.
[16] S. A. Strobel, K. Shetty, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 2903.
[17] J. C. Cochrane, S. A. Strobel, *Curr. Protoc. Nucleic Acid Chem.* **2004**, Unit 6.9.

 © 2010 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim