Catalytic DNA (deoxyribozymes) for synthetic applications—current abilities and future prospects

Scott K. Silverman*

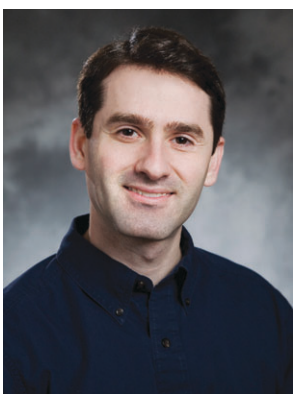
Received (in Cambridge, UK) 29th April 2008, Accepted 22nd May 2008
First published as an Advance Article on the web 1st July 2008
DOI: 10.1039/b807292m

The discovery of naturally occurring catalytic RNA (RNA enzymes, or ribozymes) in the 1980s immediately revised the view of RNA as a passive messenger that solely carries information from DNA to proteins. Because DNA and RNA differ only by the absence or presence of a 2′-hydroxyl group on each ribose ring of the polymer, the question of 'catalytic DNA?' arises. Although no natural DNA catalysts have been reported, since 1994 many artificial DNA enzymes, or 'deoxyribozymes', have been described. Deoxyribozymes offer insight into the mechanisms of natural and artificial ribozymes. DNA enzymes are also used as tools for *in vitro* and *in vivo* biochemistry, and they are key components of analytical sensors. This review focuses primarily on catalytic DNA for synthetic applications. Broadly defined, deoxyribozymes may have the greatest potential for catalyzing reactions in which the high selectivities of 'enzymes' are advantageous relative to traditional small-molecule catalysts. Although the scope of DNA-catalyzed synthesis is currently limited in most cases to oligonucleotide substrates, recent efforts have began to expand this frontier in promising new directions.

Introduction

The term 'catalytic DNA' may induce some bewilderment. Isn't DNA, which nature evolved for long-term storage of genetic information, the antithesis of chemical reactivity? How can DNA actually *do* anything? This review describes what is known about DNA as a catalyst and indicates likely future research directions for the field. The emphasis is on catalytic DNA for synthetic applications. Many recent reviews have focused on catalytic DNA for other interesting applications in biochemistry, biology, analytical chemistry, and computation.[1–5]

Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, USA. E-mail: scott@scs.uiuc.edu; Fax: +1 217 244 8024; Tel: +1 217 244 4489



Scott K. Silverman was born and raised in Los Angeles, California and received his BS degree in chemistry from UCLA in 1991. He was an NSF and ACS Organic Chemistry predoctoral fellow with Dennis Dougherty at Caltech, graduating with a PhD in chemistry in 1997. After postdoctoral research as a Helen Hay Whitney Foundation and American Cancer Society fellow with Thomas Cech at the University of Colorado at Boulder, he joined the University of Illinois at Urbana-Champaign in 2000, where he is currently Associate Professor of Chemistry, Biochemistry, and Biophysics. His laboratory studies nucleic acid structure, folding, and catalysis, including the identification and applications of deoxyribozymes.

How can DNA be a catalyst?

To both scientists and non-scientists, DNA is most easily recognized as the famous 'double helix' of Watson and Crick. In the form of a double helix (duplex), DNA adopts a relatively inflexible and regular structure that is well suited as the genetic material. The double helix and other higher-order forms of DNA are also useful as building blocks for intricate DNA nanostructures,[6–10] as the basis for nanomechanical devices,[11] and as rigid conformational control elements in macromolecular systems.[12] However, a DNA duplex is not a particularly good starting point from which to construct a catalyst, which requires a less uniform structure. (Consider that in proteins, highly regular structures form hair and fingernails but not enzymes.) When thinking about catalysts made from nucleic acids, both natural[13] and artificial[14] catalytic RNA molecules (*i.e.*, RNA enzymes, or ribozymes) offer substantial insights. Ribozymes have largely single-stranded sequence elements interspersed with double-stranded portions. Moreover, nucleic acid aptamers—which are DNA or RNA sequences that bind to target compounds, often with nanomolar or better affinity[15,16]—adopt largely single-stranded conformations.[17] These considerations suggest that DNA catalysis will require substantial single-stranded regions, whereas wholly duplex DNA is unlikely to have substantial catalytic activity.

In vitro selection to identify catalytic DNA

Over millions of years, nature has used natural selection to evolve ribozymes that participate in fundamental reactions of

modern biochemistry, such as protein synthesis in the ribosome[18] and (almost certainly) RNA splicing in the spliceosome.[19–21] Ribozymes likely played important roles in the evolution of life itself, *via* the 'RNA World'.[22] Within a few years of the discovery of natural ribozymes in the early 1980s,[23,24] artificial ribozymes were identified in the laboratory using *in vitro* selection.[25,26] This is a process in which a very large number of sequences are examined in parallel to identify those with desired catalytic activities. Although natural DNA catalysts have not been reported and may not even exist (an interesting question that is beyond the scope of this review), the general *in vitro* selection approaches used for RNA can be applied very successfully for DNA.

In vitro selection approaches are required because we do not know nearly enough to design from first principles a specific DNA sequence that can catalyze a desired chemical reaction. The same difficulty applies for ribozymes and protein enzymes. The first steps towards rational protein enzyme design have only recently been reported, either entirely *de novo*[27] or by using known proteins as scaffolds onto which computationally designed active sites are grafted.[28,29] Knowledge about catalytic DNA needs to advance considerably further for the rational design of catalytic DNA to be feasible. In the meantime, *in vitro* selection is the only practical way to identify new catalytically active DNA sequences.

Description of the *in vitro* selection process

Catalytic DNA sequences are variously termed deoxyribozymes, DNA enzymes, or DNAzymes. Here, the *in vitro* selection process is described for the identification of RNA-cleaving deoxyribozymes (Fig. 1), which were the first to be reported[30] and as a class are currently the most often studied and applied.[1,5] With appropriate technical modifications, related procedures can be used to identify catalytic DNA for other reactions. In a later section of this review are presented chemical structures and discussions of RNA cleavage and many other DNA-catalyzed reactions of potential synthetic utility.

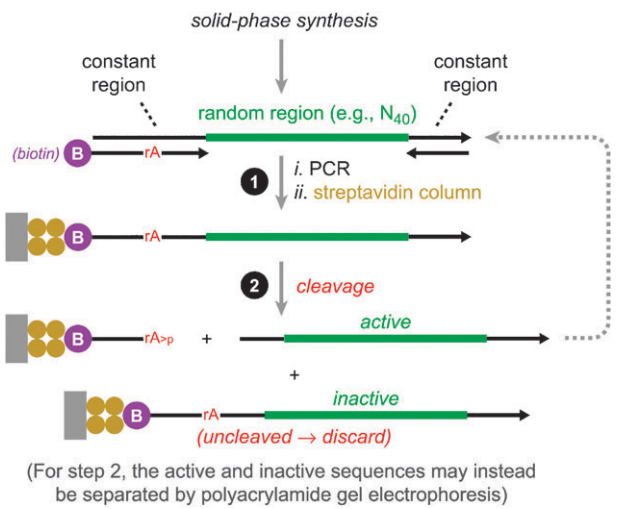


Fig. 1 *In vitro* selection strategy to identify deoxyribozymes that cleave RNA. RNA cleavage is a representative chemical reaction that can be catalyzed readily by DNA (see later section for chemical structures).

Initiating the selection process

In vitro selection generally requires solid-phase (automated) synthesis of a 'random pool' of DNA, which is a DNA oligonucleotide sample whose central portion has well-defined length but random nucleotide composition. Typical sizes for the random region length are 40–80 nucleotides. The randomness is achieved during solid-phase synthesis simply by mixing the four standard DNA phosphoramidite building blocks, one of which becomes randomly coupled at each relevant position of every growing oligonucleotide chain. Because all coupling events are independent and because sequence context does not generally affect the chemical coupling efficiency, using a mixture of the four phosphoramidites leads to a random collection of oligonucleotide sequences. The concentrations of the phosphoramidites in the mixture are adjusted according to the known relative reactivities, T > G > C > A (where T is about 1.5 times as reactive as A), so that each of the four monomers has an equal chance of being incorporated during any particular coupling event. In the resulting pool the random portion of DNA sequence is flanked by fixed regions, which are required for a subsequent polymerase chain reaction (PCR) amplification step.

With the random DNA pool available, the selection process is initiated. For finding RNA cleavage activity, this requires attachment to the DNA of an oligonucleotide bearing one or more potentially cleavable RNA linkages (Fig. 1, step 1). This can be performed as illustrated by PCR using a DNA primer that has an embedded ribonucleotide; alternatively, an enzymatic ligation can be performed. In either case, the product is a pool of candidate DNA enzyme sequences, each attached to an oligonucleotide with one or more RNA sites that will have the opportunity to be cleaved by that particular DNA sequence.

The key selection step

In the key selection step (Fig. 1, step 2), the mixture of candidate DNA sequences is incubated under suitable reaction conditions such as 50 mM Tris, pH 7.5; a wide range of conditions are possible. Key variables include the concentrations and identities of divalent metal ions (*e.g.*, 40 mM $MgCl_2$) and monovalent metal ions (*e.g.*, 150 mM NaCl), as well as the temperature (*e.g.*, 37 °C) and the incubation time (from as long as tens of hours, to as short as only a few seconds). Although the vast majority of the random-pool DNA sequences are catalytically inactive, a small number of the DNA molecules will successfully catalyze the desired RNA cleavage reaction during the allotted time. The selection strategy is devised so that these catalytically active DNA sequences will become physically separable as a direct consequence of their catalysis. For RNA cleavage, this can be achieved either by removing a biotin moiety attached at the 5′-terminus of the ribonucleotide-bearing portion, which allows the active DNA sequences to flow through a streptavidin column (as illustrated in step 2), or by inducing a substantial

electrophoretic shift on a polyacrylamide gel such that the active DNA sequences can be isolated.

Iterating selection rounds

After isolating the active DNA sequences, PCR amplification is again performed to begin the next selection round. This process is iterated for multiple rounds until the catalytic activity of the pool as a whole is sufficiently high that individual DNA sequences can be identified ('cloned') and examined in more detail. Iteration is required to enrich the pool in reproducibly active sequences. For various catalytic activities typically 5–15 selection rounds are required, although there is no absolute rule.

Increasing the selection pressure

One option during every selection experiment is to 'increase the selection pressure'; *i.e.*, to change the incubation conditions during the key selection step such that the desired reaction becomes more difficult. Doing so increases the likelihood that the surviving DNA sequences can catalyze the desired reaction even more efficiently. Common ways to increase the selection pressure are to decrease the amount of time allotted or to decrease the concentration of divalent metal ion(s) included in the sample. The former change selects for intrinsically faster deoxyribozymes, whereas the latter change selects for DNA that has a lower divalent metal ion concentration requirement. Artificial nucleic acid catalysts generally require divalent metal ions for efficient reactivity, although exceptions are known (see further discussion of this topic below).

Terminating the selection process

Judgment of when to terminate the selection process is rather subjective. In general, a successful selection experiment will eventually provide some level of catalytic activity that does not increase with further iterations of the procedure, even when the selection pressure is held constant. If this level of activity is less than 100%, such as 10%, then this can mean one of two limiting situations applies: (1) all of the DNA sequences in the enriched pool have 10% yield under the incubation conditions; or (2) about 10% of the DNA sequences have 100% yield, and the remaining $\sim$90% of the sequences have only trace yield (*e.g.*, $<$1%). Any linear combination of these two possibilities could also correspond to the actual situation; *e.g.*, 50% of the sequences have 10% yield, 5% have 100% yield, and 45% have $<$1% yield. In all cases, the investigator must decide whether the observed level of activity after any given round is sufficiently promising to warrant cloning of individual deoxyribozymes to assess their activities.

Cloning individual deoxyribozymes

The cloning process is generally straightforward and involves inserting the PCR product after the desired selection round into a standard plasmid vector followed by screening of individual deoxyribozymes for catalytic competence. Those DNA enzymes with high activities are examined by automated sequencing and then usually prepared independently by solid-phase synthesis to verify that the catalytic activity is reproducible. The final deoxyribozymes are subsequently characterized biochemically in more detail.

Quantitative considerations for the selection procedure

In practice, the number of possible sequences in even a modest length of a DNA oligonucleotide quickly rises beyond what can actually be explored in practice. The concept of 'sequence space' describes the collection of all possible sequences of a certain length. For a random region of 40 nucleotides (*i.e.*, N_{40}), the sequence space encompasses $4^{40} \approx 10^{24}$ possible sequences. Because for practical reasons a selection experiment can begin with 'only' $\sim 10^{14}$ molecules, sequence space will be sampled only to the extent of 10^{-10} ($10^{14}/10^{24}$) for an N_{40} random region (Fig. 2). Nevertheless, many deoxyribozyme selection experiments have been quite successful using N_{40} or even much larger random regions. Because the *y*-axis of Fig. 2 covers so many orders of magnitude, starting with more molecules from the random pool (*e.g.*, by using a 10- or 100-fold larger sample volume) is relatively ineffective at exploring a meaningfully larger fraction of sequence space.

The random region length is limited only by what is accessible *via* solid-phase synthesis. Random regions smaller than N_{40} can certainly be made. However, at a sufficiently small length all of sequence space is readily sampled, and there is little reason to go smaller than this limit when searching for new deoxyribozyme activities. Because $4^{22} \approx 10^{13}$, a random region of N_{22} has tenfold *over*-sampling for a typical initial pool size of $\sim 10^{14}$. Therefore, at N_{22} one is exploring all possible DNA sequences in a single selection experiment. Going in the other direction, very large DNA random regions (up to N_{228}; ref. 31) have been prepared by stitching together as many as three discrete smaller random regions, which are separated by short fixed regions that allow PCR-based preparation of the larger random pool. The extent of sequence space coverage for an N_{228} region is truly miniscule. Because

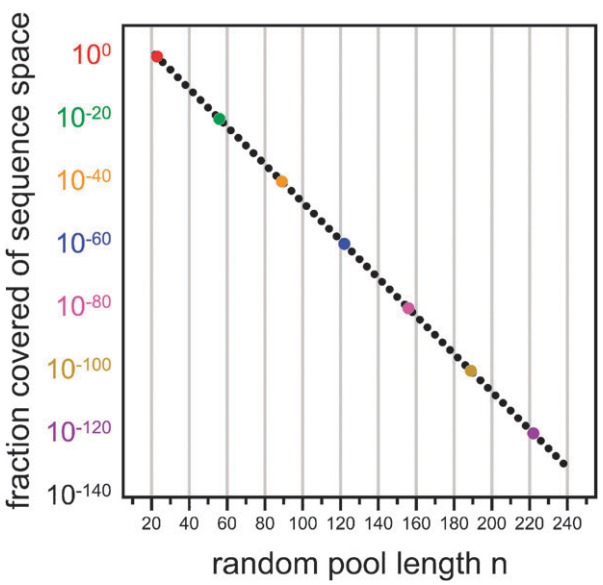


Fig. 2 Fractional coverage of sequence space as a function of random pool length, as computed for a typical 200 pmol initial random pool (10^{14} molecules). Fraction covered = $10^{14}/4^n$.

$4^{228} \approx 10^{137}$, an initial pool containing 10^{14} molecules can sample only about 10^{-123} (!) of the sequence space. Despite this mathematical nightmare, even N_{228} selection experiments have been successful. One mitigating factor is that many of the nucleotides in such large deoxyribozymes can be any of the four DNA bases or can even be deleted altogether without affecting catalysis. For every ten nucleotides with this property, $4^{10} = 10^6$ of the random sequences are equivalent in terms of catalytic ability.

The choice of the random region length in any particular deoxyribozyme selection experiment is largely subjective and empirical. As quantified above, a larger random region inherently means progressively less sampling of sequence space. On the other hand, a larger oligonucleotide can presumably adopt a more intricate three-dimensional structure that may permit more efficient catalysis of the desired reaction.[32] For certain chemical reactions, a larger random region may allow the catalysis to take place at all whereas a shorter oligonucleotide of any sequence may not be catalytically competent. It is difficult to state without detailed experimentation where the 'break point' of sequence length occurs for any particular reaction to be catalyzed.

'Reselection'—*in vitro* evolution to optimize an initially found DNA sequence

In some *in vitro* selection experiments, the starting point is not an entirely random pool of DNA sequences but rather a partially randomized, or 'biased', pool. This is often done to explore a more focused portion of sequence space near a known deoxyribozyme sequence that may itself not have optimal catalytic activity. For example, to identify a faster deoxyribozyme (*i.e.*, one with a higher rate constant), a known DNA enzyme sequence as identified by the general procedure described above can be partially randomized and subjected to 'reselection', with shorter incubation times providing the additional selection pressure. Because evolution requires both selection *and* the introduction of variation,[33] this overall process is deemed '*in vitro* evolution'. In contrast, a selection experiment in which variation is not introduced after the start of the experiment is simply '*in vitro* selection'.

The biased pool for an *in vitro* evolution experiment may be prepared by solid-phase synthesis using suitable mixtures of phosphoramidites. For example, at nucleotide positions where adenosine (A) is present in the original DNA sequence, a mixture of phosphoramidites favoring A but with perhaps 10% of each of the other three phosphoramidites (T, C, G) may be used. The probability calculations associated with such mixtures are straightforward and allow simple computation of the degree of randomization required to explore a particular extent of sequence space surrounding the original DNA sequence.[34] Alternatively, 'mutagenic PCR' can be used to introduce variation into a DNA sequence.[35] This approach has the advantage of not requiring solid-phase synthesis of a partially randomized oligonucleotide pool. Therefore, mutagenic PCR can be performed as part of an in-progress selection experiment without needing to clone individual functional DNA sequences, as is normally done at the end of a selection process. This is useful for many selection experiments, considering that the undersampling of sequence space (*e.g.*, 10^{-10} for an N_{40} region as described above) means that any initially identified deoxyribozymes are almost certainly not the optimal sequences.

Distinction between '*in vitro* selection' and 'combinatorial chemistry'—selection *vs.* screening

Using *in vitro* selection to search among many candidate DNA molecules to find the rare functionally active 'hit' may appear reminiscent of traditional small-molecule combinatorial chemistry, which is often performed as 'high-throughput screening' of individual compounds.[36] Nonetheless, *in vitro* selection and combinatorial chemistry are very different experimental approaches. The selection process does not examine each candidate sequence separately, as is inherently the case for combinatorial screens. Rather, in a selection experiment all of the contending molecules are surveyed simultaneously, and only those able to perform a predetermined challenge—in the example given above, cleave an RNA substrate—become physically separable and are declared as 'winning' sequences. This strategy allows a very large number of sequences to be examined efficiently in parallel; typically 10^{14} sequences are explored as described above.

In contrast, combinatorial chemistry screening experiments require the separate evaluation of each candidate compound, whether on microarrays, multi-well plates, or distinct beads. In practice, therefore, the number of different compounds that can be examined combinatorially is much smaller, such that a 10^3-member combinatorial library is already rather large. Of course, the compounds examined in combinatorial chemistry experiments may have quite varied structures depending on the nature of the combinatorial library, whereas *in vitro* selection of nucleic acids is inherently restricted to DNA or RNA. The potential trade-off in function, which in principle disfavors DNA relative to small-molecule catalysts, must be assessed by vigorously exploring the scope of DNA catalysis with more *in vitro* selection experiments.

DNA catalysis of potential synthetic interest—oligonucleotide substrates

Essentially all deoxyribozymes are identified using a multi-step selection process as described in Fig. 1 for RNA cleavage, with appropriate technical modifications depending on the particular catalytic activity being sought. In this section are discussed deoxyribozymes that catalyze reactions of oligonucleotide substrates. The reactions of potential synthetic interest for such substrates include cleavage and ligation reactions; reactions that modify the oligonucleotide termini (*e.g.*, phosphorylation); and reactions involving the nucleobases. Table 1 provides rate constants and rate enhancements for representative individual deoxyribozymes that catalyze the reactions discussed throughout this review.

RNA cleavage

The prototypical DNA-catalyzed reaction is the cleavage of an RNA oligonucleotide substrate, because this is the first reaction for which catalytic DNA was identified.[30] RNA cleavage

Table 1 Representative deoxyribozymes with their rate constants and rate enhancements

Reaction	k_{obs}/min^{-1}[a]	Rate enh.[b]	Ref.
RNA cleavage	1	$\sim 10^5$	30
	4	n.d.	40
	0.01	$\sim 10^8$	126
DNA cleavage (oxidative)	0.1	$\sim 10^6$	46–48
RNA ligation (3′–5′ & other)	0.1	2×10^4	59
RNA ligation (3′–5′)	0.06	$\sim 10^5$	64
RNA ligation (branch)	0.4	5×10^6	65,67
	1	$\sim 10^5$	69,70
	0.03	$\sim 10^5$	72
RNA ligation (lariat)	0.08	$\sim 10^5$	80
DNA phosphorylation	3	$\sim 10^9$	82,83
DNA adenylation (capping)	0.004	2×10^{10}	84
DNA ligation	0.06	3×10^3	86
	0.0001	$\sim 10^5$	85
DNA depurination	0.2	9×10^5	87
Thymine dimer photoreversion	5	3×10^4	89
Porphyrin metalation	1	1×10^3	31,92
Diels–Alder reaction	3	4×10^5	103
Nucleopeptide linkage formation	0.06	5×10^5	106

[a] To one significant figure, this is the maximal value of k_{obs} as reported in the indicated manuscript(s). [b] n.d. = not determined.

is catalyzed by all natural ribozymes except for the ribosome, which is the only natural catalytic RNA known to ligate rather than cleave two substrates as its biological role.[37] Cleavage of RNA is most readily catalyzed by facilitating attack of the 2′-hydroxyl group of a particular nucleotide into the adjacent phosphodiester linkage, resulting in 2′,3′-cyclic phosphate and 5′-hydroxyl termini (Fig. 3). This is the same cleavage mechanism that is promoted by ribonuclease A and several natural ribozymes including the hammerhead, hairpin, and HDV ribozymes; the larger group I intron and RNase P provide different products by a separate mechanism.[13] Uncatalyzed (spontaneous) RNA cleavage also occurs by the pathway of Fig. 3, and geometrical effects upon the reaction rate[38] can be used as a sensitive probe of RNA conformation *via* the technique of 'in-line probing'.[38,39] Nevertheless, uncatalyzed RNA cleavage is not synthetically useful due to insufficient selectivity, because essentially all RNA sites can be cleaved by this mechanism. This has led to the development of several deoxyribozymes that are useful for practical sequence-selective RNA cleavage *via* catalysis of the reaction of Fig. 3.

The first deoxyribozyme selection experiment as reported by Breaker and Joyce identified a Pb^{2+}-dependent deoxyribozyme that catalyzes cleavage at a ribonucleotide embedded within an otherwise all-DNA strand.[30] Subsequent efforts identified DNA enzymes that can cleave an all-RNA substrate. At present the most well-known DNA enzymes are the RNA-cleaving 10–23 and 8–17 deoxyribozymes, which were named

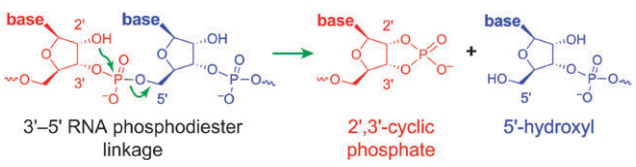


Fig. 3 RNA cleavage by attack of a 2′-hydroxyl group at a neighboring phosphodiester linkage. This reaction occurs spontaneously with relatively low rate and can also be catalyzed by DNA.

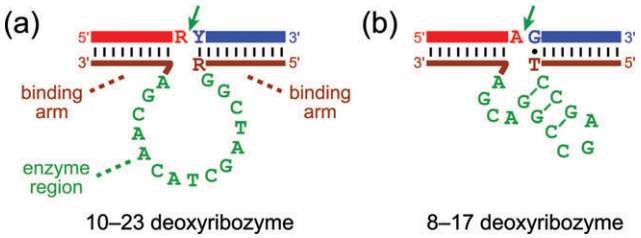


Fig. 4 RNA cleavage catalyzed by the (a) 10–23 and (b) 8–17 deoxyribozymes.[40] The 10–23 has cleavage site requirement 5′–R↓Y–3′ (R = purine A or G, Y = pyrimidine U or C), and the 8–17 has cleavage site requirement 5′–A↓G–3′. Other nearby nucleotides in the RNA substrate can have any identities, as long as the corresponding DNA nucleotides in the 'binding arms' are Watson–Crick complementary.

by Santoro and Joyce to denote the 23rd clone from round 10 and the 17th clone from round 8 of the particular selection experiment.[40] (Each laboratory uses its own nomenclature for its deoxyribozymes.) The 10–23 deoxyribozyme can cleave any RNA sequence that satisfies two simple requirements (Fig. 4): (1) the cleavage site must be a purine nucleotide (adenosine A or guanosine G, commonly written as R for 'puRine') followed by a pyrimidine nucleotide (uridine U or cytosine C, abbreviated as Y for 'pYrimidine'), or 5′-R↓Y-3′; and (2) a sufficient length of the remaining RNA nucleotides outside the cleavage site must be Watson–Crick base-paired with the deoxyribozyme 'binding arms'. The 8–17 deoxyribozyme has the same properties, except that the cleavage site requirement is more restrictive, 5′-A↓G-3′. These properties mean that the 10–23 and 8–17 DNA enzymes are useful synthetically for practical RNA cleavage, either as analytical- or preparative-scale tools.[1]

Later experiments showed that 8–17 is the simplest deoxyribozyme motif capable of RNA cleavage;[41] consistent with this, several independent selection experiments have resulted in 8–17-like structures.[40–44] A set of such deoxyribozymes as reported by Li and co-workers is collectively capable of cleaving nearly any (14 out of 16) RNA dinucleotide junctions, and these catalytic DNAs are therefore useful for the practical cleavage of essentially any RNA sequence at a predetermined site.[41] The same authors have recently reported additional deoxyribozymes that are useful in the same general fashion. Three optimal 8–17 variants can cleave 10 out of the 16 possible RNA dinucleotide junctions with rates 0.1 min^{-1} or higher, which are preparatively useful (k_{obs} of 0.1 min^{-1} corresponds to $t_{1/2}$ of merely 7 min).[45] Other reviews cover more comprehensively the general topic of DNA-catalyzed RNA cleavage.[1,5]

Oxidative DNA cleavage

Two classes of DNA catalyst have been identified for catalysis of oxidative DNA cleavage.[46] The 'class I' deoxyribozymes require both Cu^{2+} and ascorbate, the latter added as a reducing agent. The 'class II' deoxyribozymes require only Cu^{2+}, although they function 10^3-fold faster with ascorbate as well. The class II DNA enzymes were optimized to work with Cu^{2+} alone and engineered to function by cleaving a separate single-stranded DNA substrate.[47,48] Recognition of the

substrate strand by the deoxyribozyme appears to involve both duplex and triplex formation, and only a restricted set of substrate sequences can be cleaved. Moreover, because cleavage appears to involve a diffusable hydroxyl radical, the cleavage site is poorly defined (*i.e.*, not at the resolution of a single nucleotide), which limits the synthetic utility. Because DNA cannot be cleaved by a route analogous to that for RNA in Fig. 3 and is therefore ~400-fold less susceptible to cleavage,[49,50] identifying a site-selective hydrolytic DNA-cleaving deoxyribozyme remains as an unsolved and likely difficult problem.

RNA ligation—formation of linear RNA

Joining two RNA substrates is of particular synthetic interest because directly preparing large chemically modified RNAs is often challenging. The most common way to incorporate chemical modifications into large RNAs (*e.g.*, >100 nucleotides) is to prepare by solid-phase synthesis a short (<40 nt) oligoribonucleotide containing the desired modification(s), followed by one or more RNA ligation reactions using either T4 DNA ligase or T4 RNA ligase.[51–54] T4 DNA ligase requires a DNA 'splint' oligonucleotide to hold together the two RNA substrate strands, whereas T4 RNA ligase can function with or without a splint. Although this general approach has been widely used, it is idiosyncratic because predicting the suitability for any specific RNA ligation junction is difficult. With the additional consideration that preparing the RNA substrates necessary to test any particular ligation reaction can be time-consuming and expensive, alternative strategies for RNA ligation are welcome. As explored by Silverman and co-workers in a series of studies, deoxyribozymes offer hope in this regard.

For maximal practical utility, deoxyribozymes that ligate RNA should use combinations of RNA substrate functional groups that are readily obtained and do not require extensive organic synthesis, which is generally avoided by most biochemists. One such combination is the joining of a 2′,3′-cyclic phosphate with a 5′-hydroxyl group, in a reaction that is the reverse of the RNA cleavage event depicted in Fig. 3. The cyclic phosphate is readily obtained by cleavage of a precursor RNA using a deoxyribozyme such as 10–23 or 8–17, or by incorporating a self-cleaving ribozyme sequence directly into the RNA.[55,56]

In the synthetic (*i.e.*, ligation) direction for opening of the 2′,3′-cyclic phosphate by attack of a 5′-hydroxyl, the issue of regioselectivity immediately arises. Which oxygen of the cyclic phosphate, 2′ or 3′, acts as the leaving group? If the 2′-oxygen departs, then the resulting linkage is a native 3′–5′ phosphodiester bond (Fig. 5, path i). In contrast, if the 3′-oxygen departs, then the product has a non-native 2′–5′ linkage (Fig. 5, path ii). For many applications, the formation of a native 3′–5′ linkage is much more desirable, because this will leave no sign that a ligation event had ever taken place—the ligation will be 'traceless'.

Several selection experiments to identify DNA-catalyzed RNA ligation using a 2′,3′-cyclic phosphate substrate have been performed by Silverman and co-workers, leading to the discovery of several major challenges. The first study of a

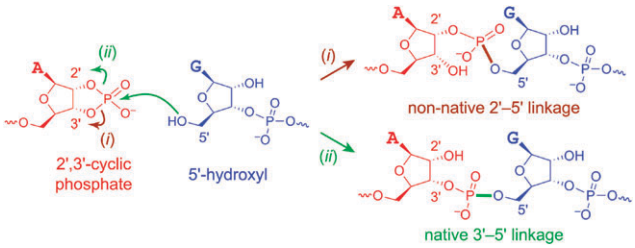


Fig. 5 Two possible RNA ligation reactions with 2′,3′-cyclic phosphate and 5′-hydroxyl termini, leading to either native 3′–5′ or non-native 2′–5′ linkages.

DNA enzyme that ligates RNA reported catalysts that create non-native 2′–5′ linkages using Mg^{2+} as the metal ion cofactor.[57] Subsequent efforts revealed that DNA-catalyzed formation of 2′–5′ linkages is strongly favored in many different contexts.[34,58] In contrast, when analogous selections were performed using Zn^{2+} as the cofactor, both 3′–5′ and 2′–5′ linkages were formed by different deoxyribozymes (as were several unnatural linkages arising from nucleophilic attack of 2′-hydroxyl groups located internally on the same substrate strand as the free 5′-hydroxyl group).[59] Taken together, these data indicate that we lack a fundamental understanding of what controls the regioselectivity in the DNA-catalyzed opening of the 2′,3′-cyclic phosphate functional group. In addition, none of the subset of Zn^{2+}-dependent deoxyribozymes that create native 3′–5′ linkages have sufficiently broad substrate sequence tolerance to be useful for practical RNA ligation.

A second useful combination of RNA functional groups is the joining of a 2′,3′-diol with a 5′-triphosphate. The latter is readily introduced at the 5′-end of an RNA strand by *in vitro* transcription using T7 RNA polymerase. (Unfortunately, appending a 5′-triphosphate onto RNA at the conclusion of solid-phase synthesis[60] is in practice rather tricky, although 5′-adenylated RNA[61] can sometimes be a functional surrogate.) In the DNA-catalyzed ligation reaction involving a 5′-triphosphate RNA substrate, the issue of site-selectivity must be addressed. Which hydroxyl group of the 2′,3′-diol substrate, 2′ or 3′, acts as the nucleophile to attack the 5′-triphosphate? As is the case for the cyclic phosphate substrate combination, both native 3′–5′ and non-native 2′–5′ linkages can be formed (Fig. 6, paths i and ii).

In some selections with the 5′-triphosphate RNA substrate, neither the terminal 2′-hydroxyl group nor the 3′-hydroxyl group of the second RNA substrate served as the nucleophile in deference to an internal 2′-hydroxyl group, with formation

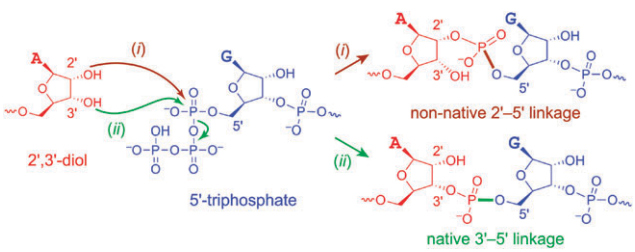


Fig. 6 Two possible RNA ligation reactions with 2′,3′-diol and 5′-triphosphate termini, leading to either native 3′–5′ or non-native 2′–5′ linkages.

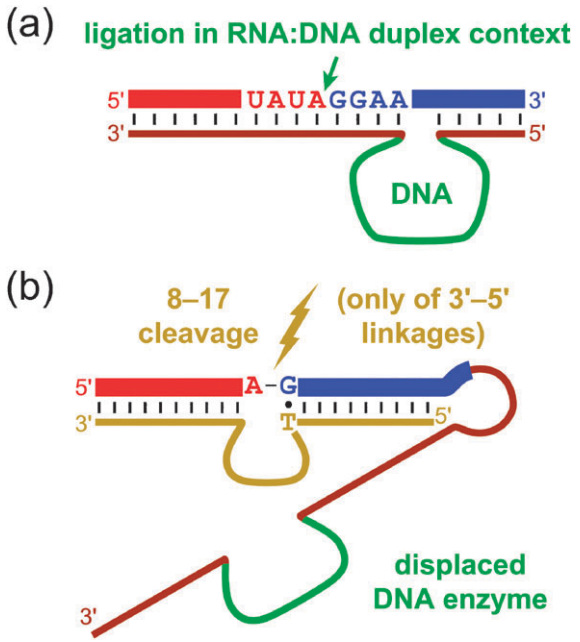


Fig. 7 Two selection strategies to favor the linear 3′–5′-linked RNA product. (a) Situating the ligation site (3′-hydroxyl + 5′-triphosphate) within the context of an RNA:DNA duplex.[62] Although this approach successfully created the desired native 3′–5′ linkages, numerous unwanted sequence requirements were induced within the two RNA substrates. (b) Adding a selection step to cleave newly formed 3′–5′ linkages with the highly selective 8–17 deoxyribozyme when the DNA sequence that catalyzed the RNA ligation is still attached.[63]

of branched RNA (these reactions are discussed in the next section). Two general approaches were developed to enforce the reactivity of the terminal 3′-hydroxyl group, leading to the desired native 3′–5′ linkages. First, the ligation junction was situated within an RNA:DNA duplex architecture (Fig. 7a), which was anticipated to favor the linear ligation product due to spatial constraint of the two reactive groups.[62] Moreover, this architecture was anticipated specifically to favor 3′–5′ rather than 2′–5′ linkages, because the former are more stable in the context of an RNA:DNA duplex and this stability should be felt (at least partially) in the transition state for ligation. The deoxyribozymes that emerged from this first approach indeed create only native 3′–5′ linkages. However, there are numerous sequence requirements for the two RNA substrates near the ligation junction, presumably because the deoxyribozyme interacts directly with certain RNA functional groups on the nucleobases, and this limits the general applicability of these DNA enzymes for RNA ligation.

Second, a selection pressure was devised that stringently favors formation of only linear 3′–5′ linkages, regardless of the architecture of the deoxyribozyme and its RNA substrates (Fig. 7b).[63] After the key step 2 of Fig. 1, in which DNA sequences capable of RNA ligation were separated by polyacrylamide gel electrophoresis (PAGE) without regard to which particular linkage was formed, a new step was introduced in which the RNA-cleaving 8–17 deoxyribozyme was applied. The 8–17 is highly selective for cleaving 3′–5′ linkages. Therefore, only the newly joined RNA strands that are connected *via* a 3′–5′ bond are cleaved by the 8–17, which

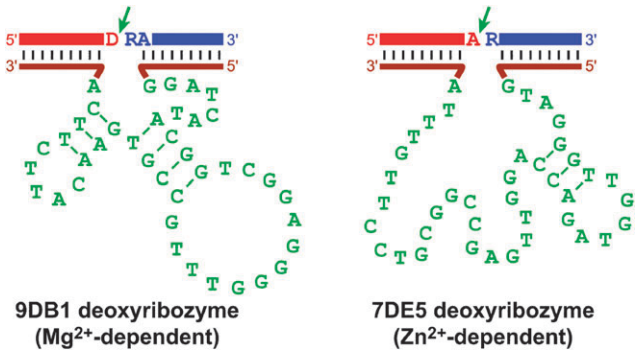


Fig. 8 The Mg^{2+}-dependent 9DB1 and Zn^{2+}-dependent 7DE5 deoxyribozymes. These DNA catalysts create native 3′–5′ RNA linkages from 3′-hydroxyl and 5′-triphosphate groups with useful rate, yield, and substrate sequence generality.[64]

physically displaces the DNA sequence that is attached to the RNA. Deoxyribozymes that specifically create 3′–5′ linkages are thus separable by an additional PAGE step, this time selecting for a decrease (rather than increase) in size. This approach successfully led to 3′–5′ linkages yet placed no demands on the architecture of the interaction between the deoxyribozyme and the two RNA substrates.

Finally, the lessons learned from the above studies were integrated to obtain deoxyribozymes that simultaneously meet several important requirements for practical RNA-ligating DNA catalysts.[64] Using either Mg^{2+} or (separately) Zn^{2+} as the divalent metal ion cofactor, two deoxyribozymes were found that ligate RNA rapidly and in high yield with the formation of native 3′–5′ linkages and useful tolerance of various RNA substrate sequences (Fig. 8). The Mg^{2+}-dependent 9DB1 deoxyribozyme requires the ligation site D↓RA, where D denotes any nucleotide except C, and the Zn^{2+}-dependent 7DE5 deoxyribozyme requires only A↓R. Ongoing efforts are seeking to complete a set of DNA enzymes that collectively can be used for ligation of nearly any combination of RNA substrates, much like the set of 8–17 variants described above that can cleave almost any RNA linkage.

RNA ligation—formation of branched RNA

As mentioned briefly above, the first study from the Silverman laboratory on DNA enzymes for ligation of 5′-triphosphate RNA reported that internal 2′-hydroxyl groups preferentially act as the nucleophiles to attack the 5′-triphosphate (Fig. 9).[65] The products are 2′,5′-branched RNAs, which are created naturally during RNA splicing.[66] This observation established that DNA enzymes have the potential for synthesis of biochemically useful RNAs, because branched RNAs are challenging to prepare by other means. Another notable aspect of this finding is the very high site-selectivity inherent in catalyzing the reaction of only one 2′-hydroxyl group, although many other chemically equivalent 2′-hydroxyl groups are present.

This initial study was followed by several additional efforts to elucidate important aspects of branched RNA formation by deoxyribozymes. A full description is beyond the scope of this review, but a summary of the key findings is provided. Characterization of the initial set of branch-forming deoxyribozymes showed that numerous sequence requirements

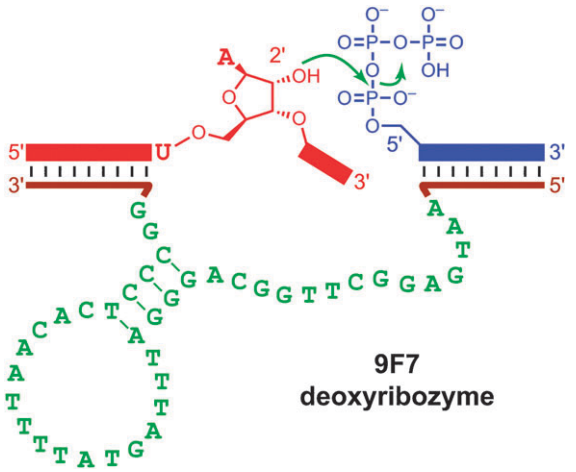


Fig. 9 Formation of 2′,5′-branched RNA by reaction of an internal 2′-hydroxyl group with a 5′-triphosphate, catalyzed by the 9F7 deoxyribozyme.

existed for the two RNA substrates, indicating that the DNA enzymes would not be useful for practical branch formation.[67] Soon thereafter, a different deoxyribozyme named 7S11 was identified that creates RNA by forming a three-helix-junction (3HJ) architecture between the DNA enzyme and the two RNA substrate strands (Fig. 10). Because numerous natural ribozymes have multi-helix junctions,[68] 7S11 offers the opportunity of a model system for understanding reactivity in this structural framework.

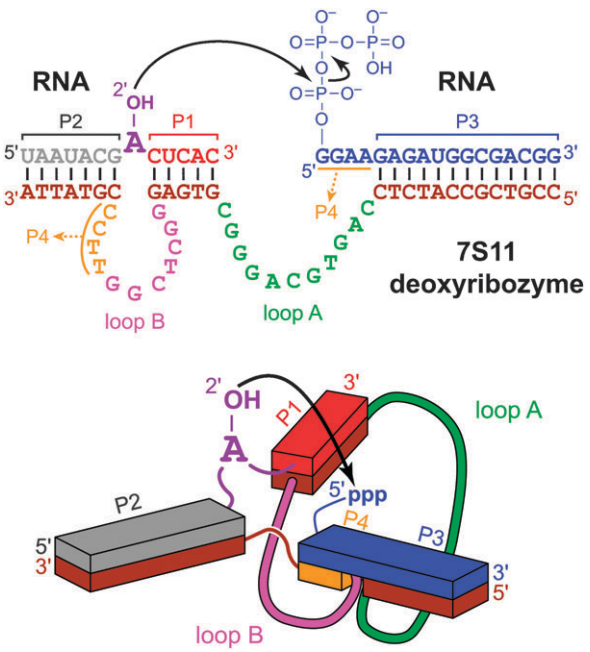


Fig. 10 The 7S11 deoxyribozyme, which has a three-helix-junction (3HJ) architecture among the DNA enzyme and its two RNA substrates.[69,70] This arrangement provides preorganization for the nucleophile (2′-hydroxyl) and electrophile (5′-triphosphate, ppp) *via* formation of the four paired regions denoted P1–P4. The detailed three-dimensional structure of the deoxyribozyme, including the two loops A and B, is not yet known.

The 7S11 deoxyribozyme has been a platform for exploring numerous aspects of DNA catalysis. Because many of the Watson–Crick base pairs in the 3HJ structure can have any identity as long as pairing is maintained, 7S11 is rather general for the synthesis of branched RNAs of many different sequences.[70] This generality enabled a biochemical study of RNA proofreading pathways;[71] a different non-multi-helix-junction deoxyribozyme[72] was used to create branched RNAs for studies of Ty1 retrotransposition.[73] In other work, the 3HJ architecture was used to demonstrate empirically that branch-site adenosine is chemically favored during branch formation, which is the same preference shown by natural splicing processes.[74] The atomic-level explanation for the adenosine preference is not yet known. This same study identified 10DM24, a variant of 7S11 that has improved generality for RNA branch formation.

An interesting synthetic application of DNA enzymes that form branched RNA is deoxyribozyme-catalyzed labeling (DECAL).[75] This approach treats one RNA substrate as the 'target' RNA to be labeled, and the second RNA substrate as a 'tag' to be attached by a deoxyribozyme such as 10DM24 at a predefined 2′-position of the target (Fig. 11). When the 'tagging' RNA has a biophysical label (*e.g.*, fluorescein or biotin) present at its second nucleotide, the DECAL strategy effectively labels the target RNA at a specific site. In the initial report, Baum and Silverman used DECAL to append both fluorescein and tetramethylrhodamine as a commonly used pair of chromophores to the independently folding 160-nucleotide P4–P6 RNA domain, thereby enabling the first fluorescence resonance energy transfer (FRET) experiment for this otherwise well-studied RNA.[75] In general, one concern is that the appended tagging strand could potentially disrupt RNA structure or function. However, control experiments showed that this did not occur for P4–P6 in the initial application of DECAL.

A second notable use of branch-forming deoxyribozymes is their application to attach DNA constraint strands to RNA

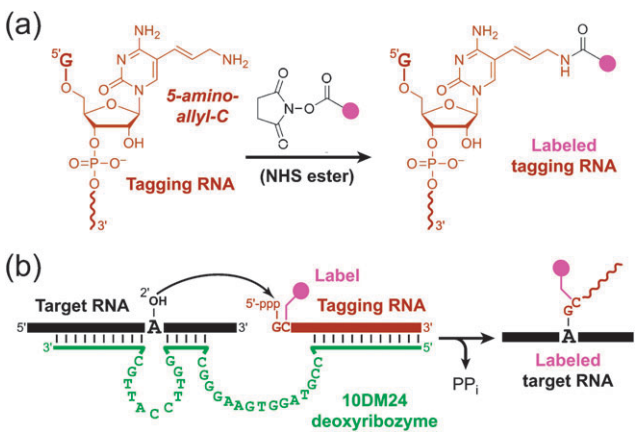


Fig. 11 Deoxyribozyme-catalyzed labeling (DECAL) of RNA.[75] (a) Generation of the labeled 'tagging' RNA by reaction of an aminoallyl-modified nucleotide with the *N*-hydroxysuccinimide (NHS) ester of a biophysical label (*e.g.*, fluorescein or biotin). The aminoallyl-modified nucleotide is introduced site-specifically into the tagging RNA by *in vitro* transcription. (b) Attachment of the tagging RNA to the target RNA by the 10DM24 deoxyribozyme.

macromolecules.[76] In these experiments, the 5′-triphosphorylated RNA was replaced with activated 5′-adenylated DNA,[77] in which the leaving group upon 2′-hydroxyl attack is adenosine 5′-monophosphate (AMP) rather than pyrophosphate (PP_i). The resulting RNA–DNA branches allow control over RNA conformation and catalysis by establishment of double-stranded DNA constraints that are incompatible with the native RNA structure.[12,78,79]

RNA ligation—formation of lariat RNA

Lariat RNAs are a subclass of 2′,5′-branched RNAs in which two of the oligonucleotide arms emerging from the branch site are connected to form a closed loop. The branched RNAs formed during biological splicing are actually lariat RNAs, which provide an even greater challenge due to their particular topology (*e.g.*, lariats cannot readily be prepared by direct solid-phase synthesis). Deoxyribozymes have been used in synthetic approaches specifically to form lariat RNAs, including the synthesis of biologically derived lariat sequences.[80,81] Shown in Fig. 12 are the sequence and possible secondary structure of one deoxyribozyme, 6BX22, that is especially proficient at creating lariat RNAs, including the ability to tolerate any nucleotide at the branch site.[80] Using 6BX22, lariats that have 266 nucleotides in the loop (1597 atoms) were created with extremely high site-selectivity, because the deoxyribozyme only uses one 2′-hydroxyl group as the nucleophile to attack the 5′-triphosphate.[80]

DNA phosphorylation, adenylation (capping), and ligation

Breaker, Li and co-workers have described deoxyribozymes for DNA-catalyzed self-phosphorylation,[82,83] self-adenylation (capping),[84] and ligation[85] (Fig. 13a). Because of the importance of these reactions for biotechnology (*e.g.*, in DNA cloning), the identification of these deoxyribozymes suggests that DNA-based alternatives can be found for the protein enzymes T4 polynucleotide kinase (for DNA phosphorylation) and T4 DNA ligase (for adenylation and subsequent ligation). Although the likely sequence requirements of the currently known DNA enzymes do not allow general practical utility in the manner of the available protein enzymes, further development may be useful in this

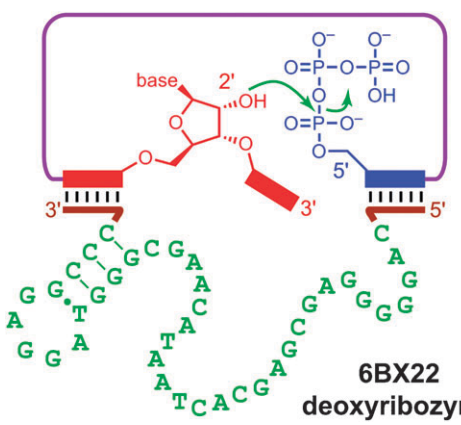


Fig. 12 Lariat RNA formation catalyzed by the 6BX22 deoxyribozyme.[80] Here the two reacting functional groups (2′-hydroxyl and 5′-triphosphate) are part of the same substrate molecule, leading to the closed loop that is characteristic of lariat RNAs.

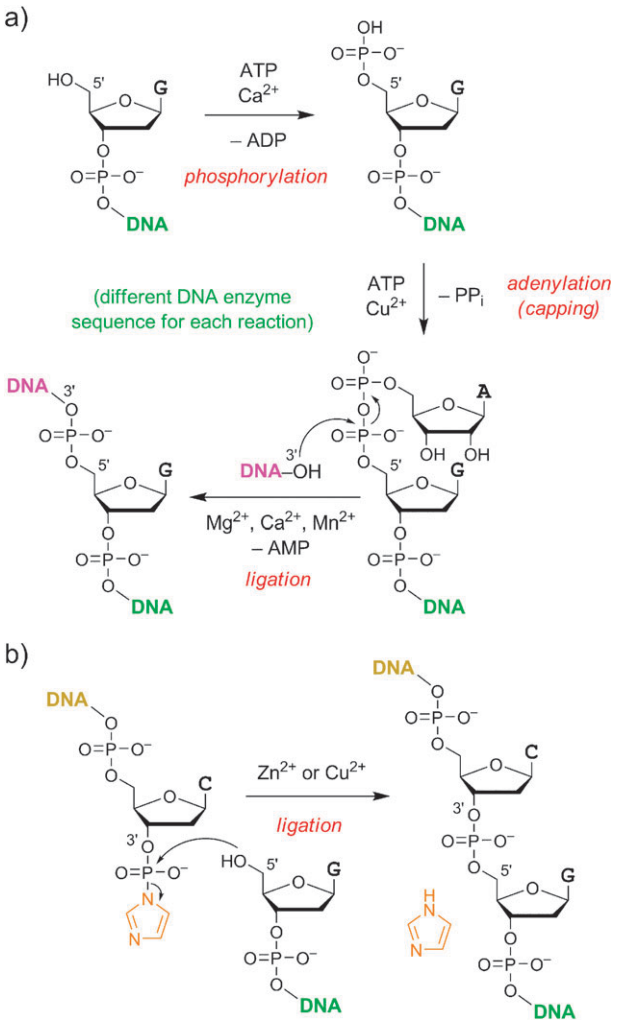


Fig. 13 DNA phosphorylation, adenylation (capping), and ligation catalyzed by DNA. (a) All three reactions as performed by a series of distinct deoxyribozymes under different reaction conditions.[82–85] (b) DNA-catalyzed DNA ligation using an activated 3′-phosphorimidazolide substrate.[86]

regard. DNA ligation was also reported by Cuenoud and Szostak using an activated 3′-phosphorimidazolide substrate that reacts with a 5′-hydroxyl group (Fig. 13b) rather than a 5′-adenylated substrate that reacts with a 3′-hydroxyl group.[86] However, the sequence tolerance of this deoxyribozyme was not reported and is likely to be low.

Deglycosylation (depurination)

Two different deoxyribozymes have been reported for catalysis of site-specific DNA deglycosylation, which is a reaction that breaks a C–N rather than an O–P bond (as in all of the reactions described to this point). First, Joyce and co-workers identified the 10–28 DNA enzyme that catalyzes the removal of the guanine nucleobase from a particular internal DNA position.[87] The apurinic (AP) site revealed after depurination subsequently undergoes β-elimination, leading to strand scission (Fig. 14a). This reaction requires a divalent metal ion such as Ca^{2+} and proceeds with multiple turnover upon strategic separation of the enzyme and substrate strands, but the sequence dependence of

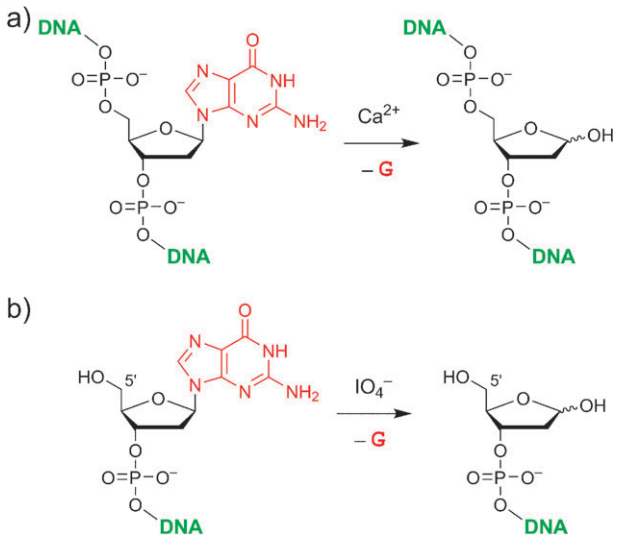


Fig. 14 Self-deglycosylation (depurination) reactions catalyzed by DNA. (a) Deglycosylation of a specific internal guanosine by the 10–28 deoxyribozyme.[87] The reaction requires a divalent metal ion (2 mM Ca^{2+} is optimal). (b) Deglycosylation of the 5′-terminal guanosine by the 10FN10 deoxyribozyme.[88] The reaction needs no divalent metal ion but requires 40 mM IO_4^- (periodate) as a cofactor.

the reaction was not reported. The authors suggest that further development of such DNA catalysts could be useful to induce site-specific deglycosylation (and subsequent repair) of particular sequences within genomic DNA, including at positions of oxidative damage or other modified bases. However, such applications are impractical at present.

Second, Silverman and co-workers identified the 10FN10 deoxyribozyme that catalyzes loss of the guanine nucleobase from its own 5′-terminal nucleotide.[88] The reaction is independent of any divalent metal ions but requires 40 mM $NaIO_4$ (sodium periodate) as an obligatory cofactor (Fig. 14b). A plausible mechanistic hypothesis is that periodate oxidizes the 5′-terminal G1 nucleobase, which then spontaneously depurinates. Similar to the 10–28 deoxyribozyme, the chemical reactivity of the 10FN10 DNA enzyme could in principle be useful with further development.

Thymine dimer photocleavage

Finally, a particularly intriguing DNA catalytic activity is thymine dimer photoreversion (Fig. 15).[89,90] The UV1C deoxyribozyme reported by Chinnapen and Sen does not require any metal ion or small-molecule cofactor and uses light of λ_{max} 305 nm to promote the photochemical reaction. A guanine quadruplex (G-quadruplex) within the DNA enzyme appears to act as an 'antenna' for subsequent electron transfer to the thymine dimer substrate. Because detailed experiments to evaluate the substrate nucleotide requirements surrounding the thymine dimer have not been reported, the generality of the UV1C deoxyribozyme is unknown for cleaving thymine dimers in various sequence contexts. One implication of these studies is that DNA has the overall ability to catalyze photochemical reactions, and it seems likely that other examples will be developed. In addition, cofactor-*dependent* photochemical processes catalyzed by DNA (which were actually the original

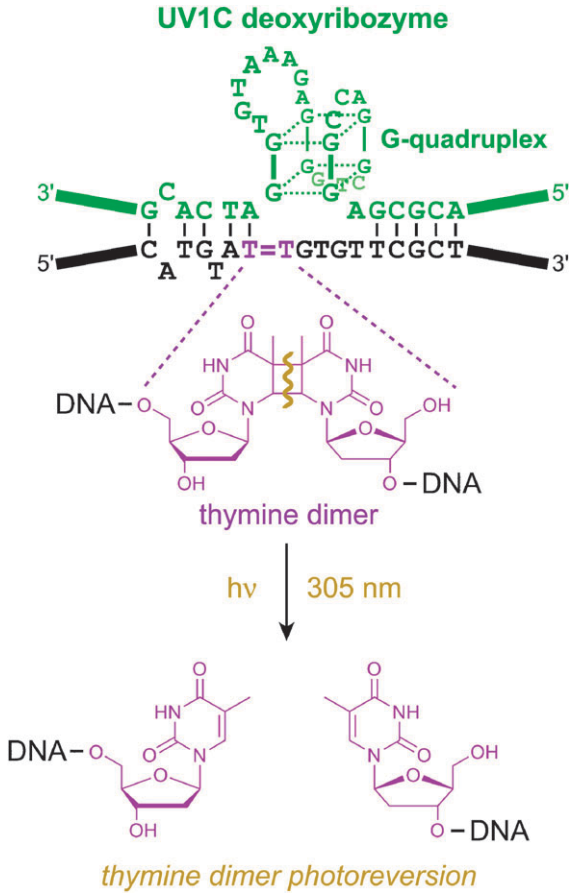


Fig. 15 Thymine dimer photoreversion catalyzed by the UV1C deoxyribozyme.[89,90] The two-tiered G-quadruplex of UV1C acts as the 'antenna' for light of λ_{max} 305 nm; no cofactors are required.

goal in the selection experiment that led to UV1C) remain as viable reaction targets.

Summary of synthetic applications of deoxyribozymes for oligonucleotide substrates

Of all of the deoxyribozymes discussed in this section, those used most frequently for practical synthetic applications catalyze RNA cleavage or RNA ligation. For RNA cleavage by the 10–23 or 8–17 deoxyribozymes (or variants of the latter), the concise substrate sequence requirements combined with a lack of other site-selective methods for accomplishing RNA cleavage have focused increasing attention on the DNA-catalyzed approach. For all other DNA-catalyzed reactions of oligonucleotide substrates, including RNA ligation, the practical need for synthesizing the reaction products and the availability (or lack thereof) of alternative preparative methods will drive any further development of useful deoxyribozymes.

DNA catalysis of potential synthetic interest—substrates that are not oligonucleotides

For many chemists, 'synthesis' implies small-molecule substrates, and using DNA to cleave or ligate RNA—or to catalyze any of the other reactions described in the previous section—does not qualify.[91] To address this small-molecule

chemistry audience, deoxyribozymes are being developed for substrates that are not oligonucleotides. These efforts address the scope and potential limitations of DNA catalysis and help to define the future directions in which research should proceed. Despite recent advances, the gold standard of catalyzing small-molecule chemistry is still largely unachieved for deoxyribozymes. This remains an important frontier for future experiments. A primary challenge for obtaining catalytic DNA that functions with non-oligonucleotide substrates is to replace the built-in binding energy that is supplied by a series of Watson–Crick base pairs between a DNA enzyme and an oligonucleotide substrate, as described for all reactions in the previous section. The first three examples described below all involve small-molecule substrates for deoxyribozymes, and the final example describes the first exploration of DNA-catalyzed amino acid sidechain reactivity.

Porphyrins as small-molecule substrates

Li and Sen described a guanine-rich DNA sequence that catalyzes the metalation of mesoporphyrin IX by insertion of Cu^{2+} or Zn^{2+} (Fig. 16a).[31] Several related porphyrins can be used as a substrate, and the folded structure of the DNA enzyme (which requires K^+ for activity) likely involves G-quartets.[92,93] Due to its potential for intercalation, a large and flat porphyrin may be an especially suitable non-oligonucleotide substrate for a deoxyribozyme. A notable feature of the selection effort that led to the porphyrin-metalating deoxyribozyme is that the structurally distorted *N*-methylmesoporphyrin was used as a transition state analogue (TSA) for the metalation reaction (Fig. 16b). This was done rather than the general approach of Fig. 1, for which an implementation is not obvious in this case. In general, the TSA approach has not been particularly successful for identifying ribozymes,[94] and the same is likely true for deoxyribozymes.

NTPs as small-molecule substrates

Several deoxyribozymes have been identified that use NTPs as small-molecule substrates. The adenylating and phosphorylating DNA enzymes described above (Fig. 13) are in this category; these deoxyribozymes use one oligonucleotide substrate and a second ATP substrate. However, no general lesson for how to use small-molecule substrates can be obtained from

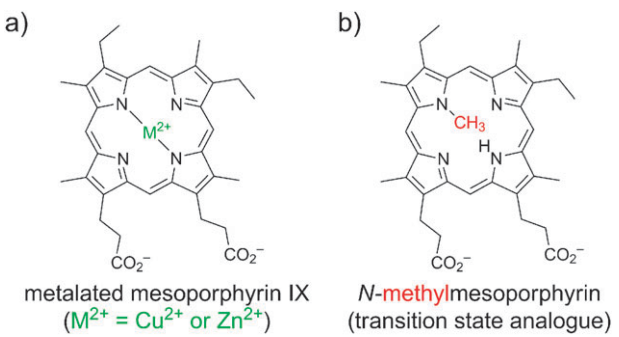


Fig. 16 DNA-catalyzed metalation of a porphyrin substrate.[31] (a) The product of the metalation reaction using mesoporphyrin IX as the small-molecule substrate. (b) *N*-Methylmesoporphyrin, which was used as a transition-state analogue during the selection process.

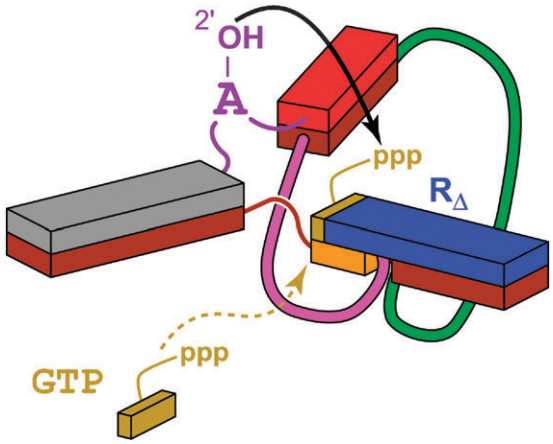


Fig. 17 Rationally engineered version of the 10DM24 deoxyribozyme that uses an NTP rather than a 5′-triphosphate-RNA as a substrate (compare with Fig. 10).[95] Disconnecting the 5′-terminal G nucleoside along with its 5′-triphosphate from the remainder of the RNA substrate allows GTP to be used as the small-molecule electrophile, along with the shortened oligoribonucleotide R_Δ as an obligatory cofactor.

those studies, other than the observation that an NTP binding site can be obtained as part of a functional deoxyribozyme when starting from an initially random region.

Höbartner and Silverman sought a more directed approach towards engineering an NTP binding site into a deoxyribozyme (Fig. 17).[95] Using the 3HJ architecture of the 10DM24 DNA enzyme as a starting point, the 5′-terminal nucleotide that provides the 5′-triphosphate was separated from the remainder of the RNA. Remarkably, 10DM24 was still catalytically active, now with GTP as a small-molecule substrate (along with the remainder of the original RNA oligonucleotide as an obligatory cofactor). Moreover, when the DNA nucleotide that pairs with GTP was changed from C to T, the selectivity of the deoxyribozyme was switched entirely from GTP to ATP. The efficiency of the catalysis was tunable by adjusting the DNA:NTP interaction energy. For example, although ATP was a relatively poor substrate, the analogous 2,6-diaminopurine ribonucleoside triphosphate ('DTP') that makes three rather than two hydrogen bonds with the T nucleotide of the deoxyribozyme was a much better substrate. These findings offer considerable hope for the rational engineering of deoxyribozyme active sites to function with different small-molecule substrates.

This approach will likely succeed only for small-molecule compounds that resemble NTPs. As a more generalizable strategy, one longer-term need is to integrate aptamers with catalysis by including pre-formed aptamer regions along with random enzyme regions at the outset of selection. The history of this approach for ribozymes is limited and mixed, with reports of both success[96] and failure[97] in terms of the designed aptamer ultimately becoming involved in catalysis. Many experiments are needed to determine the viability of the aptamer-based approach for DNA catalysis.

C–C bond formation—the Diels–Alder reaction

The range of ribozyme-catalyzed 'organic' reactions is rather limited.[14,15] As reported by Eaton and co-workers, the first

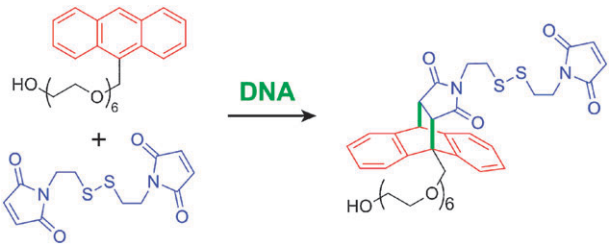


Fig. 18 The intermolecular Diels–Alder reaction between anthracene and maleimide substrates, as catalyzed by the 66-nucleotide DAB22 deoxyribozyme.[103]

C–C bond-forming ribozyme catalyzes a Diels–Alder reaction and requires pyridine-modified uridine monomers.[98] Seelig and Jäschke subsequently described a Diels–Alder ribozyme that uses only the four standard RNA nucleotides.[99] The latter ribozyme has been studied biochemically,[100,101] and its crystal structure has been obtained.[102] Inspired by these experiments, Chandra and Silverman reported a Diels–Alder deoxyribozyme as the first DNA catalyst that creates C–C bonds (Fig. 18).[103] An important motivation for this study was to assess the relative efficiencies of DNA and RNA catalysts for the same chemical reaction; this comparison is discussed more completely below. A deoxyribozyme that catalyzes the Diels–Alder reaction between anthracene and maleimide substrates 'in trans' (*i.e.*, intermolecularly) was found. The identification of this multiple-turnover DNA enzyme establishes that DNA has the capacity to catalyze bimolecular small-molecule reactions of potential synthetic importance. Although the enantioselectivity of the Diels–Alder deoxyribozyme has not yet been investigated in detail, preliminary data indicate a high enantiomeric ratio (M. Chandra and S.K.S., unpublished results), similar to the Diels–Alder ribozyme studied by Jäschke and co-workers.[100,104]

Jäschke's original Diels–Alder ribozyme was subsequently immobilized on an agarose solid support for investigations towards practical synthesis, although only submilligram quantities of substrates were reported.[105] Such experiments likely point the way to analogous use of deoxyribozymes for practical synthetic reactions, presuming that the necessary catalytic DNA sequences can be identified.

DNA-catalyzed reactions of amino acid sidechains

Finally, Silverman and co-workers reported a deoxyribozyme that joins a tyrosine sidechain to an RNA strand *via* a nucleopeptide linkage (Fig. 19).[106] This was achieved by placing the tyrosine at the intersection of a three-helix junction and performing the selection for nucleophilic attack of the tyrosine hydroxyl group onto a 5′-triphosphate-RNA. This work established that DNA can catalyze reactions of amino acid sidechains, although reactions of serine and lysine sidechains were not observed. Three challenges for future studies in this area are (1) to decrease the preorganization inherent to the 3HJ architecture while retaining catalytic activity; (2) to obtain reactivity of the sidechains of Ser and Lys as well as Tyr; and (3) to use a full protein rather than a small peptide as the substrate.

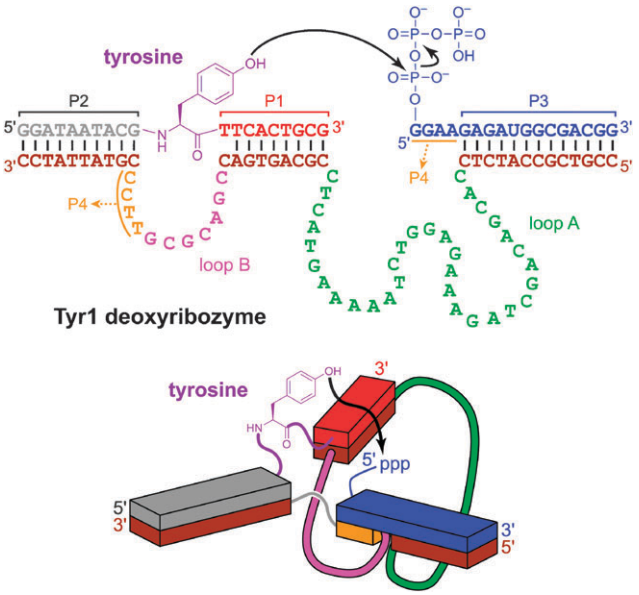


Fig. 19 Formation of a nucleopeptide linkage between a tyrosine sidechain and a 5′-triphosphate-RNA, as catalyzed by the Tyr1 deoxyribozyme.

The need to explore more DNA-catalyzed reactions

The relative brevity of this section on non-oligonucleotide substrates is evidence that the field of DNA catalysis is quite young. Some important general questions emerge when considering where to focus future research in the area. What reactions need catalysis in the first place? When might DNA be particularly suitable as a catalyst? For what reactions do we specifically need 'enzymes', rather than more traditional small-molecule catalysts? Addressing all of these questions will be important as DNA catalysts are increasingly applied to substrates that are not oligonucleotides.

DNA *vs.* RNA—do 2′-hydroxyl groups matter for catalysis?

When compared with RNA as a catalyst, DNA has two practical advantages. First, DNA is generally less expensive than RNA to prepare by solid-phase synthesis, with a cost differential of approximately one order of magnitude. This is related to the expense and ease of synthesizing the nucleotide monomers; the efficiency of the iterated solid-phase coupling reactions; and the simplicity of deprotection after synthesis, all of which favor DNA. Second, DNA is more stable than RNA, both chemically and biochemically, because the intramolecular RNA cleavage reaction of Fig. 3 cannot occur for DNA. Avoiding ubiquitous ribonucleases requires sufficient care when working with RNA both *in vitro* and *in vivo*, whereas these concerns are lower for DNA (although deoxyribonucleases of course exist as well).

Do these practical advantages of DNA come with any catalytic drawbacks due to the absence of the 2′-hydroxyl groups found in RNA? If so, then the wisdom of focusing on catalytic DNA would be unclear. Early speculations were that DNA should be less catalytically active than RNA,[107,108] if not

altogether incompetent. Since 1994 we have recognized that the absence of known natural DNA catalysts does not mean that DNA cannot actually catalyze reactions, but perhaps there is a catalytic disadvantage relative to RNA? One way to address this issue is to compare deoxyribozyme rate enhancement values to the analogous range of values for ribozymes. As tabulated elsewhere, DNA enzyme rate enhancements are in the same range as for various RNA enzymes,[14,15] which suggests that DNA has no general catalytic disadvantage relative to RNA. In this context it should also be noted that DNA aptamers appear to bind as tightly as RNA aptamers to a range of targets,[15–17] which is also consistent with no general functional difference between DNA and RNA.

An even better way to assess the catalytic prowess of DNA is to identify new DNA and RNA catalysts for the same (or very similar) reactions and assess whether there is any consistent difference between DNA and RNA. For reactions of oligonucleotide substrates, such as RNA ligation, artificial deoxyribozymes and ribozymes both generally have rate enhancements in the 10^5 to 10^6 range.[14] For other catalyzed reactions, the data is currently rather limited. The identification of the Diels–Alder deoxyribozyme as described above was motivated by precisely this question.[103] The results indicate that DNA and RNA can have comparable rate enhancements—both on the order of 10^5 or higher—for this particular C–C bond-forming reaction.

In summary, despite some speculations to the contrary, experimental evidence fails to validate that DNA is any less competent than RNA at catalyzing chemical reactions. In contrast, the available evidence indicates that DNA and RNA are comparable in catalytic ability, even though DNA lacks 2′-hydroxyl groups. Therefore, the most reasonable conclusion at present is that the practical advantages of DNA for catalysis are not mitigated by any catalytic disadvantages, and on this basis increased attention to DNA catalysis is warranted.

How do deoxyribozymes achieve catalysis?

Given our current knowledge, understanding how DNA catalyzes chemical reactions is a fertile area for future studies. Little is known mechanistically about how any deoxyribozymes actually achieve their rate enhancements. As usual, we can turn to ribozymes and protein enzymes for guidance. Catalysis generally involves some combination of directly lowering the transition-state energy—commonly termed 'chemical catalysis'[109]—and precisely positioning (aligning) the reactive functional groups of the substrates.[110] Importantly, the phenomenon of templating, *i.e.*, effective molarity, is markedly insufficient to explain the observed catalysis by deoxyribozymes, as discussed further in the section below on DNA-templated synthesis.

Ribozyme mechanisms and implications for DNA catalysis—the need for structural biology data

Compelling evidence indicates that one of the largest natural ribozymes, the ribosome, is an 'entropic catalyst' that creates peptide bonds primarily by carefully positioning its substrates (roles may also be played by exclusion of water from the active site and electrostatic stabilization of the transition state).[37,109,111] In contrast, the small natural ribozymes, all of which cleave RNA phosphodiester linkages, rely upon chemical catalysis. Although divalent metal ions were originally thought to be required by these ribozymes,[112] many experiments have revealed that nucleobase functional groups are sufficient.[113] The mechanisms of artificial ribozymes are less well understood, largely because less high-resolution structural information has been obtained. Among the few cases where such information is available, both chemical and entropic catalysis have been invoked.[102,114] Unlike the small natural ribozymes, numerous artificial ribozymes have a strict requirement for divalent metal ions such as Mg^{2+} (ref. 114). However, the detailed roles of such metal ions in folding or catalysis (or both) are not always known.

In contrast to the situation with proteins or even ribozymes, no high-resolution structural data exist for deoxyribozymes to provide a framework for more detailed biochemical studies. The 10–23 deoxyribozyme has been crystallized, but the crystals contain a 2:2 substrate:deoxyribozyme complex and the structure does not correspond to a catalytically relevant conformation.[115] No other X-ray data and no NMR data at all have been reported for deoxyribozymes. Presumably this overall lack of high-resolution structural information for catalytic DNA will be remedied at some point, which in turn will enable more detailed mechanistic investigations. In the meantime, although metal ion requirements and kinetic parameters have been studied in certain cases (*e.g.*, ref. 116), the general mechanistic view of how deoxyribozymes catalyze chemical reactions is largely incomplete.

One deoxyribozyme, one product—as expected for 'enzymes'

For nearly all deoxyribozymes, a consistent observation is that any particular DNA enzyme sequence creates only one reaction product, even when multiple products are chemically possible. For example, the various RNA-cleaving deoxyribozymes cut their RNA target at only one specific phosphodiester linkage, even though every other linkage has a 2′-hydroxyl group and could potentially be cleaved by the common mechanism of Fig. 3. Similarly, the deoxyribozymes that create branched or lariat RNA selectively use only one particular 2′-hydroxyl group as the nucleophile, even though many chemically equivalent groups are present (*e.g.*, Fig. 9–11). Templating alone should not obey this 'one deoxyribozyme, one product' rule, whereas enzymatic catalysis *via* some combination of transition-state stabilization and substrate alignment should readily do so.

Two known exceptions to the rule are readily explained. First, branched RNA formation can occur by reaction of a single 2′-hydroxyl nucleophile at either of two different electrophilic sites, when each of the latter sites is suboptimal relative to the originally selected activity.[117] This was observed when one electrophilic site had suboptimal reactivity due to a poorer leaving group, and the other site had suboptimal spatial positioning relative to the nucleophile. Second, Cu^{2+}-dependent oxidative DNA cleavage catalyzed by DNA is not

site-selective,[46–48] but this is due to the diffusible nature of the hydroxyl radical that is the active cleaving agent.

Intrinsic limitation on catalysis by nucleic acids?

Because proteins have twenty chemically varied amino acid side chains whereas nucleic acids have only four similar nucleobases, one may speculate that nucleic acids should be less catalytically capable than proteins. However, contrary to this viewpoint, Breaker and co-workers have argued that ribozymes and deoxyribozymes should have catalytic abilities rivaling those of protein enzymes, if a suitable combination of mechanistic strategies is employed.[118,119] Additional experiments to explore the wider functional abilities of nucleic acid enzymes are needed to resolve this issue.

Multiple turnover

Strictly speaking (although colloquial usage may be different), the phenomenon of 'catalysis' does not demand multiple turnover. Of the known deoxyribozymes, some can achieve multiple turnover and some cannot. For example, the 7S11 and 10DM24 deoxyribozymes that create branched RNA (Fig. 10) are only single-turnover, because the ligated product binds more tightly to the DNA than do the unligated substrates. In contrast, when 10DM24 is used with the mononucleotide GTP as the electrophilic substrate in place of 5′-triphosphate-RNA (Fig. 17), then multiple turnover is readily observed, because the substrates and product have approximately equivalent binding energies with the DNA. The Diels–Alder deoxyribozyme of Fig. 18 also achieves multiple turnover.

Three general comments are warranted on the topic of multiple turnover as related to catalytic DNA. (1) The correlation between relative binding strength (DNA:substrate *vs.* DNA:product) and turnover is sensible and is observed for protein enzymes as well. For example, T4 DNA ligase is used stoichiometrically (*i.e.*, without turnover) to join two RNA substrates *via* 'splint ligation' due to product inhibition.[120] (2) For many practical synthetic applications of deoxyribozymes, such as any instances of RNA cleavage and ligation, the DNA is not the most expensive component. In these cases, achieving multiple turnover is not particularly important as a practical matter. (3) The nature of the *in vitro* selection process (*e.g.*, Fig. 1) makes it impossible to select directly for multiple turnover due to the required physical linkage between information and catalysis. The only obvious way around this problem is to spatially segregate each potential catalyst sequence; *e.g.*, by *in vitro* compartmentalization (IVC).[121,122] This has been successful for a small number of ribozymes[123–125] and could presumably be achieved for deoxyribozymes, if a compelling need warrants the experimental effort.

Improving DNA catalysis by using non-DNA components

One way for DNA to use more catalytic strategies is to increase the variety of functional groups that are available to participate directly in 'chemical catalysis'. Because unmodified DNA starts with relatively few functional groups, many options can expand this repertoire.

Metal ions as deoxyribozyme cofactors

The most obvious non-DNA component to include is metal ions. Of course, DNA as a polyanion inherently binds to metal ions, but charge neutralization can be achieved with monovalent ions (*e.g.*, Na^+, K^+) that are not likely to promote catalytic activity. Nearly all deoxyribozymes have been selected to require one or more divalent metal ions such as Mg^{2+}, Ca^{2+}, or Mn^{2+} for efficient catalysis, although some exceptions have been described.[126,127]

Small-molecule compounds as deoxyribozyme cofactors

Less obvious than metal ions as deoxyribozyme cofactors are small-molecule compounds. Two observations suggest that low-molecular-weight cofactors are plausible participants in DNA catalysis. First, the natural *glmS* riboswitch-ribozyme uses glucosamine 6-phosphate (GlcN6P) not only as an allosteric regulator[128] but also as an obligatory catalytic cofactor (coenzyme), with the amine moiety of GlcN6P likely engaging in proton transfer during catalysis.[129] Second, both RNA and DNA are highly competent as aptamers for binding to small-molecule ligands, as described above.[15–17] Two cofactor-dependent deoxyribozymes have been described: an RNA-cleaving DNA enzyme that requires histidine,[130] and the deoxyribozyme that catalyzes site-selective depurination in the presence of periodate (Fig. 14b).[88] It seems likely that future studies will use other small molecules as cofactors for new DNA catalysts.

Unnatural nucleotides to improve DNA catalysis

Incorporating unnatural nucleotides directly into nucleic acids can readily provide nonstandard functional groups, such as imidazole and aliphatic amines (Fig. 20). This has been achieved in several cases for ribozymes[131] as well as for deoxyribozymes. In all such cases, the unnatural nucleotides must be incorporated randomly *via* a polymerase that tolerates the chemical modification. Because site-specific modification is impossible by this approach (this requires direct solid-phase synthesis of the modified DNA, whereas the selection process requires PCR during each round), typically one of the four standard DNA nucleotides is entirely replaced by a chemically modified variant. A thermostable DNA polymerase must be able to accept the corresponding modified deoxynucleoside triphosphate, or alternatively a dNTP bearing a functional group 'handle' such as an azide or alkyne that can be derivatized after the DNA synthesis is performed.[132,133] Chemical modifications are particularly well tolerated by suitable polymerases at the C7 position of the 7-deazaadenine nucleobase[134] and the C5 position of the deoxyuridine nucleobase[135] (Fig. 20). In one striking example, commercially available DNA polymerases were used to synthesize DNA for which all four nucleobases were chemically modified.[136,137] Although such thoroughly modified DNA has not yet been reported in the context of selection experiments, RNA-cleaving deoxyribozymes that make use of 'extended chemical functionality' on one or two of the four nucleotides have been

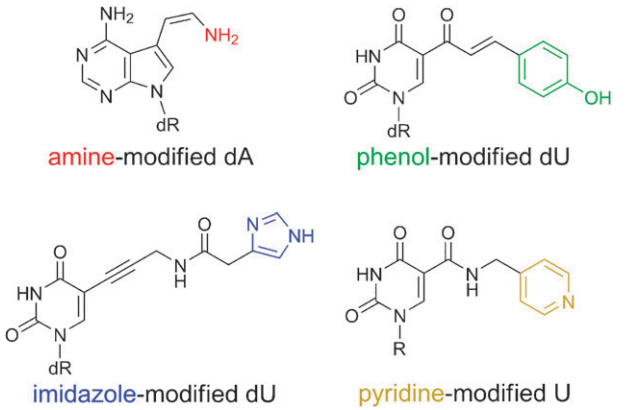


Fig. 20 Examples of nonstandard nucleotides that have been enzymatically incorporated into deoxyribozymes and ribozymes. dR = deoxyribose; R = ribose. Structures from refs. 134 and 135 (DNA) and ref. 98 (RNA).

described.[1] In general, chemically modified deoxyribozymes have substantial potential for improving DNA catalysis, but this potential is largely unexplored.

When is it not 'DNA' anymore?

For all of the catalytic improvement strategies mentioned in this section, one can question the fairness of including the additional molecular component(s) and still claiming to have 'catalytic DNA'. A comparison with protein enzymes is again useful. Many proteins are obligatory metalloenzymes, but without the concern that the metal ion requirement invalidates the assignment of catalytic function to the protein. Similarly, proteins with small-molecule cofactors are widespread and well-accepted as 'enzymes'. Proteins that have unnatural or modified amino acid residues are less common but also readily acknowledged as enzymes. In all of these cases, the additional component (metal, cofactor, modified amino acid) does not disqualify the protein from being considered as the catalyst. Therefore, it seems reasonable to assert that any catalytic DNA which uses additional molecular components is still fundamentally a 'deoxyribozyme'. Of course, certain additional components—particularly, it seems likely, the incorporation of unnatural nucleotides—raise practical concerns such as the need to synthesize the modified deoxynucleoside triphosphate. These concerns may refocus researchers' attention on catalytic DNA that avoids these additional components in the first place. On the other hand, a requirement for a metal or small-molecule cofactor will not pose a practical difficulty if the cofactor is inexpensive and readily obtained, especially if the DNA itself includes only the standard four nucleotides.

Interesting phenomena involving DNA that are not examples of 'catalytic DNA'

Several interesting applications of DNA that are distinct from deoxyribozymes have been reported. Although there may be specific exceptions, in general these other phenomena cannot be termed 'catalytic DNA'. This attention to terminology enables focus on the underlying chemical challenges that face deoxyribozymes. Other DNA-based applications for which 'catalytic DNA' is *not* generally involved are DNA-encoded chemical libraries; DNA-templated synthesis (DTS) along with the related DNA-directed catalysis; and DNA-based asymmetric catalysis. A common general feature of these other applications is that the DNA sequences are arbitrarily chosen, which is clearly not the case for deoxyribozymes. As is well appreciated for natural selection,[33] the *in vitro* selection process—and therefore each DNA sequence that emerges from the selection process—is decidedly non-arbitrary.

DNA-encoded chemical libraries

The basis of DNA-encoded chemical libraries is the covalent conjugation of encoding DNA sequences to other compounds that may have useful biochemical or biological properties. The desired compounds are identified *via* these properties, and their structures are discerned indirectly from the DNA sequence. As described in several reviews,[138,139] this approach has close similarities with phage display and other biologically related technologies, although those methods are generally restricted to the encoding of polypeptides. In the original conception, DNA-encoded chemical libraries are constructed by stepwise split-and-pool synthesis of polypeptide sequences on beads, with the DNA that encodes each polypeptide synthesized in parallel on the same beads.[140–142] In principle, molecules other than polypeptides could similarly be attached to beads alongside the encoding DNA strands, although this has not been reported.

'DNA display' is a particular DNA-encoded library strategy developed by Halpin and Harbury in which functionalized single-stranded DNA molecules are used to encode synthetic non-peptide small molecules.[143–146] A key feature of DNA display is that the DNA is directly involved in library synthesis; the DNA sequence that encodes each small molecule remains covalently attached to that molecule. As a recent application of DNA display, novel SH3 domain ligands were obtained by selection from a combinatorial library of 10^8 DNA-encoded 8-mer peptoids.[147]

Another variant of DNA-encoded chemical libraries is the 'encoded self-assembling chemical' (ESAC) library approach of Neri and co-workers. This approach is based on physically associating potential pharmacophore moieties by hybridization of two or three attached encoding DNA strands, followed by *in vitro* selection for specific protein binding.[148] Self-assembly of the encoding DNA strands provides a non-catalytic scaffold upon which the pharmacophores are brought together, potentially increasing the affinity or selectivity for the protein target beyond that of the individual pharmacophores alone.[149] ESAC libraries have been used to identify small-molecule compounds that bind to trypsin,[150] albumin,[150] and several other proteins.

A key feature of all of these DNA-encoded chemical libraries is that the DNA sequences are arbitrarily chosen, with the provision that each sequence must have minimal Watson–Crick overlap with the other encoding strands. The power of encoding lies in the informational connection between DNA sequence and small-molecule (or polypeptide) constitution. However, this does not have any relationship

with 'deoxyribozymes' in the catalytic sense, because no catalysis is occuring in the DNA-encoded libraries.

DNA-templated synthesis

The field of DNA-templated synthesis (DTS) using non-nucleotide functional groups was initiated in 1992 by Lynn and co-workers, who used a hexamer DNA oligonucleotide template with two trimer substrates that aligned on the template and became joined by reductive amination of the amine and aldehyde moieties attached to the substrates.[151] This work was based on many studies since the 1960s on nucleic acid-templated synthesis of nucleic acids by Orgel and others (reviewed in ref. 152). Since then, several laboratories have reported DNA as a template for various chemical reactions,[153–155] notably efforts from Liu and co-workers who have used DTS for intriguing applications such as reaction discovery.[152,156,157] Although DTS is powerful within its range of applications, it is important to recognize that DTS and catalytic DNA are distinct phenomena. In DTS (as in the DNA-encoded library approaches described above), the specific DNA sequence is arbitrary as long as Watson–Crick base pairing is maintained between the templating elements; indeed, this arbitrary aspect is what makes DTS so useful. The chemical basis of DTS is solely 'effective molarity', in which chemical reactions are controlled entirely by juxtaposition of suitable functional groups as dictated by the DNA template.

In contrast to the situation for DTS, the origin of catalysis by deoxyribozymes goes far beyond the effective molarity phenomenon. A useful way to compare DTS and catalytic DNA is to compare how the corresponding rate enhancements are assessed quantitatively. For DTS, the rate enhancement is due solely to the templating effect and can be calculated as the ratio of $k_{templated}$ and $k_{untemplated}$, where the former is a pseudo-first-order rate constant (s^{-1}) and the latter is a second-order rate constant (M^{-1} s^{-1}; Fig. 21a). The rate enhancement for DTS therefore has units of concentration (M), as expected for 'effective molarity'. In sharp contrast, for catalytic DNA the rate enhancement can be determined as the ratio of k_{cat} and $k_{templated}$ (Fig. 21b), where k_{cat} pertains to the DNA-catalyzed reaction and $k_{templated}$ is for the reaction in the presence of DNA but without any catalysis taking place (*e.g.*, assayed by using a mutant nonfunctional DNA sequence, or simply a DNA splint). Note that $k_{templated}$ appears in the numerator of the rate enhancement expression for DTS but in the denominator for catalytic DNA. If only templating (*i.e.*, effective molarity) were occurring for a deoxyribozyme, then the rate enhancement $k_{cat}/k_{templated}$ would be merely 1 because k_{cat} would be identical to $k_{templated}$; there would be no rate enhancement beyond the templating effect. The observation that deoxyribozyme rate enhancements, computed as $k_{cat}/k_{templated}$, routinely exceed 10^5 demonstrates clearly that unlike for DTS, catalytic DNA does not derive its power and utility from effective molarity.

DNA-directed catalysis

A recent report described a G-quadruplex DNA that binds a porphyrin-aldehyde substrate without using Watson–Crick base pairs.[158] The substrate is held by the DNA near a catalytic proline moiety that is covalently tethered to the DNA; acetone is then the reaction partner with the aldehyde in an aldol reaction. This variation of DTS can be termed 'DNA-directed catalysis', to emphasize that DNA is acting as a scaffold but apparently is not performing the actual catalysis (whereas in DTS itself, there is no catalysis at all by any component of the system). Although the DNA sequence in this case is not entirely arbitrary because of the requirement to form a G-quadruplex structure, the DNA itself still appears to act solely as an effective molarity template rather than directly as a catalyst; the latter role is played by the DNA-tethered proline. Further experimental data is needed for this single reported example of DNA-directed catalysis.

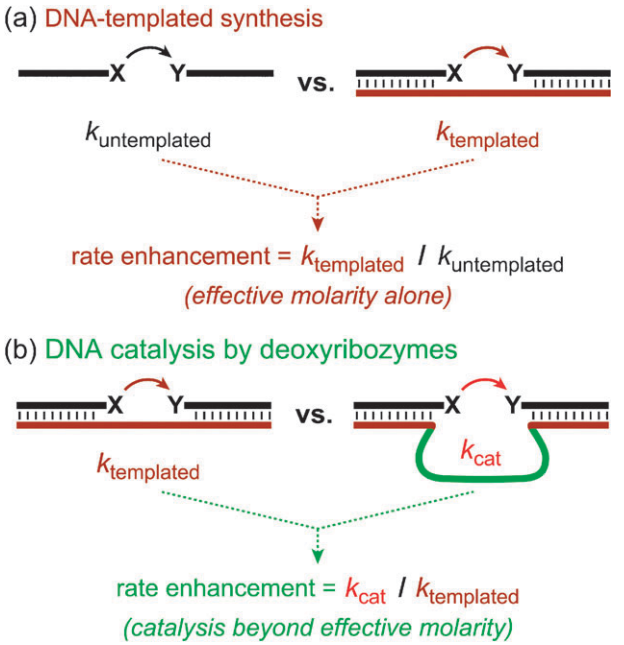


Fig. 21 Comparison of rate enhancements as calculated for (a) DNA-templated synthesis, or (b) deoxyribozymes.

DNA-based asymmetric catalysis

Several recent reports have described the use of DNA as a chiral ligand for metal-based synthetic transformations. This 'DNA-based asymmetric catalysis' is a very promising approach for transferring the stereochemical information inherent within the DNA double helix to small-molecule substrates.[159–161] Asymmetric induction has also been observed during DTS.[162] Nevertheless, in none of these cases does the DNA actually participate as a 'catalyst'. Instead, the DNA plays the role of a large (but nonparticipatory) chiral ligand, by establishing a distinctly chiral environment inside of which a chemical reaction is catalyzed. Situations are conceivable in which a DNA ligand can directly participate in catalysis, in addition to its primary role as the source of a chiral environment. However, unlike for deoxyribozymes, any such direct catalytic role of DNA in DNA-based asymmetric catalysis (at least as originally conceived[159]) would presumably be fortuitous and not an intentional part of the experimental design.

A 'hot spot' for spontaneous DNA depurination

Finally, Amosova *et al*. reported enhanced deguanylation at a specific nucleotide in the loop of a stem-loop DNA structure.[163] This reaction, which occurs in the absence of divalent metal ions or any other cofactors, appears to be an example of a spontaneous DNA depurination event that is $\sim 10^5$ faster than the typical background rate (albeit with k_{obs} of only $\sim 10^{-4}$ min^{-1} at pH 5). This is similar conceptually if not mechanistically to rapid spontaneous RNA backbone cleavage at certain nucleotides dependent upon the internucleotide geometry.[38] For each of the deguanylation or backbone cleavage reactions considered separately, the same mechanism appears to operate at all nucleotide positions, but the rate can be much faster at certain special sites. Until specific mechanistic evidence arises to the contrary, a reasonable explanation of the deguanylation finding[163] is that a conformational preference within the DNA stem-loop sequence promotes an enhanced rate of spontaneous depurination at a particular site, similar to the general situation with RNA backbone cleavage. Neither type of reaction is an example of 'nucleic acid catalysis'.

Conclusion—prospects and challenges for catalytic DNA in synthetic applications

Since the first report in 1994, deoxyribozymes have been shown to catalyze a variety of reactions, including several transformations of potential synthetic interest. This includes numerous reactions of oligonucleotides and, more recently, several reactions of non-oligonucleotide substrates. DNA catalysts will almost certainly not supplant traditional small-molecule catalysts for most synthetic chemistry applications. Nonetheless, in cases where the high selectivity of an 'enzyme' is required along with a modest reaction scale (*e.g*., milligrams or less of product), use of catalytic DNA is plausible to contemplate. Two principal challenges that currently limit the practical synthetic applications of deoxyribozymes are (1) broadening the scope of reactions known to be catalyzed by DNA, combined with (2) expanding the ability of DNA catalysts to function with small molecules and proteins rather than oligonucleotide substrates. The studies and concepts described in this review represent the first steps along the path from the initial identification of deoxyribozymes to, we envision, their utility in a range of synthetic applications.

Acknowledgements

Research in the Silverman laboratory on deoxyribozymes has been supported by the National Institutes of Health, the David and Lucile Packard Foundation, the March of Dimes Birth Defects Foundation, the Burroughs Wellcome Fund, and the American Chemical Society Petroleum Research Fund.

Notes and references

1. S. K. Silverman, *Nucleic Acids Res*., 2005, **33**, 6151–6163.
2. A. Peracchi, *ChemBioChem*, 2005, **6**, 1316–1322.
3. Y. Lu and J. Liu, *Curr. Opin. Biotechnol*., 2006, **17**, 580–588.
4. C. Höbartner and S. K. Silverman, *Biopolymers*, 2007, **87**, 279–292.
5. D. A. Baum and S. K. Silverman, *Cell. Mol. Life Sci*., 2008, **65**, , DOI: 10.1007/s00018-008-8029-y.
6. N. C. Seeman, *Nature*, 2003, **421**, 427–431.
7. P. W. Rothemund, *Nature*, 2006, **440**, 297–302.
8. H. Yang and H. F. Sleiman, *Angew. Chem., Int. Ed*., 2008, **47**, 2443–2446.
9. Y. He, T. Ye, M. Su, C. Zhang, A. E. Ribbe, W. Jiang and C. Mao, *Nature*, 2008, **452**, 198–201.
10. F. A. Aldaye and H. F. Sleiman, *J. Am. Chem. Soc*., 2007, **129**, 13376–13377.
11. N. C. Seeman, *Trends Biochem. Sci*., 2005, **30**, 119–125.
12. S. K. Silverman, *Mol. Biosyst*., 2007, **3**, 24–29.
13. J. A. Doudna and T. R. Cech, *Nature*, 2002, **418**, 222–228.
14. S. K. Silverman, *Wiley Encyclopedia of Chemical Biology*, ed. T Begley, John Wiley & Sons, 2008, see http://www.amazon.co.uk/Wiley-Encyclopedia-Chemical-Biology-Begley/dp/0471754773; chapter http://mrw.interscience.wiley.com/emrw/9780470048672/wecb/article/wecb406/current/abstract.
15. S. K. Silverman, in *Functional Nucleic Acids for Sensing and Other Analytical Applications*, eds. Y. Lu and Y. Li, Springer, New York, 2007.
16. M. Famulok, J. S. Hartig and G. Mayer, *Chem. Rev*., 2007, **107**, 3715–3743.
17. T. Hermann and D. J. Patel, *Science*, 2000, **287**, 820–825.
18. T. R. Cech, *Science*, 2000, **289**, 878–879.
19. C. A. Collins and C. Guthrie, *Nat. Struct. Biol*., 2000, **7**, 850–854.
20. T. Villa, J. A. Pleiss and C. Guthrie, *Cell*, 2002, **109**, 149–152.
21. S. Valadkhan, *Biol. Chem*., 2007, **388**, 693–697.
22. U. F. Müller, *Cell. Mol. Life Sci*., 2006, **63**, 1278–1293.
23. K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling and T. R. Cech, *Cell*, 1982, **31**, 147–157.
24. C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace and S. Altman, *Cell*, 1983, **35**, 849–857.
25. D. L. Robertson and G. F. Joyce, *Nature*, 1990, **344**, 467–468.
26. G. F. Joyce, *Annu. Rev. Biochem*., 2004, **73**, 791–836.
27. J. Kaplan and W. F. DeGrado, *Proc. Natl. Acad. Sci. U. S. A*., 2004, **101**, 11566–11570.
28. L. Jiang, E. A. Althoff, F. R. Clemente, L. Doyle, D. Röthlisberger, A. Zanghellini, J. L. Gallaher, J. L. Betker, F. Tanaka, C. F. BarbasIII, D. Hilvert, K. N. Houk, B. L. Stoddard and D. Baker, *Science*, 2008, **319**, 1387–1391.
29. D. Röthlisberger, O. Khersonsky, A. M. Wollacott, L. Jiang, J. DeChancie, J. Betker, J. L. Gallaher, E. A. Althoff, A. Zanghellini, O. Dym, S. Albeck, K. N. Houk, D. S. Tawfik and D. Baker, *Nature*, 2008, **453**, 190–195.
30. R. R. Breaker and G. F. Joyce, *Chem. Biol*., 1994, **1**, 223–229.
31. Y. Li and D. Sen, *Nat. Struct. Biol*., 1996, **3**, 743–747.
32. P. C. Sabeti, P. J. Unrau and D. P. Bartel, *Chem. Biol*., 1997, **4**, 767–774.
33. R. Dawkins, *The Blind Watchmaker*, W. W. Norton & Company, Inc., New York, 1986.
34. A. Flynn-Charlebois, T. K. Prior, K. A. Hoadley and S. K. Silverman, *J. Am. Chem. Soc*., 2003, **125**, 5346–5350.
35. R. C. Cadwell and G. F. Joyce, *PCR Methods Appl*., 1994, **3**, S136–S140.
36. P. J. Hergenrother, *Curr. Opin. Chem. Biol*., 2006, **10**, 213–218.
37. M. V. Rodnina, M. Beringer and W. Wintermeyer, *Trends Biochem. Sci*., 2007, **32**, 20–26.
38. G. A. Soukup and R. R. Breaker, *RNA*, 1999, **5**, 1308–1325.
39. M. Mandal and R. R. Breaker, *Nat. Rev. Mol. Cell Biol*., 2004, **5**, 451–463.
40. S. W. Santoro and G. F. Joyce, *Proc. Natl. Acad. Sci. U. S. A*., 1997, **94**, 4262–4266.
41. R. P. G. Cruz, J. B. Withers and Y. Li, *Chem. Biol*., 2004, **11**, 57–67.
42. D. Faulhammer and M. Famulok, *Angew. Chem., Int. Ed. Engl*., 1996, **35**, 2837–2841.
43. J. Li, W. Zheng, A. H. Kwon and Y. Lu, *Nucleic Acids Res*., 2000, **28**, 481–488.
44. K. Schlosser and Y. Li, *Biochemistry*, 2004, **43**, 9695–9707.
45. K. Schlosser, J. Gu, L. Sule and Y. Li, *Nucleic Acids Res*., 2008, **36**, 1472–1481.
46. N. Carmi and R. R. Breaker, *Bioorg. Med. Chem*., 2001, **9**, 2589–2600.

47. N. Carmi, L. A. Shultz and R. R. Breaker, *Chem. Biol.*, 1996, **3**, 1039–1046.
48. N. Carmi, S. R. Balkhi and R. R. Breaker, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 2233–2237.
49. A. Radzicka and R. Wolfenden, *Science*, 1995, **267**, 90–93.
50. Y. Li and R. R. Breaker, *J. Am. Chem. Soc.*, 1999, **121**, 5364–5372.
51. M. J. Moore and P. A. Sharp, *Science*, 1992, **256**, 992–997.
52. M. J. Moore and C. C. Query, in *RNA-Protein Interactions: A Practical Approach*, ed. C. W. J. Smith, Oxford University Press, Oxford, 1998, pp. 75-108.
53. J. D. Bain and C. Switzer, *Nucleic Acids Res.*, 1992, **20**, 4372.
54. M. R. Stark, J. A. Pleiss, M. Deras, S. A. Scaringe and S. D. Rader, *RNA*, 2006, **12**, 2014–2019.
55. C. A. Grosshans and T. R. Cech, *Nucleic Acids Res.*, 1991, **19**, 3875–3880.
56. A. R. Ferré-D'Amaré and J. A. Doudna, *Nucleic Acids Res.*, 1996, **24**, 977–978.
57. A. Flynn-Charlebois, Y. Wang, T. K. Prior, I. Rashid, K. A. Hoadley, R. L. Coppins, A. C. Wolf and S. K. Silverman, *J. Am. Chem. Soc.*, 2003, **125**, 2444–2454.
58. D. R. Semlow and S. K. Silverman, *J. Mol. Evol.*, 2005, **61**, 207–215.
59. K. A. Hoadley, W. E. Purtha, A. C. Wolf, A. Flynn-Charlebois and S. K. Silverman, *Biochemistry*, 2005, **44**, 9217–9231.
60. N. Paul, G. Springsteen and G. F. Joyce, *Chem. Biol.*, 2006, **13**, 329–338.
61. Y. Wang and S. K. Silverman, *RNA*, 2006, **12**, 1142–1146.
62. R. L. Coppins and S. K. Silverman, *J. Am. Chem. Soc.*, 2004, **126**, 16426–16432.
63. Y. Wang and S. K. Silverman, *Biochemistry*, 2005, **44**, 3017–3023.
64. W. E. Purtha, R. L. Coppins, M. K. Smalley and S. K. Silverman, *J. Am. Chem. Soc.*, 2005, **127**, 13124–13125.
65. Y. Wang and S. K. Silverman, *J. Am. Chem. Soc.*, 2003, **125**, 6880–6881.
66. C. L. Peebles, P. S. Perlman, K. L. Mecklenburg, M. L. Petrillo, J. H. Tabor, K. A. Jarrell and H. L. Cheng, *Cell*, 1986, **44**, 213–223.
67. Y. Wang and S. K. Silverman, *Biochemistry*, 2003, **42**, 15252–15263.
68. D. M. J. Lilley, *Q. Rev. Biophys.*, 2000, **33**, 109–159.
69. R. L. Coppins and S. K. Silverman, *Nat. Struct. Mol. Biol.*, 2004, **11**, 270–274.
70. R. L. Coppins and S. K. Silverman, *J. Am. Chem. Soc.*, 2005, **127**, 2900–2907.
71. Y. Wang and S. K. Silverman, *ACS Chem. Biol.*, 2006, **1**, 316–324.
72. E. D. Pratico, Y. Wang and S. K. Silverman, *Nucleic Acids Res.*, 2005, **33**, 3503–3512.
73. E. D. Pratico and S. K. Silverman, *RNA*, 2007, **13**, 1528–1536.
74. E. Zelin, Y. Wang and S. K. Silverman, *Biochemistry*, 2006, **45**, 2767–2771.
75. D. A. Baum and S. K. Silverman, *Angew. Chem., Int. Ed.*, 2007, **46**, 3502–3504.
76. E. Zelin and S. K. Silverman, *ChemBioChem*, 2007, **8**, 1907–1911.
77. W. Chiuman and Y. Li, *Bioorg. Chem.*, 2002, **30**, 332–349.
78. C. V. Miduturu and S. K. Silverman, *J. Am. Chem. Soc.*, 2005, **127**, 10144–10145.
79. C. V. Miduturu and S. K. Silverman, *Angew. Chem., Int. Ed.*, 2006, **45**, 1918–1921.
80. Y. Wang and S. K. Silverman, *Angew. Chem., Int. Ed.*, 2005, **44**, 5863–5866.
81. Y. Wang and S. K. Silverman, *RNA*, 2006, **12**, 313–321.
82. Y. Li and R. R. Breaker, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 2746–2751.
83. W. Wang, L. P. Billen and Y. Li, *Chem. Biol.*, 2002, **9**, 507–517.
84. Y. Li, Y. Liu and R. R. Breaker, *Biochemistry*, 2000, **39**, 3106–3114.
85. A. Sreedhara, Y. Li and R. R. Breaker, *J. Am. Chem. Soc.*, 2004, **126**, 3454–3460.
86. B. Cuenoud and J. W. Szostak, *Nature*, 1995, **375**, 611–614.
87. T. L. Sheppard, P. Ordoukhanian and G. F. Joyce, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 7802–7807.
88. C. Höbartner, P. I. Pradeepkumar and S. K. Silverman, *Chem. Commun.*, 2007, 2255–2257.
89. D. J. Chinnapen and D. Sen, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 65–69.
90. D. J. Chinnapen and D. Sen, *J. Mol. Biol.*, 2007, **365**, 1326–1336.
91. A. Jäschke, C. Frauendorf and F. Hausch, *Synlett*, 1999, 825–833.
92. Y. Li and D. Sen, *Biochemistry*, 1997, **36**, 5589–5599.
93. Y. Li and D. Sen, *Chem. Biol.*, 1998, **5**, 1–12.
94. K. N. Morris, T. M. Tarasow, C. M. Julin, S. L. Simons, D. Hilvert and L. Gold, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 13028–13032.
95. C. Höbartner and S. K. Silverman, *Angew. Chem., Int. Ed.*, 2007, **46**, 7420–7424.
96. J. R. Lorsch and J. W. Szostak, *Nature*, 1994, **371**, 31–36.
97. C. Wilson and J. W. Szostak, *Nature*, 1995, **374**, 777–782.
98. T. M. Tarasow, S. L. Tarasow and B. E. Eaton, *Nature*, 1997, **389**, 54–57.
99. B. Seelig and A. Jäschke, *Chem. Biol.*, 1999, **6**, 167–176.
100. B. Seelig, S. Keiper, F. Stuhlmann and A. Jäschke, *Angew. Chem., Int. Ed.*, 2000, **39**, 4576–4579.
101. F. Stuhlmann and A. Jäschke, *J. Am. Chem. Soc.*, 2002, **124**, 3238–3244.
102. A. Serganov, S. Keiper, L. Malinina, V. Tereshko, E. Skripkin, C. Höbartner, A. Polonskaia, A. T. Phan, R. Wombacher, R. Micura, Z. Dauter, A. Jäschke and D. J. Patel, *Nat. Struct. Mol. Biol.*, 2005, **12**, 218–224.
103. M. Chandra and S. K. Silverman, *J. Am. Chem. Soc.*, 2008, **130**, 2936–2937.
104. R. Wombacher, S. Keiper, S. Suhm, A. Serganov, D. J. Patel and A. Jaschke, *Angew. Chem., Int. Ed.*, 2006, **45**, 2469–2472.
105. J. C. Schlatterer, F. Stuhlmann and A. Jaschke, *ChemBioChem*, 2003, **4**, 1089–1092.
106. P. I. Pradeepkumar, C. Höbartner, D. A. Baum and S. K. Silverman, *Angew. Chem., Int. Ed.*, 2008, **47**, 1753–1757.
107. T. R. Cech, *Science*, 1987, **236**, 1532–1539.
108. J. H. Cate, A. R. Gooding, E. Podell, K. Zhou, B. L. Golden, C. E. Kundrot, T. R. Cech and J. A. Doudna, *Science*, 1996, **273**, 1678–1685.
109. M. V. Rodnina, M. Beringer and P. Bieling, *Biochem. Soc. Trans.*, 2005, **33**, 493–498.
110. S. Hur and T. C. Bruice, *J. Am. Chem. Soc.*, 2003, **125**, 1472–1473.
111. A. Sievers, M. Beringer, M. V. Rodnina and R. Wolfenden, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 7897–7901.
112. A. M. Pyle, *Science*, 1993, **261**, 709–714.
113. P. C. Bevilacqua and R. Yajima, *Curr. Opin. Chem. Biol.*, 2006, **10**, 455–464.
114. M. P. Robertson and W. G. Scott, *Science*, 2007, **315**, 1549–1553.
115. J. Nowakowski, P. J. Shim, G. S. Prasad, C. D. Stout and G. F. Joyce, *Nat. Struct. Biol.*, 1999, **6**, 151–156.
116. S. W. Santoro and G. F. Joyce, *Biochemistry*, 1998, **37**, 13330–13342.
117. R. L. Coppins and S. K. Silverman, *Biochemistry*, 2005, **44**, 13439–13446.
118. G. M. Emilsson, S. Nakamura, A. Roth and R. R. Breaker, *RNA*, 2003, **9**, 907–918.
119. R. R. Breaker, G. M. Emilsson, D. Lazarev, S. Nakamura, I. J. Puskarz, A. Roth and N. Sudarsan, *RNA*, 2003, **9**, 949–957.
120. M. J. Moore and C. C. Query, *Methods Enzymol.*, 2000, **317**, 109–123.
121. D. S. Tawfik and A. D. Griffiths, *Nat. Biotechnol.*, 1998, **16**, 652–656.
122. O. J. Miller, K. Bernath, J. J. Agresti, G. Amitai, B. T. Kelly, E. Mastrobattista, V. Taly, S. Magdassi, D. S. Tawfik and A. D. Griffiths, *Nat. Methods*, 2006, **3**, 561–570.
123. J. J. Agresti, B. T. Kelly, A. Jaschke and A. D. Griffiths, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 16170–16175.
124. M. Levy, K. E. Griswold and A. D. Ellington, *RNA*, 2005, **11**, 1555–1562.
125. H. S. Zaher and P. J. Unrau, *RNA*, 2007, **13**, 1017–1026.
126. C. R. Geyer and D. Sen, *Chem. Biol.*, 1997, **4**, 579–593.
127. D. Faulhammer and M. Famulok, *J. Mol. Biol.*, 1997, **269**, 188–202.

128. W. C. Winkler, A. Nahvi, A. Roth, J. A. Collins and R. R. Breaker, *Nature*, 2004, **428**, 281–286.
129. T. J. McCarthy, M. A. Plog, S. A. Floy, J. A. Jansen, J. K. Soukup and G. A. Soukup, *Chem. Biol.*, 2005, **12**, 1221–1226.
130. A. Roth and R. R. Breaker, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 6027–6031.
131. J. A. Bittker, K. J. Phillips and D. R. Liu, *Curr. Opin. Chem. Biol.*, 2002, **6**, 367–374.
132. P. M. Gramlich, C. T. Wirges, J. Gierlich and T. Carell, *Org. Lett.*, 2008, **10**, 249–251.
133. S. H. Weisbrod and A. Marx, *Chem. Commun.*, 2007, 1828–1830.
134. T. Gourlain, A. Sidorov, N. Mignet, S. J. Thorpe, S. E. Lee, J. A. Grasby and D. M. Williams, *Nucleic Acids Res.*, 2001, **29**, 1898–1905.
135. K. Sakthivel and C. F. Barbas, III, *Angew. Chem., Int. Ed.*, 1998, **37**, 2872–2875.
136. O. Thum, S. Jager and M. Famulok, *Angew. Chem., Int. Ed.*, 2001, **40**, 3990–3993.
137. S. Jäger, G. Rasched, H. Kornreich-Leshem, M. Engeser, O. Thum and M. Famulok, *J. Am. Chem. Soc.*, 2005, **127**, 15071–15082.
138. S. Melkko, C. E. Dumelin, J. Scheuermann and D. Neri, *Drug Discovery Today*, 2007, **12**, 465–471.
139. J. Scheuermann, C. E. Dumelin, S. Melkko and D. Neri, *J. Biotechnol.*, 2006, **126**, 568–581.
140. S. Brenner and R. A. Lerner, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**, 5381–5383.
141. J. Nielsen, S. Brenner and K. D. Janda, *J. Am. Chem. Soc.*, 1993, **115**, 9812–9813.
142. M. C. Needels, D. G. Jones, E. H. Tate, G. L. Heinkel, L. M. Kochersperger, W. J. Dower, R. W. Barrett and M. A. Gallop, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 10700–10704.
143. D. R. Halpin and P. B. Harbury, *PLoS Biol.*, 2004, **2**, E173.
144. D. R. Halpin and P. B. Harbury, *PLoS Biol.*, 2004, **2**, E174.
145. D. R. Halpin, J. A. Lee, S. J. Wrenn and P. B. Harbury, *PLoS Biol.*, 2004, **2**, E175.
146. S. J. Wrenn and P. B. Harbury, *Annu. Rev. Biochem.*, 2007, **76**, 331–349.
147. S. J. Wrenn, R. M. Weisinger, D. R. Halpin and P. B. Harbury, *J. Am. Chem. Soc.*, 2007, **129**, 13137–13143.
148. S. Melkko, J. Scheuermann, C. E. Dumelin and D. Neri, *Nat. Biotechnol.*, 2004, **22**, 568–574.
149. S. Melkko, C. E. Dumelin, J. Scheuermann and D. Neri, *Chem. Biol.*, 2006, **13**, 225–231.
150. S. Melkko, Y. Zhang, C. E. Dumelin, J. Scheuermann and D. Neri, *Angew. Chem., Int. Ed.*, 2007, **46**, 4671–4674.
151. J. T. Goodwin and D. G. Lynn, *J. Am. Chem. Soc.*, 1992, **114**, 9197–9198.
152. X. Li and D. R. Liu, *Angew. Chem., Int. Ed.*, 2004, **43**, 4848.
153. J. L. Czlapinski and T. L. Sheppard, *J. Am. Chem. Soc.*, 2001, **123**, 8618–8619.
154. A. T. Poulin-Kerstien and P. B. Dervan, *J. Am. Chem. Soc.*, 2003, **125**, 15811–15821.
155. Z. Tang and A. Marx, *Angew. Chem., Int. Ed.*, 2007, **46**, 7297–7300.
156. Z. J. Gartner and D. R. Liu, *J. Am. Chem. Soc.*, 2001, **123**, 6961–6963.
157. M. W. Kanan, M. M. Rozenman, K. Sakurai, T. M. Snyder and D. R. Liu, *Nature*, 2004, **431**, 545–549.
158. Z. Tang, D. P. Goncalves, M. Wieland, A. Marx and J. S. Hartig, *ChemBioChem*, 2008, **9**, 1061–1064.
159. G. Roelfes and B. L. Feringa, *Angew. Chem.,Int. Ed.*, 2005, **44**, 3230–3232.
160. G. Roelfes, A. J. Boersma and B. L. Feringa, *Chem. Commun.*, 2006, 635–637.
161. D. Coquière, B. L. Feringa and G. Roelfes, *Angew. Chem., Int. Ed.*, 2007, **46**, 9308–9311.
162. X. Li and D. R. Liu, *J. Am. Chem. Soc.*, 2003, **125**, 10188–10189.
163. O. Amosova, R. Coulter and J. R. Fresco, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 4392–4397.