cGMP Capture

DOI: 10.1002/ange.201403579

Endless: A Purine-Binding RNA Motif that Can Be Expressed in Cells**

*Christoph Kröner, Martin Thunemann, Sven Vollmer, Manuela Kinzer, Robert Feil, and Clemens Richert**

Abstract: *It is becoming increasingly clear that nature uses RNAs extensively for regulating vital functions of the cell, and short sequences are frequently used to suppress gene expression. However, controlling the concentration of small molecules intracellularly through designed RNA sequences that fold into ligand-binding structures is difficult. The development of "endless", a triplex-based folding motif that can be expressed in mammalian cells and binds the second messenger 3′,5′-cyclic guanosine monophosphate (cGMP), is described. In vitro, DNA or RNA versions of endless show low micromolar to nanomolar dissociation constants for cGMP. To test its functionality in vivo, four endless RNA motifs arranged in tandem were co-expressed with a fluorescent cGMP sensor protein in murine vascular smooth muscle cells. Nitric oxide induced endogenous cGMP signals were suppressed in endless-expressing cells compared to cells expressing a control motif, which suggests that endless can act as a genetically encoded cGMP sink to modulate signal transduction in cells.*

RNA performs a number of key intracellular functions. Ribosomal RNA, tRNA, and mRNA are examples of functional ribonucleic acids that are part of the cellular machinery. However, RNA and oligonucleotides derived from RNA sequences are also used to regulate the expression of specific genes. One way in which this regulation is achieved is through binding to complementary sequences through Watson–Crick base pairing. For example, antisense oligonucleotides[1,2] and short interfering RNAs (siRNAs)[3,4] can be designed to interfere with the translation of mRNAs to which their sequences are complementary. In fact, siRNAs have become important tools for chemical biology and enable the probing of intracellular processes and the validation of therapeutic strategies.[5] The mechanism of action of these sequences is based on that of natural microRNAs used by cells to regulate protein synthesis.

The oligonucleotides named above bind RNA targets through duplex formation. However, RNAs that form binding pockets for small molecules also exist. For example, riboswitches are natural nontranslated single-stranded RNAs, the gene regulatory activity of which is modulated through effector molecules. Riboswitches that bind low-molecular-weight metabolites like cyclic di-cGMP,[6] SAM,[7–11] THF,[12,13] FMN,[14] and glucosamine-6-phosphate[15] in bacteria are known. The TPP responsive riboswitch is an example of a riboswitch found in eukaryotes.[16–19] Artificial riboswitches have been developed by fusing target-binding sequences with aptamers. Theophylline-binding switches are used in cells to influence splicing,[20] mRNA stability,[21–25] RNA silencing,[26–30] translation,[31–38] or posttranslational activity.[39] One of these artificial switches affects the proliferation of T cells in mice.[22] Intracellularly expressed RNA can form functional nanostructures in bacteria,[40] and the intracellular synthesis of short triplex-forming RNA strands can downregulate the activity of endogenous gene promoters in human cells.[41]

Despite these impressive advances in RNA technology, to our knowledge, no de novo designed RNA motifs have been reported that act as sinks for small molecules in the cell. One of the small molecules that regulate cell activity and that are closely linked to human disease is 3′,5′-cyclic guanosine monophosphate (cGMP), an important second messenger in eukaryotes.[42] Pharmacologic and genetic manipulation of cGMP-generating enzymes and cGMP-binding effector proteins have demonstrated a central role for the cGMP signaling system in physiology, pathophysiology, and pharmacotherapy in mmamals. However, the spatiotemporal dynamics of cGMP and its functional relevance are not well understood. A recent technological advance was the generation of transgenic mice that express a fluorescent cGMP-sensor protein to visualize endogenous cGMP signals in living mammalian cells.[43] However, cGMP research would also benefit from the development of methods that enable the experimental manipulation of endogenous cGMP levels in cells. We sought to achieve this goal by expressing a genetically encoded RNA-based cGMP sink.

We noted that triplex structures may form in DNA[44–47] or in RNA transcripts.[48] We had recently developed triplex-based binding motifs for purines.[49] The binding pockets of these motifs are gaps in the central strand of parallel DNA or RNA triplexes.[50] Being derived from a canonical triple helix, the de novo designed motifs lend themselves to systematic optimization. The selectivity of the motifs for A or G nucleotides or selected cofactors has been tuned through variation of the bases at or near the binding pocket.[51]

Unfortunately, the motifs developed thus far are unsuitable for intracellular expression because both the 3′ and 5′ termini are part of the binding pockets. Since mRNAs

[*] Dipl.-Chem. C. Kröner, Dipl.-Chem. S. Vollmer, Prof. C. Richert
Institut für Organische Chemie, Universität Stuttgart
70569 Stuttgart (Germany)
E-mail: lehrstuhl-2@oc.uni-stuttgart.de
Dr. M. Thunemann, M. Kinzer, Prof. R. Feil
Interfakultäres Institut für Biochemie, Universität Tübingen
72076 Tübingen (Germany)

[**] This work was supported by DFG (grant No. RI 1063/13-1 to C.R.). We thank Dr. Frank Essmann for help with flow cytometry.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201403579.

typically feature a 5′ cap and a 3′ polyadenylate tail, they would be unable to form the binding pocket or would require elaborate processing with designed cleavage steps. Herein, we report a de novo designed triplex motif termed "*endless*", which can fold to form binding pockets for cGMP within a continuous sequence. A DNA sequence encoding a quadruple version of *endless* was cloned into an expression plasmid that was used to transiently transfect murine vascular smooth muscle cells, a well-known model for cGMP signaling.[43,52] Upon *endless* expression, NO-induced endogenous cGMP levels were reduced, thus indicating that *endless* acts as a sink for cGMP in intact cells.

Designed binding motifs with canonical base pairing form on both the DNA and RNA level.[49] Figure 1 shows the sequences tested in the in vitro stages of our sequence development, which was on the DNA level since oligodeoxynucleotides are less costly than oligoribonucleotides. We first asked how the designed binding pocket may be formed without involving the termini. Unlike known triplex-based motifs,[49,51,53,54] *endless*-type sequences contain a bridging

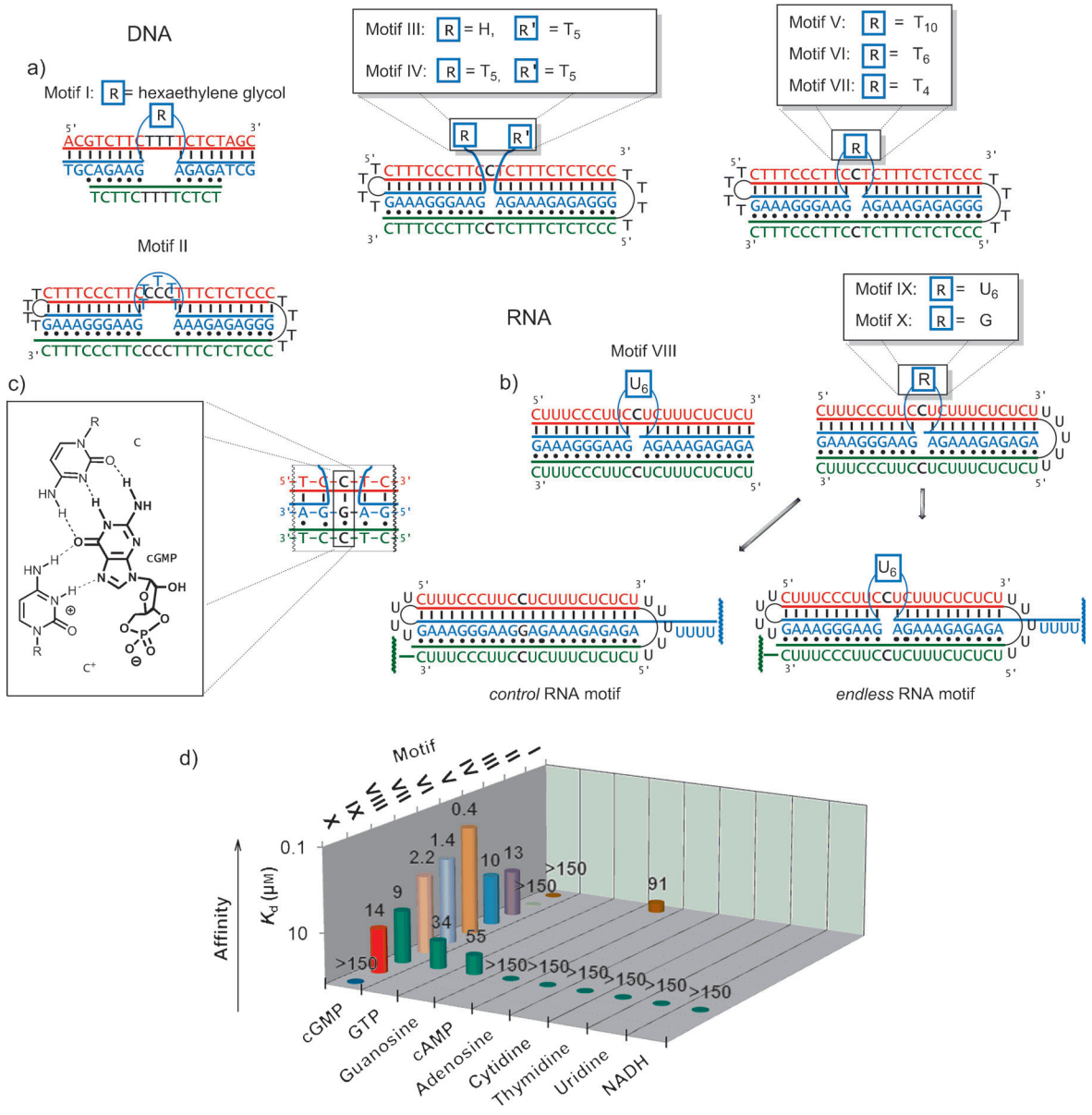

Figure 1. Sequences of binding motifs and their affinities for purine-containing small molecules. a) Sequences studied on the DNA level with binding site sizes of three nucleotides (motifs **I** and **II**) or one nucleotide (motifs **III**–**VII**); b) RNA binding motifs for in vitro characterization (**VIII** and **IX**), control RNA motif **X**, which lacks the binding pocket, and the *endless* and *control* RNA motifs for in vivo studies; c) Designed cGMP binding interactions. d) Dissociation constants for complexes with small molecules, determined by filtration assay at 10 mM phosphate buffer, pH 7.0, 1 M NaCl (DNA motifs) or 10 mM phosphate buffer, pH 7.0, 150 mM NaCl, 10 mM $MgCl_2$ (RNA motifs) at 4 °C. Numerical values for the dissociation constants are given at the top of each column. See Chapter 3 of the Supporting Information for further details.

segment in the oligopurine strand, which is designed to form a bulge. Our binding pockets featured two pyrimidines that can engage the target purine in Watson–Crick and Hoogsteen base pairing. Two assays were employed to detect ligand binding. One assay involved membrane filtration (3 kDa cut-off) to separate the solution with unbound nucleotides from the folding motif or its complex with the ligand. In some instances, UV-melting assays were employed, since some triplex motifs release the ligand simultaneously with the triplex-forming segment, resulting in a shift in the triplex-to-duplex transition.[49]

The first assays were run with motif **I**, the three-nucleotide adenine-specific binding site of which was bridged by a phosphodiester-linked hexaethylene glycol chain. This motif gave an apparent dissociation constant of 91 μM for cAMP (Figure 1d and Table S2 in the Supporting Information), suggesting that the bridge with its two phosphodiester linkages does not block binding. Control experiments showed the expected selectivity for cAMP over cGMP, with no detectable binding to the latter (Table S2). Subsequent experiments were performed with motif **II**, which has a three-nucleotide gap and was designed to bind cGMP. Motif **II** features a bridge of T residues and gave disappointing results, with no detectable binding to cGMP.

In order to determine whether nucleotides as bridging units were tolerated at all, we next studied motifs **III** and **IV**, which contain single-stranded flaps and a one-nucleotide binding site but no continuous bridge. Both motifs bound cGMP with double-digit micromolar affinity, thus suggesting that nucleotides protruding from the central oligopurine strand do not suppress binding. This encouraged us to test motifs with a one-nucleotide binding pocket and a longer bridge containing at least four bases (motifs **V**–**VII**). Very high affinity for cGMP was measured, with K_d values down to the nanomolar range (motif **V**, Figure 2a). The affinity of the bridged motifs for cGMP was also apparent from UV-melting curves, which showed increases in the triplex melting point (Table S1).

Motif **VI** was a compromise of target affinity and overall length. Its binding site was therefore transferred to the RNA level, leading to motifs **VIII** and **IX** (Figure 1). The latter is a two-strand motif, whereas the former was produced from three strands because very long synthetic RNAs were not accessible. Gratifyingly, RNA motif **VIII** gave a shift in the triplex melting point of 9.1 °C in the presence of cGMP, indicating tight binding (Figure 2b). Moreover, the dissociation constant of the complex with cGMP, determined by filtration assays, was 9 μM (Figure 1d). Importantly, control filtrations with RNA motif **VIII** and other nucleosides or nucleoside phosphates showed selectivity for cGMP over other purine-containing structures (GTP, guanosine, adenosine, NADH) or pyrimidine-containing compounds (cytosine, thymidine, uridine; Figure 1d and Table S2). The actual cGMP affinity of RNA motif **VIII** in intact cells was inaccessible but the micromolar dissociation constants determined in vitro are in the range of the binding affinities of intracellular cGMP receptor proteins and of the cGMP concentrations reached in cells.[43] A control experiment with RNA motif **X**, which lacks the binding pocket, containing a continuous oligopurine strand instead, showed no detectable affinity for cGMP (Table S2). These results for RNA motifs **VIII** and **IX** led to the final *endless* RNA motif for expression in cells (Figure 3a), where it could act as a cGMP sink. Multiple copies of this motif can be arranged in a seemingly endless sequence, hence the name.

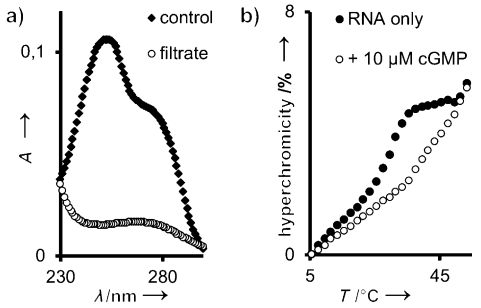

Figure 2. Folding motifs of the *endless* type bind cGMP tightly. a) UV spectra from filtration assay. Conditions: 10 mM phosphate buffer, pH 7.0, 1 M NaCl. b) Expansion of the UV-melting curve of motif VIII in the presence or absence of cGMP, showing the shift in the triplex-to-duplex transition. Conditions: 0.5 μM oligonucleotide, 10 μM cGMP, 10 mM phosphate buffer, pH 6.0, 1 M NaCl, heating rate 1 °C min^{-1}.

To test the functionality of *endless* in mammalian cells, a DNA sequence encoding four consecutive motifs (*endless*^{4}) was cloned into an expression vector containing the strong CAG (chicken actin/β-globin) promoter. The *endless*^{4} RNA constituted the 5′ part of a longer transcript that also encodes the fluorescence resonance energy transfer (FRET)-based cGMP sensor protein cGi500[55] (Figure 3a). This strategy allowed the co-expression of *endless*^{4} RNA with the cGMP sensor used to monitor intracellular cGMP concentrations. Primary vascular smooth muscle cells from mouse aorta, which generate cGMP transients upon stimulation with nitric oxide,[43,52] were transiently transfected with the *endless*^{4}–cGi500 construct or with a control construct encoding four consecutive copies of a motif lacking a cGMP-binding pocket (*control*^{4}–cGi500). Transfected cells were recognized by their fluorescence owing to cGi500 expression and were used for FRET-based cGMP imaging. During imaging, cells were superfused with increasing concentrations of the nitric oxide donor DEA/NO (Figure 3b). Peak heights of the resulting cGMP transients were taken to estimate half maximal effective concentration (EC_{50}) values for the nitric oxide induced cGMP responses. A shift of the EC_{50} value from 132 nM for cells transfected with *control*^{4}–cGi500 to 236 nM for cells transfected with *endless*^{4}–cGi500 was observed (Figure 3c). Similar results were obtained in an independent experiment with a different cell preparation (Figure S8 in the Supporting Information). These results indicate that *endless*^{4} acts as an intracellular sink for cGMP, thereby reducing the amount of cGMP available for binding to cGi500 and other target proteins. Apparently, *endless*^{4} RNA is able to compete well for cGMP binding with the cGMP sensor, a protein-based cGMP binder.

Neither the cellular and nuclear morphology of the transfected cells (Figure S7 and data not shown) nor the evaluation of surface-exposed phosphatidylserine, a marker

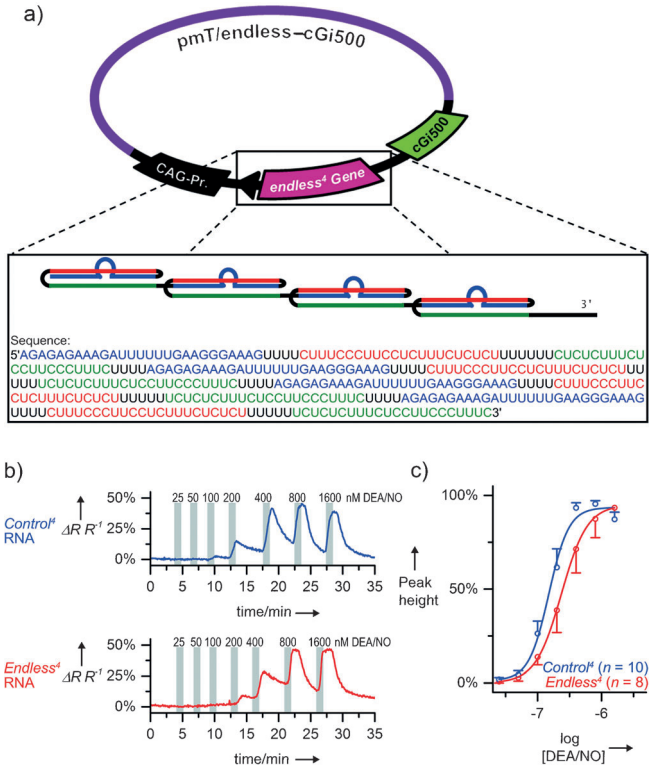

Figure 3. Evaluation of NO-induced cGMP transients in vascular smooth muscle cells expressing *endless*^{4} or *control*^{4} RNA. a) The *endless*^{4}–cGi500 expression vector and its transcript. The expanded area shows the sequence of the RNA transcript of the *endless*^{4} gene (see Figure S7). b) Representative FRET traces for two individual cells expressing either *endless*^{4} or *control*^{4} RNA together with cGi500. Shown is the baseline-normalized FRET ratio ΔR/R, which represents the intracellular cGMP concentration measured with cGi500. Gray bars represent periods of DEA/NO superfusion with the indicated concentrations. c) Peak-height analysis of the cGMP transients. Shown are mean (±SEM) values for 8 cells expressing *endless*^{4} RNA and 10 cells expressing *control*^{4} RNA. Curves are from nonlinear regression with a Hill function. The *p*-value for a comparison of peak heights (at 100, 200, 400, and 800 nM DEA/NO) for *endless*^{4} versus *control*^{4} RNA expressing cells with two-way repeated-measures ANOVA is 0.08. Similar results were obtained in an independent experiment (see Figure S8).

for apoptotic cells, with annexin V by using flow cytometry (data not shown) indicated adverse effects of *endless*^{4}–cGi500 expression on cell viability as compared to the *control*^{4}–cGi500 construct. Studies are under way to correlate altered intracellular cGMP levels upon expression of *endless* RNA with biological responses in the affected cells.

In conclusion, we have developed the purine-binding folding motif *endless*, which binds cGMP without engaging the termini of the strand. There are several known ways to stimulate cGMP production in the cell,[42] but there is no known inducible sink for second messengers that we are aware of. The results from the expression of *endless* in eukaryotic cells suggest that it can be used as inducible sink for endogenous cGMP in living cells. Experiments aimed at developing other chemical biology tools based on ligand-binding RNA sequences are under way.

Received: March 21, 2014
Revised: April 29, 2014
Published online: July 9, 2014

Keywords: aptamers · cGMP · nucleotides · oligonucleotides · RNA

[1] P. C. Zamecnik, M. L. Stephenson, *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 280–284.
[2] a) E. Uhlmann, A. Peyman, *Chem. Rev.* **1990**, *90*, 543–584; b) C. F. Bennett, E. E. Swayze, *Annu. Rev. Pharmacol. Toxicol.* **2010**, *50*, 259–293; c) M. A. Campbell, J. Wengel, *Chem. Soc. Rev.* **2011**, *40*, 5680–5689.
[3] A. Reynolds, D. Leake, Q. Boese, S. Scaringe, W. S. Marshall, A. Khvorova, *Nat. Biotechnol.* **2004**, *22*, 326–330.
[4] M. Manoharan, A. Akinc, R. K. Pandey, J. Qin, P. Hadwiger, M. John, K. Mills, K. Charisse, M. A. Maier, L. Nechev, E. M. Greene, P. S. Pallan, E. Rozners, K. G. Rajeev, M. Egli, *Angew. Chem.* **2011**, *123*, 2332–2336; *Angew. Chem. Int. Ed.* **2011**, *50*, 2284–2288.
[5] R. Kittler, G. Putz, L. Pelletier, I. Poser, A.-K. Heninger, D. Drechsel, S. Fischer, I. Konstantinova, B. Habermann, H. Grabner, M.-L. Yaspo, H. Himmelbauer, B. Korn, K. Neugebauer, M. T. Pisabarro, F. Buchholz, *Nature* **2004**, *432*, 1036–1040.
[6] N. Sudarsan, E. R. Lee, Z. Weinberg, R. H. Moy, J. N. Kim, K. H. Link, R. R. Breaker, *Science* **2008**, *321*, 411–413.
[7] R. K. Montange, R. T. Batey, *Nature* **2006**, *441*, 1172–1175.
[8] W. C. Winkler, A. Nahvi, N. Sudarsan, J. E. Barrick, R. R. Breaker, *Nat. Struct. Biol.* **2003**, *10*, 701–707.
[9] B. A. McDaniel, F. J. Grundy, I. Artsimovitch, T. M. Henkin, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3083–3088.
[10] V. Epshtein, A. S. Mironov, E. Nudler, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5052–5056.
[11] R. T. Fuchs, F. J. Grundy, T. M. Henkin, *Nat. Struct. Mol. Biol.* **2006**, *13*, 226–233.
[12] J. J. Trausch, P. Ceres, F. E. Reyes, R. T. Batey, *Structure* **2011**, *19*, 1413–1423.
[13] T. D. Ames, D. A. Rodionov, Z. Weinberg, R. R. Breaker, *Chem. Biol.* **2010**, *17*, 681–685.
[14] A. Serganov, L. Huang, D. J. Patel, *Nature* **2009**, *458*, 233–237.
[15] D. J. Klein, A. R. Ferré-D'Amaré, *Science* **2006**, *313*, 1752–1756.
[16] N. Sudarsan, J. E. Barrick, R. R. Breaker, *RNA* **2003**, *9*, 644–647.
[17] M. T. Croft, M. Moulin, M. E. Webb, A. G. Smith, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 20770–20775.
[18] S. Thore, M. Leibundgut, N. Ban, *Science* **2006**, *312*, 1208–1211.
[19] A. Serganov, A. Polonskaia, A. T. Phan, R. R. Breaker, D. J. Patel, *Nature* **2006**, *441*, 1167–1171.
[20] D. S. Kim, V. Gusti, K. J. Dery, R. K. Gaur, *BMC Mol. Biol.* **2008**, *9*, 23.
[21] M. N. Win, C. D. Smolke, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 14283–14288.
[22] Y. Y. Chen, M. C. Jensen, C. D. Smolke, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 8531–8536.
[23] M. N. Win, C. D. Smolke, *Science* **2008**, *322*, 456–460.
[24] A. H. Babiskin, C. D. Smolke, *Nucleic Acids Res.* **2011**, *39*, 5299–5311.
[25] S. Ausländer, P. Ketzer, J. S. Hartig, *Mol. BioSyst.* **2010**, *6*, 807–814.
[26] C. L. Beisel, Y. Y. Chen, S. J. Culler, K. G. Hoff, C. D. Smolke, *Nucleic Acids Res.* **2011**, *39*, 2981–2994.
[27] D. Kumar, C. An, Y. Yokobayashi, *J. Am. Chem. Soc.* **2009**, *131*, 13906–13907.
[28] C. An, V. B. Trinh, Y. Yokobayashi, *RNA* **2006**, *12*, 710–716.

[29] N. Tuleuova, C. An, E. Ramanculov, A. Revzin, *Biochem. Biophys. Res. Commun.* **2008**, *376*, 169 – 173.
[30] C. L. Beisel, T. S. Bayer, K. G. Hoff, C. D. Smolke, *Mol. Syst. Biol.* **2008**, *4*, 224.
[31] S. A. Lynch, S. K. Desai, H. K. Sajja, J. P. Gallivan, *Chem. Biol.* **2007**, *14*, 173 – 184.
[32] S. A. Lynch, J. P. Gallivan, *Nucleic Acids Res.* **2009**, *37*, 184 – 192.
[33] S. Topp, J. P. Gallivan, *Chem. Biochem.* **2008**, *9*, 210 – 213.
[34] S. Topp, J. P. Gallivan, *RNA* **2008**, *14*, 2498 – 2503.
[35] S. Topp, C. M. Reynoso, J. C. Seeliger, I. S. Goldlust, S. K. Desai, D. Murat, A. Shen, A. W. Puri, A. Komeili, C. R. Bertozzi, *Appl. Environ. Microbiol.* **2010**, *76*, 7881 – 7884.
[36] B. Suess, B. Fink, C. Berens, R. Stentz, W. Hillen, *Nucleic Acids Res.* **2004**, *32*, 1610 – 1614.
[37] A. Verhounig, D. Karcher, R. Bock, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 6204 – 6209.
[38] V. Sharma, Y. Nomura, Y. Yokobayashi, *J. Am. Chem. Soc.* **2008**, *130*, 16310 – 16315.
[39] D. Ausländer, M. Wieland, S. Ausländer, M. Tigges, M. Fussenegger, *Nucleic Acids Res.* **2011**, *39*, 155.
[40] C. J. Delebecque, A. B. Lindner, P. A. Silver, F. A. Aldaye, *Science* **2011**, *333*, 470 – 474.
[41] S. Cogoi, C. Suraci, E. Del Terra, S. Diviacco, G. van der Marel, J. van Boom, F. Quadrifoglio, L. Xodo, *Antisense Nucleic Acid Drug Dev.* **2000**, *10*, 283 – 295.
[42] a) J. A. Beavo, L. L. Brunton, *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 710 – 718; b) B. Kemp-Harper, R. Feil, *Sci. Signaling* **2008**, *1*, pe12.
[43] M. Thunemann, L. Wen, M. Hillenbrand, A. Vachaviolos, S. Feil, T. Ott, X. Han, D. Fukumura, R. K. Jain, M. Russwurm, C. de Wit, R. Feil, *Circ. Res.* **2013**, *113*, 365 – 371.
[44] R. Besch, C. Giovannangeli, T. Schuh, C. Kammerbauer, K. Degitz, *J. Mol. Biol.* **2004**, *341*, 979 – 989.
[45] Y. M. Agazie, G. D. Burrkholder, J. S. Lee, *J. Biochem.* **1996**, *316*, 461 – 466.
[46] G. Wang, K. M. Vasquez, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13448 – 13453.
[47] G. P. Schroth, P. Shing Ho, *Nucleic Acids Res.* **2004**, *32*, 1977 – 1983.
[48] J. A. Brown, M. L. Valenstein, T. A. Yario, K. T. Tycowski, J. A. Steitz, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 19202 – 19207.
[49] C. Kröner, M. Roethlingshoefer, C. Richert, *J. Org. Chem.* **2011**, *76*, 2933 – 2936.
[50] N. T. Thuong, C. Hélène, *Angew. Chem.* **1993**, *105*, 697 – 723; *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 666 – 690.
[51] C. Kröner, A. Göckel, W. Liu, C. Richert, *Chem. Eur. J.* **2013**, *19*, 15879 – 15887.
[52] R. Feil, N. Gappa, M. Rutz, J. Schlossmann, C. R. Rose, A. Konnerth, S. Brummer, S. Kuhbandner, F. Hofmann, *Circ. Res.* **2002**, *90*, 1080 – 1086.
[53] M. Patel, A. Dutta, H. Huang, *Anal. Bioanal. Chem.* **2011**, *400*, 3035 – 3040.
[54] Q. Zhang, Y. Wang, X. Meng, R. Dar, H. Huang, *Anal. Chem.* **2013**, *85*, 201 – 207.
[55] M. Russwurm, F. Mullershausen, A. Friebe, R. Jager, C. Russwurm, D. Koesling, *Biochem. J.* **2007**, *407*, 69 – 77.

