Light-Activatable Nucleic Acids 'Caged' at the Nucleobases

Maximilian Buff, Timo Mack, and Alexander Heckel[§]*

[§]Grammaticakis-Neumann Prize Winner 2008

Abstract: The attachment of photolabile groups to biologically active molecules offers a very versatile way to put a biological effect under the control of an external trigger signal and thus confers spatiotemporal and dose control to this effect. Over the last years we have prepared a number of oligonucleotide derivatives that are modified in such a way. These derivatives are usually referred to as 'caged compounds' and in particular our photolabile 'caging groups' are located at the nucleobases of oligonucleotides so that the Watson–Crick interaction is temporarily impossible. Thus several nucleic acid-based applications have now become controllable with light.

Keywords: Aptamers · Caged compounds · Light activation · Light trigger · RNAi

Correspondence: Prof. Dr. A. Heckel
Goethe University Frankfurt
Cluster of Excellence Macromolecular Complexes
Max-von-Laue-Str. 9
D-60438 Frankfurt am Main
Tel.: + 49 69 798 29821
Fax: + 49 69 798 29823
E-mail: heckel@uni-frankfurt.de

Alexander Heckel was born in 1972 in Lindau (B) and studied chemistry at the University of Constance where he made his diploma thesis under the supervision of Prof. Dr. R. R. Schmidt in the field of oligosaccharide solid-phase synthesis. For his PhD thesis he joined the group of Prof. Dr. D. Seebach at the ETH Zurich and worked on the enantioselective heterogeneous catalysis with TADDOL and Salen on silica gel. During his postdoctoral period with Prof. Dr. P. B. Dervan at Caltech, US, he moved into the field of nucleic acid recognition and chemistry. With a Liebig fellowship (Fonds der Chemischen Industrie) and later an Emmy Noether fellowship (DFG) he became independent group leader under the mentorship of Prof. Dr. M. Famulok at the University of Bonn. Since this time he is interested in the control of biological processes with light and the assembly of nucleic acid nanoarchitectures using secondary structural elements. Alexander Heckel has received many fellowships and awards – recently for example the 'Dozentenstipendium' of the German Fonds der Chemischen Industrie and currently holds a position as professor for chemical biology and medicinal chemistry in the Cluster of Excellence Macromolecular Complexes at the Goethe University Frankfurt.

1. Introduction

When we study nature in the form of experiments the answers we get can only be as precise as the question we ask. While *in vitro* studies can be formidable tools to investigate a phenomenon without the influence of its context it is often exactly in this context that the phenomenon becomes meaningful for 'biological systems'.[1] On the other hand such biological systems are complex by definition and it would be desirable to prepare the system in a certain state and then instantly be able to generate a certain non-equilibrium condition and see how the system behaves. Also biological systems are often compartmentalized – like a cell, a tissue or an entire organism. This means that spatially addressable questions become equally important. It is exactly in these spatiotemporal contexts in living systems – but not only there – where induction methods for biological processes become interesting.

While there are quite a number of inducible systems known, light is a quite exquisite trigger signal. First of all it can be easily generated and manipulated with a huge set of well-established technologies – such as laser light sources and microscopes or endoscopes. Secondly, many of the biological systems which are currently studied are light-accessible *e.g.* the commonly used model organisms fruit fly, nematode worm or zebrafish, and even in higher organisms many external or internal surfaces can be reached with established technologies. Thirdly, if applied well it is possible to avoid phototoxic effects and most importantly light is an orthogonal trigger since the majority of biological systems do not already react to light themselves. Hence it can be applied at a time of arbitrary choice to a region which is freely selectable and the magnitude of the stimulus can potentially be controlled by carefully adjusting the light dose.

The first attempts to control biological processes with light date back to the late 1970s. At roughly the same time both Engels and Schlaeger[2] and the group of Hoffman[3] modified biologically active compounds with photolabile groups. The first chose cAMP (→**1**) as their target molecule while the latter used ATP (→**2**) and called their derivatives 'caged compounds' (Fig. 1). This choice of word is very unfortunate since it inevitably leads – in chemists and biologists alike these days – to the association of one molecule residing inside the cavity of another and it makes literature searches for this approach very difficult since the word 'caged' and its derivatives appear very frequently also in other contexts. Hence the past 30 years have seen a plethora of synonyms like 'light-triggered', 'light-inducible' *etc.* or even 'photocaged' which make literature searches even more laborious.

Until now many compounds have been caged and they mainly belonged to the domain of 'small molecules' and peptides or proteins.[4–9] While quite a number of outstanding studies have appeared in these
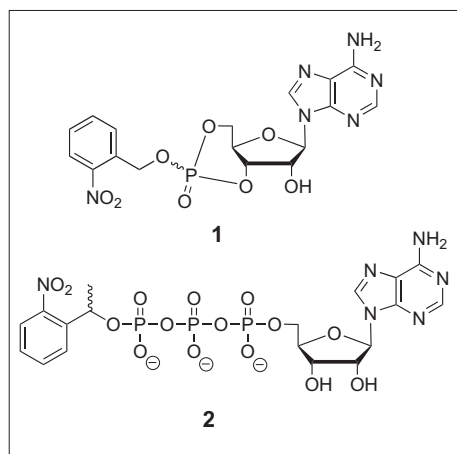
Fig. 1. cAMP (by Engels and Schlaeger[2]) and ATP (by Hoffman and coworkers[3]) were the first two biologically active molecules to be 'caged', *i.e.* of which derivatives with photolabile groups were synthesized with the intention of light-triggered release in biological systems.

Fig. 2. Early studies by other groups on caged DNA or RNA. (a) Pitsch et al.[10] (n = 1, R = adenine) and Chaulk and MacMillan[11] (n = 0, R = cytosine) introduced RNA derivatives with caging groups at the 2'-OH position. (b) Haselton, Monroe and coworkers,[13,14] Okamoto and coworkers[15] and the Friedman group[16] modified whole nucleic acids under benzylating conditions 'statistically'.

Fig. 3. (a) Transcription triggered with a light signal becomes possible if nucleobase-caged nucleotides are used in the promoter of a gene.[17] As demonstrated in the graph the system showed a clean ON/OFF behavior. (b) An siRNA with a caged residue at position ten of the guide strand is inactive until it is irradiated with light. The recovery of the activity is complete within error limits.[18]

fields there were only very few studies on caged nucleic acids until several years ago. These included for example work by Pitsch *et al.* or MacMillan and coworkers on a synthetic derivative of the hammerhead ribozyme with a photocleavable group at one of the 2'-OH groups (Fig. 2a).[10,11] Upon irradiation with light the activity of the ribozyme could be triggered. In a similar study MacMillan used this approach to study the assembly of the spliceosome complex.[12] Quite differently to this synthesis of well-defined nucleic acid derivatives other groups had chosen to subject (deoxy)nucleic acids to benzylating conditions using the reagents **3** or **4** (Fig 2b) in which photolabile groups were randomly attached at various positions. Thus for example Haselton, Monroe and coworkers[13,14] or the Okamoto group[15] modified plasmid DNA or mRNA and Friedman modified siRNAs.[16] In each case it was difficult to obtain a good ON/OFF-behavior which is most likely due to the fact that the modified positions in the nucleic acids were not well defined. Therefore we decided to make new derivatives of DNA and RNA which should bear photolabile groups at the nucleobases in such a way that the formation of Watson–Crick interactions becomes temporarily impossible. These types of derivatives can be seen as temporary mismatches. Since in almost all of the many applications which are known for nucleic acids the nucleobases play the central role it should be possible to put all these applications under the control of a light trigger signal.

2. Results and Discussion

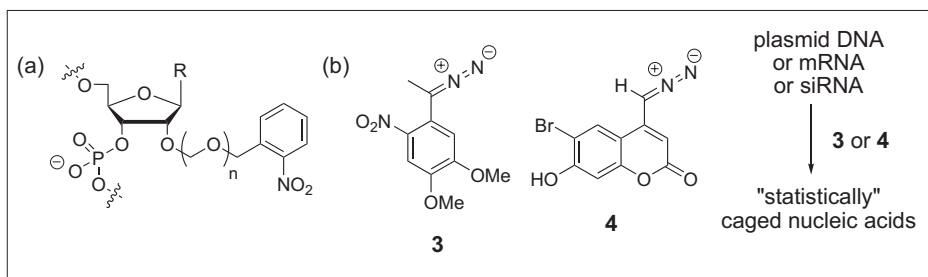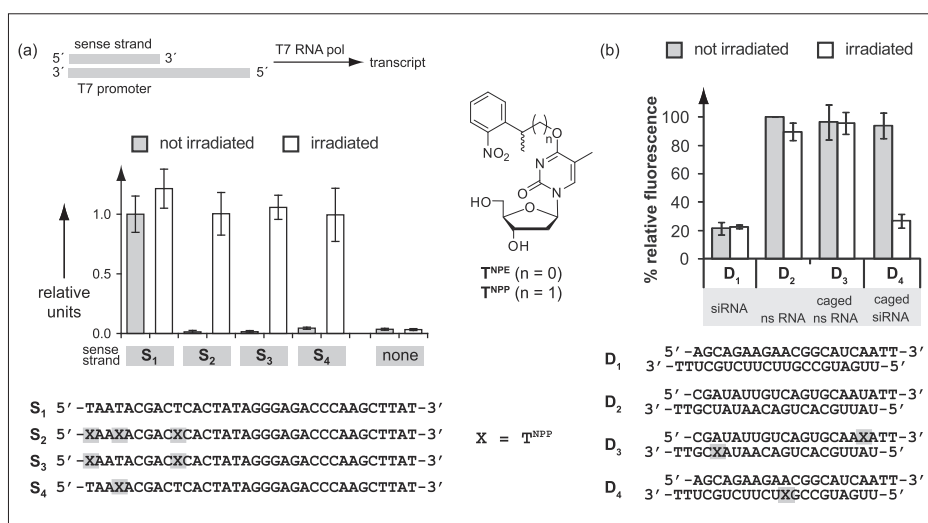The first derivatives we prepared were thymidine residues with caging groups derived from the 2-nitrobenzyl group (Fig. 3a).[17] The synthesis was very straightforward and involved the generation of a leaving group at the O[4] of thymidine. Apart from the commonly used NPE (1-(2-nitrophenyl)-1-ethyl) group we also used its homologous NPP derivative which had been introduced by Pfleiderer *et al.* and does not yield a harmful nitroso compound upon photolysis but rather a nitrostyryl derivative.[19] As one of the first studies we showed that it is possible to trigger transcription with a light signal. Therefore we chose the first nucleotides of a luciferase gene which was under the control of a T7 promoter. This allowed us to perform very simple *in vitro* transcription studies. The rationale was that the T7 promoter needs to be present in an unperturbed fashion in order for the T7 RNA polymerase to be able to recognize it and produce the RNA transcript. We introduced between one and five caging groups in the sense strand at various thymidines. The synthesis of the required DNA oligonucleotides turned out to be possible under standard conditions but due to the

inherent lability of the caging groups on the O[4] of thymidine towards ammonia it was necessary to use so-called 'ultramild' protecting groups on the other nucleobases. The problem is mainly the commonly used iso-butyroyl protecting group for guanosine which can only be removed with ammonia at elevated temperature. However, this is not a limitation to the technology since not only the 'ultramild' protecting groups turned out to be compatible but also the much cheaper dmf- or acetyl derivatives of guanosine and the acetyl derivatives of adenosine and cytidine. Melting point studies showed that the caged thymidine residues which we used were indeed severe mismatches. *In vitro* transcription experiments showed that only one caging group at the right residue position is enough to completely prevent transcription in this case but after irradiation with UV light (366 nm) the transcription could be fully recovered. This is also reflected by HPLC studies which clearly showed that the oligonucleotides could be cleanly and quantitatively uncaged when NPP was used as a protecting group.
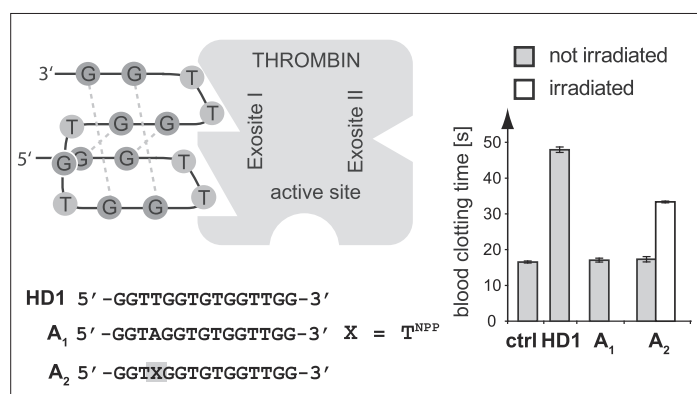
Fig. 4. The thrombin-binding aptamer **HD1** becomes light-inducible if a residue that is involved in the interaction is caged like in the derivative **A₂**.[25]

Fig. 5. The formation of G-quadruplex structures can be induced by light if the correct residue is caged – as for example in the sequence **T₂**.[28]

Encouraged by these results we proceeded to investigate if it is also possible to modify siRNAs in such a way that they are temporarily inactive. siRNAs are the key players in RNA interference which is a very versatile and general way to modify gene expression. For its discovery the Nobel Prize in Medicine or Physiology was awarded to Fire and Mello in 2006.[20,21] However, it is more complicated to develop inactive caged siRNAs since for most of the nucleotide residues the introduction of caging groups interferes with the siRNA's activity but does not turn it off. This is reasonable since siRNAs with mismatches are in fact rather miRNAs which can still interfere with gene expression, however in a different mechanism. But there was one particular study which had shown that the positions ten and eleven in the guide strand of an siRNA are very sensitive to modifications on the nucleobases.[22] This is also reasonable because these positions end up in the active center of the RISC complex next to the position where the mRNA is cleaved. On the other hand the same study showed that it is possible to have deoxynucleotides at these positions without any loss in siRNA activity. Therefore we chose these residues for the introduction of caged deoxynucleotides (Fig. 3b).[22] As test system an siRNA against EGFP was used and transfected into HeLa cells. The unmodified siRNA showed its effect both with irradiation and without. A caged or unmodified double stranded nonsense RNA never showed any effect – also not upon irradiation. This negative control experiment is important to discriminate from phototoxic effects. Finally the siRNA caged at the right position was completely inactive before activation but its activity could be completely recovered within error limits by irradiation with light.

Another very powerful application for nucleic acids is their use as aptamers. Aptamers are either DNA or RNA single strands that can fold into certain active conformations and bind to many target structures with usually high affinities and good selectivities and can hence be used for example as inhibitors.[23,24] Therefore we got interested in preparing caged aptamers and teamed up with the group of Günter Mayer at the University of Bonn who had been studying aptamers for quite some time. One extensively studied aptamer is the aptamer **HD1** (Fig. 4) – a 15mer single-stranded DNA molecule. It binds to α-thrombin, which is the key enzyme of the hemostatic network – cleaving fibrinogen to fibrin and thus initiating the blood clot formation. When **HD1** binds to the fibrinogen-recognizing exosite I domain of thrombin it inhibits the processing of fibrinogen. Since there were X-ray structures of the **HD1**/α-thrombin complex available,[26,27] this system was an ideal starting point for a rational approach that led to a caged light-controllable aptamer. We used again a **T^NPP** residue and introduced it into the aptamer at a position which was known to be vital for the interaction with thrombin from the X-ray structure (Fig. 4).[25] Blood clotting studies showed that caging of such an important thymidine residue indeed inhibited the interaction of the aptamer with its target molecule completely as does a point mutation at this position. But irradiation of the caged aptamer **A₂** with long wave UV light restored the thrombin binding and also the blood clotting inhibition activity – however in this case not to the full extent. This is due to the formation of a byproduct upon photolysis which seems to be favored in this sequence context.[19]

While the above-described approach based on a steric hindrance between the caged thymidine residue and the thrombin surface requires detailed knowledge of the interaction of the aptamer with its target molecule a more general approach would be to prevent the formation of the aptamer's active conformation. Therefore we became interested in the light-triggered formation of G-quadruplex structures.[28] To prove the possibility of trigg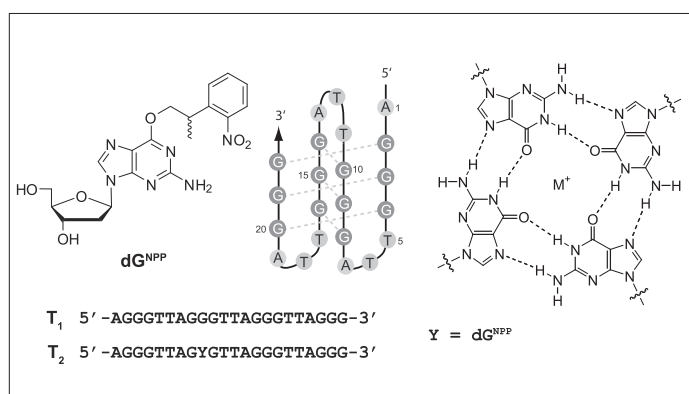ering the G-quadruplex formation with light we introduced an NPP group under *Mitsunobu* conditions in a dG nucleoside and used this **dG^NPP** residue for the solid-phase DNA synthesis of a caged telomeric DNA sequence and the **HD1** aptamer. The telomeric sequence **T₁** has been shown to fold into a three layer G-quadruplex compared to the two layer G-quadruplex of the **HD1** aptamer (Fig. 5). Different sequences with one, two or three **dG^NPP** residues introduced were tested with regard to their folding behavior in different buffers by CD spectroscopy. We found that in case of the three-layer G-quadruplex it is necessary to cage a dG residue that is both located in the core of the sequence and in the formation of the middle layer (such as in the sequence **T₂**, Fig. 5). With this set-up the telomeric sequence was prevented from folding into its correct shape. With respect to the **HD1** aptamer sequence it is again sufficient to modify only one nucleotide to obtain a caged version which then no longer folds into a G-quadruplex. In a very recent study we were able to show that the activity amplitude upon light triggering of a caged **HD1** aptamer can be significantly increased if it is fused with another aptamer domain targeting the Exosite II of thrombin.[29]

With the steric and the conformational approach we have demonstrated two ways to activate aptamers with light. But in terms of light-mediated regulation of biological activity it is equally attractive to be able to deactivate aptamers. For example if an aptamer works as an inhibitor, switching the aptamer off results in the recovery of the target protein's activity. To show that our approach of caged nucleic acids is also applicable for deactivation we developed a new variant of the thrombin binding aptamer **HD1**. Inspired by the work of Sullenger and coworkers[30,31] who used a separate antisense strand to form a duplex with an aptamer and thus destroyed the active conformation of the aptamer, we decided to extend the antithrombin aptamer and build a molecule with the aptamer region and the matching antisense region in the same
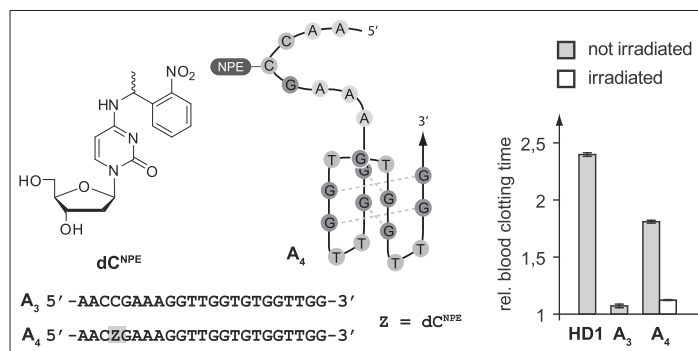
Fig. 6. An antithrombin aptamer with a caged 5'-antisense region refolds upon light activation and forms an inactive conformation. Thus the thrombin inhibiting function is turned OFF with light in this case.[32]

single strand nucleic acid (Fig. 6).[32] This has three advantages: The antidote activity is next to its target, the molecule offers the adequate amount of antisense strand by itself and there is only the necessity for one administration because the aptamer and its antidote come in the same molecule. The aptamer was elongated at the 5'-end with a GAAA-loop and four antisense nucleotides (A_3 in Fig. 6). In blood clotting experiments this molecule was complete inactive. Then we synthesized a light-activatable cytidine derivative (**dC^{NPE}**, Fig. 6) and replaced one dC residue in the antisense region with the caged variant, resulting in the aptamer A_4 (Fig. 6). In blood clotting experiments the aptamer A_4 shows a significant elongation of the blood clotting time before irradiation, albeit not as much as the progenitor molecule **HD1**. This is due to the fact that the attachment of the loop region already decreases the aptamer's affinity but can easily be compensated by a higher application dose. After irradiation with UV-light (366 nm) the aptamer A_4 is completely inactive.

Thus we were able to demonstrate in the past few years that interesting applications are possible for nucleobase-caged nucleic acids both in the domain of gene regulation as well as in the domain of the regulation of protein function. Since we started our program other groups have joined this endeavor or began to pursue complementary approaches. Among them there are

for example the groups of Schwalbe and Pitsch[33] and Silverman[34] who used nucleobase-caged nucleic acids to study folding processes. Dmochowski and coworkers presented for example a light-triggerable antisense strategy in zebrafish embryos.[35] Recently the Deiters group also used an antisense approach in mammalian cells.[36] A full account of competing and complementary research is beyond the scope of this article but we are currently preparing a book chapter in which there will be much more room for these discussions.

[1]   H. Kitano, *Science* **2002**, *295*, 1662.
[2]   J. Engels, E. J. Schlaeger, *J. Med. Chem.* **1977**, *20*, 907.
[3]   J. H. Kaplan, B. Forbush III, J. F. Hoffman, *Biochemistry* **1978**, *17*, 1929.
[4]   G. Mayer, A. Heckel, *Angew. Chem.* 2006, *118*, 5020; *Angew. Chem., Int. Ed.* **2006**, *45*, 4900.
[5]   X. Tang, I. J. Dmochowski, *Mol. Biosyst.* **2007**, *3*, 100.
[6]   G. C. R. Ellis-Davies, *Nat. Methods* **2007**, 4, 619.
[7]   D. D. Young, A. Deiters, *Org. Biomol. Chem.* **2007**, *5*, 999.
[8]   M. Goeldner, R. Givens, 'Dynamic Studies in Biology', Wiley-VCH, Weinheim, **2005**.
[9]   A. P. Pelliccioli, J. Wirz, *Photochem. Photobiol. Sci.* **2002**, *1*, 441.
[10] S. Pitsch, P. A. Weiss, X. L. Xu, D. Ackermann, T. Honegger, *Helv. Chim. Acta* **1999**, *82*, 1753.
[11] S. G. Chaulk, A. M. MacMillan, *Nucleic Acids Res.* **1998**, *26*, 3173.
[12] S. G. Chaulk, A. M. MacMillan, *Angew. Chem.* **2001**, *113*, 2207; *Angew. Chem., Int. Ed.* **2001**, *40*, 2149.
[13] W. T. Monroe, M. M. McQuain, M. S. Chang, J. S. Alexander, F. R. Haselton, *J. Biol. Chem.* **1999**, *274*, 20895.
[14] B. Ghosn, F. R. Haselton, K. R. Gee, W. T. Monroe, *Photochem. Photobiol.* **2005**, *81*, 953.
[15] H. Ando, T. Furuta, R. Y. Tsien, H. Okamoto, *Nat. Genet.* **2001**, *28*, 317.
[16] S. Shah, S. Rangarajan, S. H. Friedman, *Angew. Chem.* **2005**, *117*, 1352; *Angew. Chem., Int. Ed.* **2005**, *44*, 1328.
[17] L. Kröck, A. Heckel, *Angew. Chem.* 2005, *117*, 475; *Angew. Chem., Int. Ed.* **2005**, *44*, 471.
[18] V. Mikat, A. Heckel, *RNA* **2007**, *13*, 2341.
[19] S. Walbert, W. Pfleiderer, U. E. Steiner, *Helv. Chim. Acta* **2001**, *84*, 1601.
[20] A. Z. Fire, *Angew. Chem.* **2007**, *119*, 7094; *Angew. Chem., Int. Ed.* **2007**, *46*, 6966.
[21] C. C. Mello, *Angew. Chem.* **2007**, *119*, 7114; *Angew. Chem., Int. Ed.* **2007**, *46*, 6985.
[22] Y. L. Chiu, T. M. Rana, *RNA* **2003**, *9*, 1034.
[23] M. Famulok, J. S. Hartig, G. Mayer, *Chem. Rev.* **2007**, *107*, 3715.
[24] G. F. Joyce, *Angew. Chem.* **2007**, *119*, 6540; *Angew. Chem., Int. Ed.* **2007**, *46*, 6420.
[25] A. Heckel, G. Mayer, *J. Am. Chem. Soc.* **2005**, *127*, 822.
[26] K. Padmanabhan, A. Tulinsky, *Acta Cryst.* **1996**, *D52*, 272.
[27] K. Padmanabhan, K. P. Padmanabhan, J. D. Ferrara, J. E. Sadler, A. J. Tulinsky, *Biol. Chem.* **1993**, *268*, 17651.
[28] G. Mayer, L. Kröck, V. Mikat, M. Engeser, A. Heckel, *ChemBioChem* **2005**, *6*, 1966.
[29] G. Mayer, J. Müller, T. Mack, D. F. Freitag, T. Höver, B. Pötzsch, A. Heckel, *ChemBioChem* **2009**, in print.
[30] C. P. Rusconi, E. Scardino, J. Layzer, G. A. Pitoc, T. L. Ortel, D. Monroe, B. A. Sullenger, *Nature* **2002**, *419*, 90.
[31] C. P. Rusconi, J. D. Roberts, G. A. Pitoc, S. M. Nimjee, R. R. White, G. Quick, Jr., E. Scardino, W. P. Fay, B. A. Sullenger, *Nat. Biotechnol.* **2004**, *22*, 1423.
[32] A. Heckel, M. C. R. Buff, M. S. L. Raddatz, J. Müller, B. Pötzsch, G. Mayer, *Angew. Chem.* **2006**, *118*, 6900; *Angew. Chem., Int. Ed.* **2006**, *45*, 6748.
[33] P. Wenter, B. Fürtig, A. Hainard, H. Schwalbe, S. Pitsch, *Angew. Chem.* **2005**, *117*, 2656; *Angew. Chem., Int. Ed.* **2005**, *44*, 2600.
[34] C. Höbartner, S. K. Silverman, *Angew. Chem.* **2005**, *117*, 7471; *Angew. Chem., Int. Ed.* **2005**, *44*, 7305.
[35] X. Tang, S. Maegawa, E. S. Weinberg, I. J. Dmochowski, *J. Am. Chem. Soc.* **2007**, *129*, 11000.
[36] D. D. Young, H. Lusic, M. O. Lively, J. A. Yoder, A. Deiters, *ChemBioChem* **2008**, *9*, 2397.