VIP Modified Nucleotides Very Important Paper

International Edition: DOI: 10.1002/anie.201604641
German Edition: DOI: 10.1002/ange.201604641

Sequence-Specific Incorporation of Enzyme–Nucleotide Chimera by DNA Polymerases

*Moritz Welter, Daniela Verga, and Andreas Marx**

Abstract: *DNA polymerases select the right nucleotide for the growing polynucleotide chain based on the shape and geometry of the nascent nucleotide pairs and thereby ensure high DNA replication selectivity. High-fidelity DNA polymerases are believed to possess tight active sites that allow little deviation from the canonical structures. However, DNA polymerases are known to use nucleotides with small modifications as substrates, which is key for numerous core biotechnology applications. We show that even high-fidelity DNA polymerases are capable of efficiently using nucleotide chimera modified with a large protein like horseradish peroxidase as substrates for template-dependent DNA synthesis, despite this "cargo" being more than 100-fold larger than the natural substrates. We exploited this capability for the development of systems that enable naked-eye detection of DNA and RNA at single nucleotide resolution.*

DNA polymerases have evolved for copying genetic information in an accurate manner. They recognize the template strand and choose the canonical nucleotide from the pool of four 2′deoxy nucleotide triphosphates (dNTPs) for incorporation with high accuracy. The nucleobase selection is believed to be based on editing of the shape and geometry of the nascent nucleotide pairs building the basis for DNA replication selectivity.[1] High-fidelity DNA polymerases are thus believed to possess tight active sites that allow little deviation from the canonical structures. It is thus puzzling that DNA polymerases are capable of using modified nucleotides as substrates. These characteristics are key for employing DNA polymerases in numerous core biotechnology applications such as sequencing and diagnostic applications.[2]

To gain insight into the mechanisms by which DNA polymerases are able to process modified nucleotides despite their active sites being strictly defined, functional and structural studies have been performed. It has been shown that nucleotides bearing small-molecule modifications such as affinity,[3] spin,[4] and redox labels;[5] dyes;[2a] amino acids;[6] nucleosides;[7] or short oligonucleotides[8] can be used as substrates for DNA replication by DNA polymerases.

We now show that even protein-modified dNTPs, despite their size, can be used as substrates for the replication of nucleic acid targets by high-fidelity DNA polymerases (Figure 1 a). The protein, in this case the approximately 40 kDa glycoprotein horseradish peroxidase (HRP) from *Amoracia rusticana*,[9] was site specifically incorporated into the nascent DNA chain without compromising the ability of the enzyme to produce colorimetric signals through the oxidation of dye substrates. This enzyme is active even under harsh environmental conditions, such as elevated temperature, a wide range of pH values, and chaotropic reagents, which makes it one of the most frequently used reporter enzymes biochemical detection assays.[10] On this basis, we exploited the capability

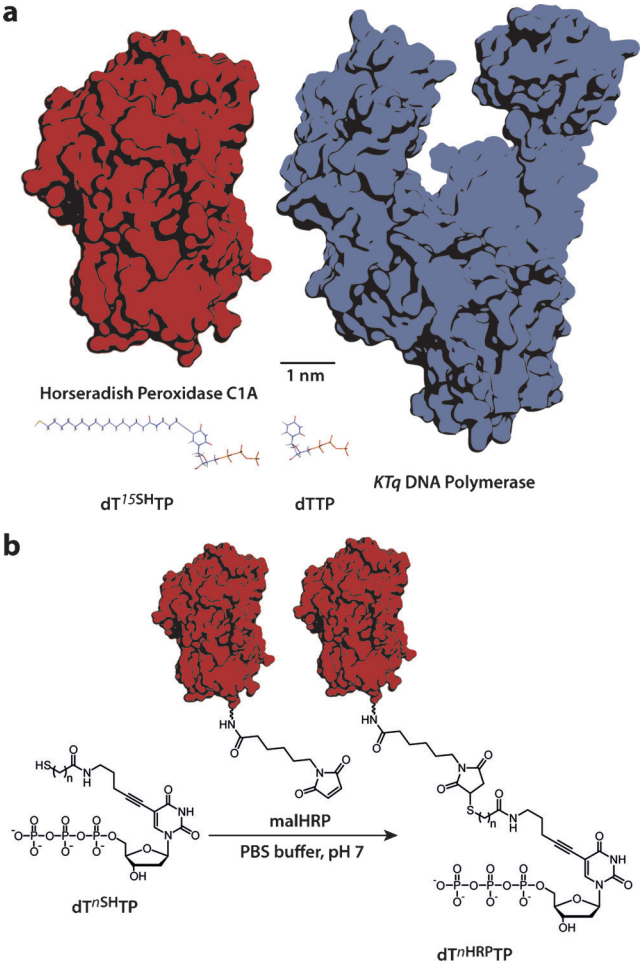

Figure 1. Scheme of the conjugation strategy and the relative proportions of the structures used in this study. A) Size comparison of *KTq* DNA Polymerase (PDB ID: 1KTQ), HRP C1A (PDB ID: 1HCH), dTTP, and a modified dTTP bearing a C_{15} thiol linker ($dT^{15SH}TP$) to scale. b) Coupling of $dT^{nSH}TP$ and malHRP (derived from HRP C1A) through a thiol–maleimide reaction.

[*] M. Sc. M. Welter, Dr. D. Verga, Prof. Dr. A. Marx
Department of Chemistry
Konstanz Research School Chemical Biology
University of Konstanz
Universitätsstraße 10, 78457 Konstanz (Germany)
E-mail: andreas.marx@uni-konstanz.de

Supporting information for this article can be found under: http://dx.doi.org/10.1002/anie.201604641.

of DNA polymerases to use HRP-modified dNTPs as substrate for the development of systems that enable the naked-eye detection of DNA and RNA at single-nucleotide resolution.

HRP species that have been activated with one to three maleimide groups through conversion of the lysine residues with maleimidocaproic acid *N*-hydroxysuccinimide ester (malHRP) are commercially available. These species allow conjugation to a target molecule through a thiol–maleimide reaction. Since this reaction is one of the fastest bioconjugation reactions in aqueous conditions, it promises fast conversion and high yields.[11] Thus, we envisioned a thiol-functionalized nucleotide for conjugation to the maleimide-activated enzyme.

Previous studies have shown that C5-deoxythymidine derivatives are often well accepted by DNA polymerases.[12] In order to study the influence of linker length on the incorporation efficiency, we synthesized two thymidine analogues bearing ω-mercaptocarboxylic acid based linkers of different lengths at the C5 position (dT$^{n\mathrm{SH}}$TP; Figure S1 in the Supporting Information). To do so, we followed the synthetic pathway that has already been reported for hydroxy-functionalized nucleotides, starting with a Sonogashira coupling of 5-trifluoroacetamidopentyne with 5-Iodo-2′-deoxyuridine (Figure S1 a) to yield the TFA-protected, amino-functionalized nucleoside **1**.[8a] The nucleoside was then converted into the triphosphate (Figure S1 b) and deprotected in an ammonium hydroxide solution to yield compound **2**. The ω-mercaptocarboxylic acid was finally introduced using the coupling reagent O-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HATU) together with *N*,*N*-diisopropylethylamine (DIPEA) in *N*,*N*-dimethylformamide (DMF) to yield the final products dT$^{7\mathrm{SH}}$TP and dT$^{15\mathrm{SH}}$TP (Figure S1 c; see Figures S10, S11 for NMR spectra). Conjugation to the enzyme was performed by simple incubation with malHRP in PBS buffer (Figure 1 b). Subsequently, the conjugates (dT$^{n\mathrm{HRP}}$TP) were purified through anion-exchange fast protein liquid chromatography (FPLC) and the conjugation was verified by ESI-MS (Figure S2).

To examine whether the synthesized conjugates are substrates for DNA polymerases, we performed primer-extension reactions on a template containing the B-type Raf kinase (BRAF) T1796A point mutation, which is strongly associated with carcinogenesis (Figure 2 a, see the Supporting Information for sequences).[13] The 21-nucleotide (nt) 5′-radioactively labeled primer was designed to end directly 5′ of the mutation site so that a single deoxythymidine analogue could be inserted. This insertion was subsequently analyzed by denaturing polyacrylamide gel electrophoresis (PAGE) and autoradiography. When employing the unconjugated nucleotides dT$^{7\mathrm{SH}}$TP and dT$^{15\mathrm{SH}}$TP, we observed the expected band of retarded mobility, thus confirming successful insertion of the bulky nucleotides by the *KlenTaq* DNA polymerase (*KTq*; Figure 2 b, Lanes 2 and 3; see Figure S3 for the full gel). The misshapen appearance of the band compared to that of natural dTTP (Lane 1) might be caused by interaction of the free thiol with the gel matrix. Interestingly, it appears that for the unconjugated nucleotides, the one with the smaller linker (dT$^{7\mathrm{SH}}$TP) is accepted better by the polymerase.

When employing the conjugates dT$^{7\mathrm{HRP}}$TP and dT$^{15\mathrm{HRP}}$TP (Lanes 4 and 5), we observed a vastly retarded migration of the band corresponding to the primer (approximately equivalent of 250 nt), thus indicating successful incorporation of the nucleotide with its protein "cargo". We therefore concluded that enzyme-labeled nucleotides, despite their size, are indeed accepted as DNA polymerase substrates. This finding is also supported by sodium dodecylsulfate (SDS)-PAGE of a primer extension reaction (Figure S4), where a shift of the protein band towards higher molecular weights could be observed. Interestingly, we found that the conjugate connected via C_{15} linker is better accepted than the

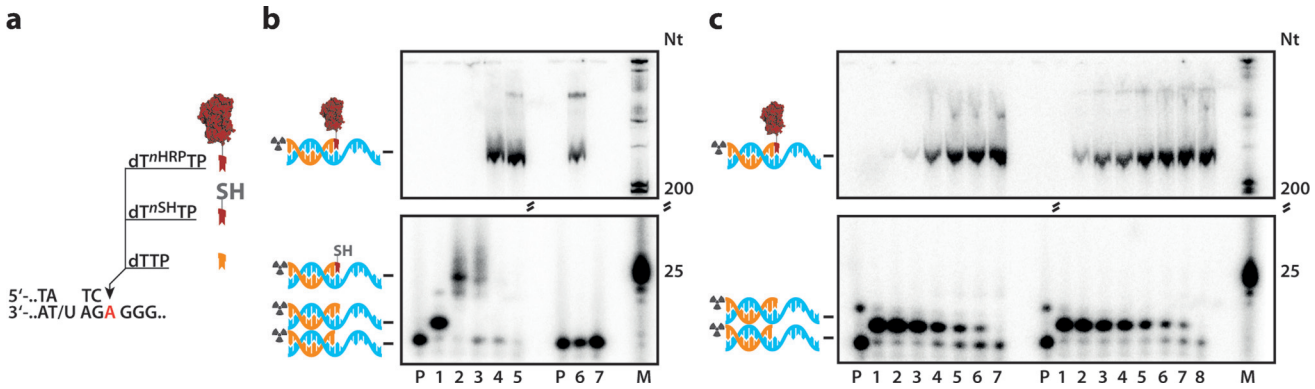

Figure 2. PAGE of primer extension reactions employing the synthesized conjugates. a) Partial sequence of the primer and the template at the incorporation site. b) Left: Autoradiography of a primer extension reaction employing 1 μM of natural dTTP (Lane 1), one of the two thiol-modified nucleotides (dT$^{7\mathrm{SH}}$TP and dT$^{15\mathrm{SH}}$TP, Lanes 2 and 3), or one of the two conjugates dT$^{7\mathrm{HRP}}$TP and dT$^{15\mathrm{HRP}}$TP (Lanes 4 and 5) with the *KTq* DNA polymerase on a template containing the BRAF T1796A point mutation site, Samples were collected after 10 min. Right: Primer extension with *RT-KTq* DNA polymerase on an RNA template of the BRAF point mutation sequence using dT$^{15\mathrm{HRP}}$TP (Lane 6) and in the absence of the template (Lane 7). Samples were collected after 10 min. P: Primer, M: Marker, Nt: Nucleotides. See Figure S3 for the full PAGE analysis. c) Incorporation competition experiments between the dT$^{n\mathrm{HRP}}$TP and dTTP with the *KTq* DNA polymerase. Ratios applied for dT$^{7\mathrm{HRP}}$TP (left) were 1) 0:1, 2) 1:1, 3) 3:1, 4) 19:1, 5) 49:1, 6) 99:1 and 7) 1:0. Ratios applied for dT$^{15\mathrm{HRP}}$TP (right) were 1) 0:1, 2) 1:1, 3) 3:1, 4) 4:1, 5) 9:1, 6) 19:1, 7) 99:1 and 8) 1:0. Samples were collected after 3 min. See Figure S3 for the PAGE analysis.

one with the shorter linker. We even observed another retarded band, which indicates a second incorporation. Therefore, we conducted primer extension on a template encoding eleven consecutive dTMP insertions to evaluate the propensity of the conjugates for multiple incorporation (Figures S5, S6, right). We found that the conjugate bearing the longer linker can be successively incorporated into a primer strand while no multiple incorporation for the shorter conjugate was observed, thus indicating that too little distance between the nucleotide and its "cargo" interferes with DNA polymerase activity.

For dT^{15HRP}TP, we then investigated the possibility of periodic incorporation into a primer strand (Figure S6). The incorporation of up to four conjugates could be observed by 9% denaturating PAGE (dPAGE) using 8,3M Urea, with the gel limiting the resolution due to highly retarded migration. Separation on less dense gels however, led to undefined bands, thereby impeding analysis. To estimate the incorporation efficiency and the influence of linker length, we carried out competition assays with the conjugates and natural dTTP (Figure 2c and Figure S7). We performed single-nucleotide incorporation experiments, with the modified nucleotides directly competing for incorporation with their natural counterparts. This experimental setup has previously been used for a similar purpose.[14] When employing *KTq* DNA polymerase, we found that the unmodified dTTP is 33-fold better processed than dT^{7HRP}TP (Figure 2c, left) and only 6-fold better in comparison to dT^{15HRP}TP (Figure 2c, right). This nicely demonstrates that the linker length has a substantial effect on polymerase acceptance by increasing the distance between the polymerase and the cargo. Furthermore, we also demonstrated that the exonuclease-deficient Klenow Fragment from *E. Coli* DNA Polymerase I (*KF exo*$^-$) is capable of using the modified nucleotides, resulting in a 68-fold diminished efficiency for dT^{7HRP}TP and 21-fold for dT^{15HRP}TP (Figure S8).

Finally, we investigated the modified nucleotide in reverse transcription and tested the *RT-KTq* DNA polymerase, a *KTq* DNA polymerase mutant with reverse transcriptase activity. This DNA polymerase is able to elongate a DNA primer strand using both, DNA and RNA templates.[15] We therefore conducted the same primer extension experiment, using the same RNA sequence context for reverse transcription (Figure 2b, Lanes 6 and 7; see the Supporting Information for sequences). Since dT^{15HRP}TP is processed with higher efficiency, we focused on this nucleotide. We found the same characteristic shift of primer mobility by PAGE analysis as described above, thus indicating that HRP-modified nucleotides can be applied to reverse transcription as well.

The conjugation of HRP to oligonucleotides has been exploited in the past to provide a colorimetric system for the detection of nucleic acids.[16] However, covalent attachment to the 5′ or 3′ end of an oligonucleotide allows hybridization-based approaches only, which have significant disadvantages in terms of sensitivity and selectivity.[17] The enzyme-modified nucleotides introduced herein have the potential to combine the colorimetric read-out of peroxidase-based assays with the versatility and fidelity of a DNA polymerase reaction.

In order to explore the potential of protein-modified nucleotides, we designed an assay to exploit primer extensions of solid-phase-immobilized primers strands (Figure 3a). It was envisioned that due to the template dependency of DNA polymerases, only primers that bind sequence-selectively to a target sequence will be extended by elongation with a HRP-modified nucleotide. So the presence or absence of the target should be detectable (Figure 3a, match and absence). After the removal of excess nucleotide–protein conjugate, the addition of the HRP substrate will result in a colorimetric read-out.[8b,c] We carried out the assay, applying a 5′-biotin-immobilized version of the BRAF sequence context used for

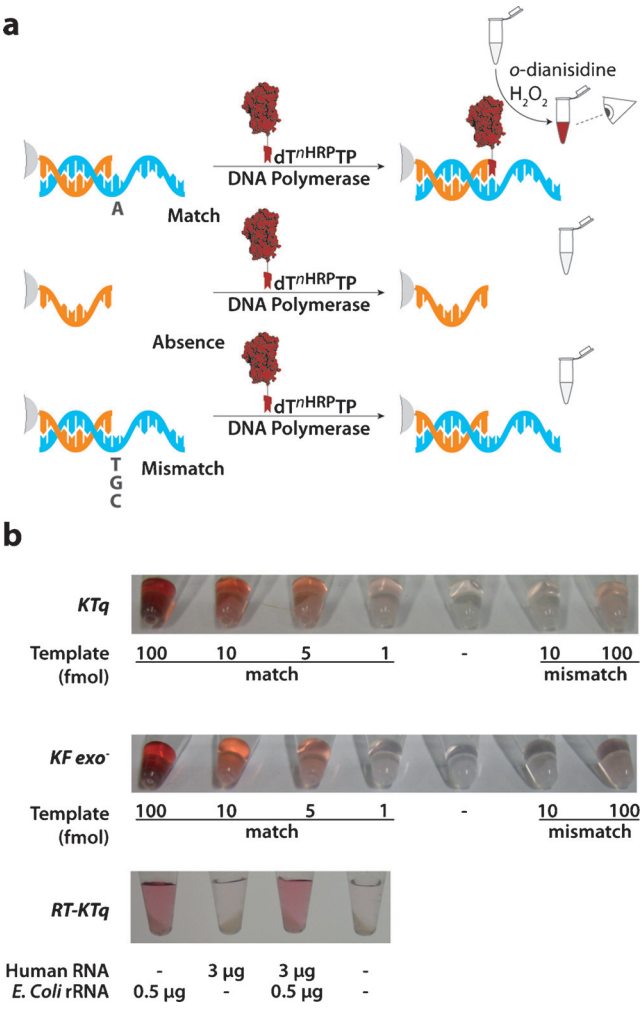

Figure 3. A possible application for the conjugates in a naked-eye detection assay. a) Scheme of the naked-eye detection assay employing the enzyme-labeled nucleotides. The primer is immobilized on a solid support via the biotin–streptavidin interaction. After annealing of the template sequence, DNA polymerase and conjugates are added. After the incubation period, the unbound conjugate is removed by repeated filtration and a dye solution is added to give a colorimetric read-out visible to the naked eye. b) Results of the assay depicted with *KTq* and *KF exo*$^-$ DNA polymerase on a matched or mismatched ssDNA template and *RT-KTq* DNA polymerase on an *E. coli* ribosomal rRNA mixture in the presence and absence of an excess of human RNA. For the assay, 1 µM of dT^{15HRP}TP was incubated in a primer-extension reaction mixture for 5 min (*KTq*/*RT-KTq*) or 15 min (*KF exo*$^-$ DNA polymerase).

the primer extension in solution. When employing dT^{15HRP}TP, we were able to reliably detect as little as 1 fmol of DNA through *KTq* DNA polymerase promoted incorporation and subsequent HRP reaction (Figure 3b and Figure S9).

Aside from detection of the presence or absence of a target sequence, the intrinsic fidelity of the DNA polymerase to produce Watson–Crick base pairs has the potential to be exploited to discriminate single-nucleotide variations at the insertion site (Figure 3a, match and mismatch). We tested this application by employing a mismatched template sequence bearing a thymidine residue at the insertion site, which leads to a T:T mismatch. A strong discrimination between the matched and mismatched template was observed with *KTq* DNA polymerase (Figure 3b, *KTq* match/mismatch) which made it possible to distinguish between 5 fmol of the matched and a 20-fold excess (100 fmol) of the mismatched template.

Next, we carried out an assay for the detection of single nucleotide variations with *KF exo*$^-$ DNA polymerase since this enzyme is active at room temperature, which is a beneficial property for point-of-care testing (POCT). We measured a detection limit of 1 fmol when *KF exo*$^-$ DNA polymerase was used. This enzyme was also able to distinguish between matched and mismatched templates (Figure 3b, see Figure S9 for quantification), thus demonstrating the selectivity of this approach.

In order to extend the naked-eye detection system to RNA diagnostics, for example, for applications in pathogen detection, we also performed the assay using *RT-KTq* DNA polymerase. We employed a primer complementary to a sequence in the *E. Coli* 16S rRNA, carrying out the assay in the presence and absence of a 6-fold excess of human total RNA. In this way, we were able to reliably detect 0.5 μg of an *E. Coli* rRNA mixture in less than 10 minutes without any interference from the human RNA (Figure 3b, *RT-KTq*). It is important to note that the experiment was carried out without any annealing steps although the primer was designed to bind to a partially double-stranded segment of the 16S rRNA. The amount of template sequence bound to the primer (and detected) might therefore be considerably lower than the amount added. We also investigated the lower detection limit for the *RT-KTq* DNA polymerase-promoted reaction and measured a detection limit of 1 fmol, which is similar to that for DNA (Figure S10).

It has been shown that terminal deoxynucleotidyl transferase is able to ligate nucleotides modified with a protein at the 3′ end of single-stranded DNA without any template control.[18] Also, efficient methods for the post-incorporation labeling of modified nucleotides with proteins have been reported.[19] Herein, we report on the direct template dependent incorporation of protein-modified nucleotides by high-fidelity DNA polymerases as a new way to generate site-specific DNA–protein conjugates. Although the cargo protein (ca. 43000 Å^3) is several orders of magnitude larger than the natural substrates (ca. 400 Å^3, see the Supporting Information for calculations), we show that protein-modified nucleoside triphosphates are substrates for DNA polymerases. This is exemplarily shown for HRP-modified nucleotides. The HRP-conjugated nucleotides are incorporated into the template in a sequence-specific manner when the DNA or RNA targets are present. Moreover, the HRP catalytic activity is not abolished by the conjugation and can thus be exploited for colorimetric read-out of the nucleotide incorporation that is detectable by the naked eye. This naked-eye assay offers major advantages compared to previously reported approaches. First, the conjugation of the enzyme to the nucleotide generates a versatile tool that bears the potential to be easily applied to any sequence context without having to change the general setup. Second, the utilization of DNA polymerases for genotyping gives a high level of discrimination between matched and mismatched templates. Finally, by using mesophilic DNA polymerases, the assay can be carried out without any sophisticated laboratory equipment, like thermocyclers, which makes it suited for in field analysis and POCT. The possibility to incorporate nucleotides that are covalently attached to whole proteins by using DNA polymerases is thus highly promising for future applications.[20]

Acknowledgements

We gratefully acknowledge support by the DFG, the Alexander von Humboldt-Stiftung (stipend to D.V.) and the Konstanz Research School Chemical Biology.

Keywords: colorimetric analysis · DNA polymerase · DNA replication · horseradish peroxidase · modified nucleotides

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 10131–10135
Angew. Chem. **2016**, *128*, 10286–10290

[1] a) T. W. Kim, J. C. Delaney, J. M. Essigmann, E. T. Kool, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15803–15808; b) E. T. Kool, H. O. Sintim, *Chem. Commun.* **2006**, 3665–3675.
[2] a) D. R. Bentley, S. Balasubramanian, H. P. Swerdlow, G. P. Smith, J. Milton, C. G. Brown, K. P. Hall, D. J. Evers, C. L. Barnes, H. R. Bignell, et al., *Nature* **2008**, *456*, 53–59; b) T. D. Harris, P. R. Buzby, H. Babcock, E. Beer, J. Bowers, I. Braslavsky, M. Causey, J. Colonell, J. Dimeo, J. W. Efcavitch, et al., *Science* **2008**, *320*, 106–109; c) S. Ishino, Y. Ishino, *Front. Microbiol.* **2014**, *5*, 465.
[3] T. Ohbayashi, M. Kuwahara, M. Hasegawa, T. Kasamatsu, T. Tamura, H. Sawai, *Org. Biomol. Chem.* **2005**, *3*, 2463–2468.
[4] S. Obeid, M. Yulikov, G. Jeschke, A. Marx, *Angew. Chem. Int. Ed.* **2008**, *47*, 6782–6785; *Angew. Chem.* **2008**, *120*, 6886–6890.
[5] J. Balintova, J. Spacek, R. Pohl, M. Brazdova, L. Havran, M. Fojta, M. Hocek, *Chem. Sci.* **2015**, *6*, 575–587.
[6] M. Kitaoka, Y. Tsuruda, Y. Tanaka, M. Goto, M. Mitsumori, K. Hayashi, Y. Hiraishi, K. Miyawaki, S. Noji, N. Kamiya, *Chem. Eur. J.* **2011**, *17*, 5387–5392.
[7] A. R. Kore, *Tetrahedron Lett.* **2009**, *50*, 793–795.
[8] a) A. Baccaro, A. L. Steck, A. Marx, *Angew. Chem. Int. Ed.* **2012**, *51*, 254–257; *Angew. Chem.* **2012**, *124*, 260–263; b) D. Verga, M. Welter, A. L. Steck, A. Marx, *Chem. Commun.* **2015**, *51*, 7379–7381; c) D. Verga, M. Welter, A. Marx, *Bioorg. Med. Chem. Lett.* **2016**, *26*, 841–844.
[9] L. M. Shannon, E. Kay, J. Y. Lew, *J. Biol. Chem.* **1966**, *241*, 2166–2172.
[10] a) J. Everse, K. E. Everse, M. B. Grisham, *Peroxidases in chemistry and biology*, CRC Press, Boca Raton, Florida etc., **1991**; b) J. W. Tams, K. G. Welinder, *FEBS Lett.* **1998**, *421*, 234–236.

[11] F. Saito, H. Noda, J. W. Bode, *ACS Chem. Biol.* **2015**, *10*, 1026 – 1033.
[12] a) M. Hocek, *J. Org. Chem.* **2014**, *79*, 9914 – 9921; b) A. Hottin, A. Marx, *Acc. Chem. Res.* **2016**, *49*, 418 – 427.
[13] H. Davies, G. R. Bignell, C. Cox, P. Stephens, S. Edkins, S. Clegg, J. Teague, H. Woffendin, M. J. Garnett, W. Bottomley, et al., *Nature* **2002**, *417*, 949 – 954.
[14] a) S. Obeid, A. Baccaro, W. Welte, K. Diederichs, A. Marx, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 21327 – 21331; b) K. Bergen, A. L. Steck, S. Strutt, A. Baccaro, W. Welte, K. Diederichs, A. Marx, *J. Am. Chem. Soc.* **2012**, *134*, 11840 – 11843.
[15] N. Blatter, K. Bergen, O. Nolte, W. Welte, K. Diederichs, J. Mayer, M. Wieland, A. Marx, *Angew. Chem. Int. Ed.* **2013**, *52*, 11935 – 11939; *Angew. Chem.* **2013**, *125*, 12154 – 12158.
[16] a) M. S. Urdea, B. D. Warner, J. A. Running, M. Stempien, J. Clyne, T. Horn, *Nucleic Acids Res.* **1988**, *16*, 4937 – 4956; b) S. S. Ghosh, P. M. Kao, A. W. McCue, H. L. Chappelle, *Bioconjugate Chem.* **1990**, *1*, 71 – 76; c) R. P. van Gijlswijk, M. P. van de Corput, V. Bezrookove, J. Wiegant, H. J. Tanke, A. K. Raap, *Histochem. Cell Biol.* **2000**, *113*, 175 – 180.
[17] M. Strerath, A. Marx, *Angew. Chem. Int. Ed.* **2005**, *44*, 7842 – 7849; *Angew. Chem.* **2005**, *117*, 8052 – 8060.
[18] R. S. Sørensen, A. H. Okholm, D. Schaffert, A. L. B. Kodal, K. V. Gothelf, J. Kjems, *ACS Nano* **2013**, *7*, 8098 – 8104.
[19] a) J. Dadová, P. Orság, R. Pohl, M. Brázdová, M. Fojta, M. Hocek, *Angew. Chem. Int. Ed.* **2013**, *52*, 10515 – 10518; *Angew. Chem.* **2013**, *125*, 10709 – 10712; b) J. Dadová, M. Vrábel, M. Adámik, M. Brázdová, R. Pohl, M. Fojta, M. Hocek, *Chem. Eur. J.* **2015**, *21*, 16091 – 16102.
[20] C. M. Niemeyer, *Angew. Chem. Int. Ed.* **2010**, *49*, 1200 – 1216; *Angew. Chem.* **2010**, *122*, 1220 – 1238.

Received: May 12, 2016
Published online: July 8, 2016

