CHEMISTRY
A EUROPEAN JOURNAL

ChemPubSoc
Europe

DOI: 10.1002/chem.201302972

Three-Pronged Probes: High-Affinity DNA Binding with Cap, β-Alanines and Oligopyrrolamides

Rüdiger Haug, Markus Kramer, and Clemens Richert*[a]

Sequence-specific recognition of DNA strands is most readily achieved through hybridization with a complementary probe strand, designed to bind following the Watson–Crick base pairing rules.[1] But, this binding mode interrogates only a very small fraction of the surface of a target strand. A limited sequence fidelity results, with a small penalty for mismatches in the duplex.[2,3] In contrast, active sites of enzymes typically bind a large fraction of the surface of substrates, and so do key proteins that act on DNA during replication or repair.[4,5] Given the importance of sequence-specific recognition of DNA, it is desirable to develop strategies for more thoroughly interrogating target sequences with synthetic probes, using a combination of base pairing, interactions based on shape complementarity, and hydrogen bonding in the grooves. The need for high-affinity probes is particularly urgent for sequences rich in weakly pairing bases (A and T) that are difficult to bind strongly with DNA alone.

Substituents that increase the affinity of oligonucleotides for target strands through additional stacking interactions and/or preorganizing effects have been described.[6–8] Further, elegant work exists on cyclic probes that use a combination of Watson–Crick and Hoogsteen base pairing[9] and that target triplex-forming sequences. Two-pronged probes can be constructed from oligonucleotides and a minor groove binder.[10] Another approach combines base pairing with stacking on the terminus of the duplex with the target. For example, we have previously shown that non-nucleosidic substituents at the termini of oligonucleotides can act as "caps" that stabilize duplexes with canonical base pairs at their termini.[11,12] Caps can give significant increases in affinity and base-pairing fidelity, but their effect fades with the distance from the terminus. A probe featuring a stilbene cap and a minor groove binder was prepared previously, but it gave disappointing mismatch discrimination at the terminus,[13] suggesting that cap and MGB did not act with proper synergy in binding the target. So, a seamless recognition of a target from three sides (Watson–Crick face, terminus, and minor groove) remained a challenge. Here we report three-pronged probes (TPPs) that bind A/T-rich target strands from three sides. The three-pronged approach results in exceptionally stable duplexes (ΔT_m of up to +44.8 °C) and high selectivity, even at the very terminus (ΔT_m of −6.4 °C for a terminal wobble base pair). Control experiments suggest that all portions of the TPP can engage in target binding, including the linker.

A high-resolution structure of a duplex with the Uaq cap[14] provided the starting point for the design of the three-pronged probes (Figure 1). The terminus is tightly capped by the anthraquinone moiety, and the uracil residue of the linker is found in the minor groove. In fact, the O4 atom of U points straight into the groove, making it an ideal attachment point for a minor groove binder (MGB). Since the carbonyl oxygen of the pyrimidine cannot be reacted itself without creating an undesirable lactim ether structure, N4 of the corresponding cytidine derivative was chosen for appending the MGB. A DNA sequence was selected for the probes that matches the preference of the oligopyrrolamides to bind A/T-rich duplexes.[15] Since individual β-alanine residues have been shown to be tolerated in minor groove binders,[16] we chose β-Ala-based linkers, hoping that they would provide the necessary flexibility.

Scheme 1 shows key intermediates of the synthesis of three-pronged probes **1**–**7**. Oligoamides **8**–**10**, featuring four pyrroles and from one to four β-alanine residues, were syn-

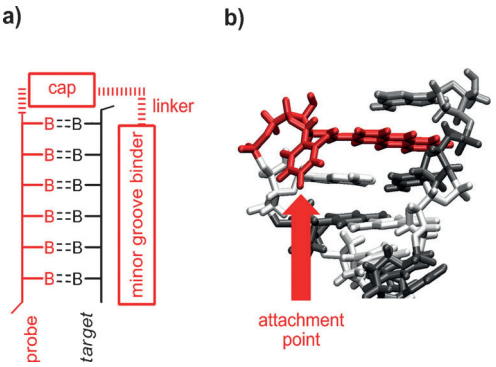


Figure 1. a) Principle of binding a DNA target strand with a three-pronged probe through molecular interrogation from three sides. b) Three-dimensional structure of the DNA duplex $(ACGCG\text{-}Uaq)_2$[14] with O4 of the uridine residue highlighted by an arrow. The anthraquinone moiety acting as cap and the uridine residue acting as linker are shown in red.

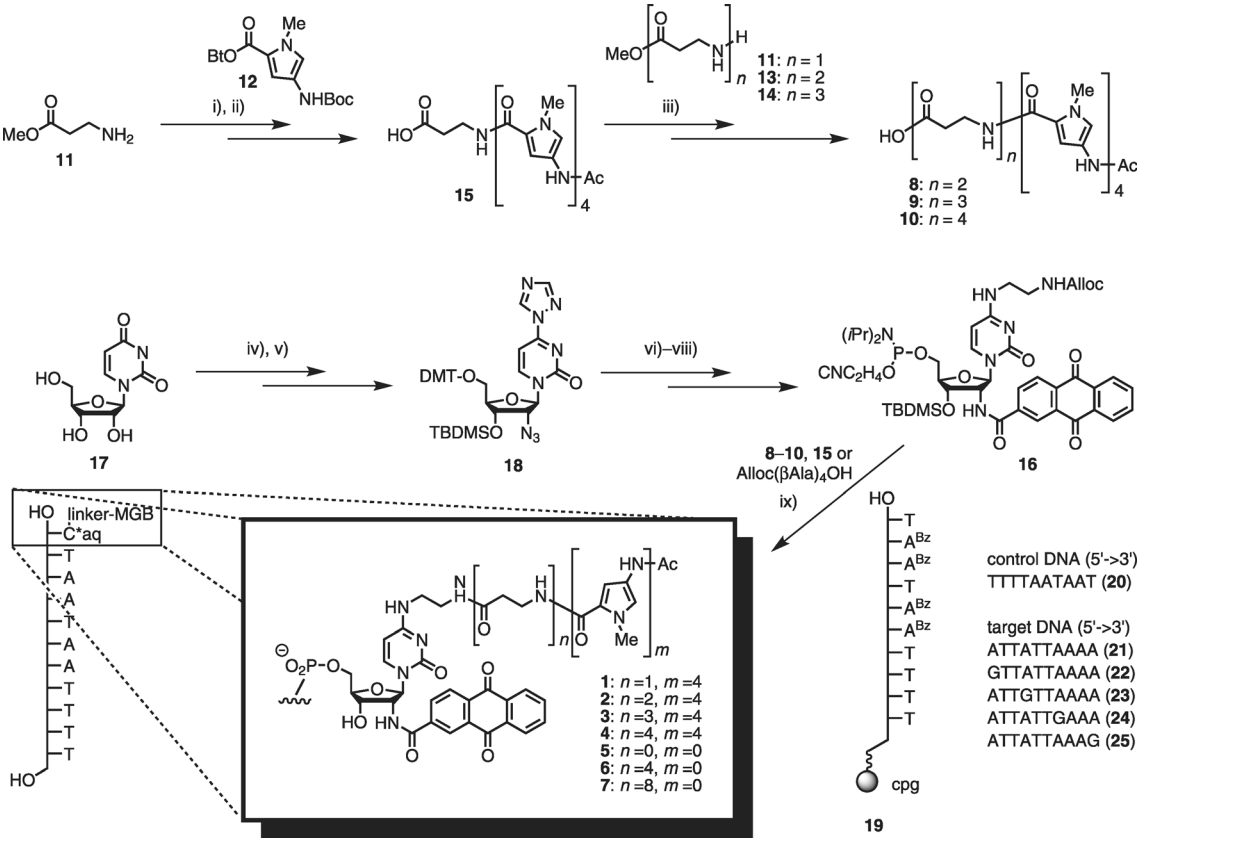
[a] Dipl.-Chem. R. Haug, Dipl.-Chem. M. Kramer, Prof. C. Richert
Institut für Organische Chemie
Universität Stuttgart
70569 Stuttgart (Germany)
Fax: (+49) 711-685-64321
E-mail: lehrstuhl-2@oc.uni-stuttgart.de

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/chem.201302972.

thesized in solution, starting from **11** and active ester **12** or methyl esters **11**, **13** or **14** as building blocks, and a protocol typical for Boc-based approaches.[17,18] Intermediate **15** was also used directly for coupling to the probe. Additionally, terminally *N*-acetylated tetramer Ac(β-Ala)$_4$ as well as octamer Ac(β-Ala)$_8$ were prepared as controls (see Supporting Information for details of those and all other syntheses). Next, we prepared the N4-aminoethylcytidine building block. Phosphoramidite **16** was elaborated from uridine **17** in 10 steps and 22% overall yield via azidotriazolide **18** as key intermediate. Automated chain assembly using commercial 5′-phosphoramidites on a DNA synthesizer produced support **19**, which was smoothly extended with **16**, followed by deprotection of the aliphatic amine, and coupling of carboxylic acids **8**–**10**, **15**, or Alloc(β-Ala)$_4$OH with subsequent removal of the Alloc moiety and acetylation. Two consecutive coupling rounds with Alloc(β-Ala)$_4$OH and acetylation produced the probe with the Ac(β-Ala)$_8$ chain (**7**). A two-stage deprotection, with TBAF followed by NH_4OH/$MeNH_2$, was used to cleave all protecting groups and release the oligonucleotides from the support.

UV-Melting experiments were performed to study the molecular recognition of target strand **20** (Table 1). With aminoethyl-terminated C*Aq cap alone (**5**), the melting point (T_m) increased by 11.5 °C, a value close to the 13 °C measured earlier for the UAq cap on a decamer.[14] Upon addition of the tetrapyrrole unit, the ΔT_m values shot up to 38.3–44.8 °C, with the T_m increase levelling off at four β-Ala residues (**4**:**21**). At this length, the minor groove binder contributed 34.3 °C to the melting point increase, beyond the 11.5 °C provided by the cap alone. The ΔT_m of 44.8 °C observed for **4**:**21** is the largest melting point increase for any terminal substituent of an oligonucleotide that we are aware of. Further, the melting transition for the duplex of three-pronged probe **4** with its target is highly cooperative (Figure 2), a desirable feature for high fidelity hybridization.[19]

We then studied the components of **4** more closely. First, probe strands containing only β-alanine residues were tested. Four β-Ala residues (**6**:**21**) increased the T_m by 4 °C over that of the cap-only version (**5**:**21**), and eight such residues by 8.3 °C (**7**:**21**). This increase of ~1 °C per β-alanine



Scheme 1. Synthesis of three-pronged probes and control compounds. i) Synthesis cycle (repeated four times): 1) **12**, DIPEA, 2) MeOH, HCl; ii) 1) Ac_2O, pyridine, 2) K_2CO_3 aq.; iii) 1) HBTU, DIPEA, **11**, **13** or **14**, 2) K_2CO_3 aq.; iv) 1) diphenyl carbonate, $NaHCO_3$, 2) DMTCl, DMAP, pyridine, 3) NaN_3, [15]crown-5, 4) TBDMSCl, imidazole; v) $POCl_3$, triazole, DIPEA; vi) $AllocNHC_2H_4NH_2$, DIPEA; vii) 1) PPh_3, H_2O, 2) anthraquinone carboxylic acid, HBTU, DIPEA; viii) 1) Cl_3CCO_2H, MeOH, 2) $NCC_2H_4OP(N(iPr)_2)_2$, DIPAT; ix) 1) **19**, tetrazole, 2) I_2, py, H_2O, 3) $[Pd(PPh_3)_4]$, $Et_2NH_2^+HCO_3^-$, 4) **8**, **9**, **10**, **15**, or Alloc-(βAla)$_4$-OH, HBTU, DIPEA, (deprotection and coupling with Alloc-(βAla)$_4$-OH twice for compound **7**); 5) for compounds **5**–**7**: $[Pd(PPh_3)_4]$, $Et_2NH_2^+HCO_3^-$; and then Ac_2O, 2,6-lutidine, *N*-methylimidazole, 6) 1) TBAF, 2) NH_4OH, $MeNH_2$. (Compound **5** was formed simultaneously under these conditions.) Abbreviations: Alloc = allyloxycarbonyl, Boc = *tert*-butyloxycarbonyl, Bt = benzotriazolyl, cpg = controlled pore glass, DMT = 4,4′-dimethoxytrityl, TBDMS = *tert*-butyldimethylsilyl.

Table 1. UV-Melting points of duplexes of ATTATTAAAA (**21**) and the probe strands given.

Probe	Melting point [°C][a]	ΔT_m [°C][b]	Hyperchromicity [%][c]
20	19.9±1.1	–	37
1	58.2±0.8	38.3	33
2	62.3±0.4	42.4	31
3	64.6±0.2	44.7	36
4	64.7±0.4	44.8	39
5	31.4±0.8	11.5	24
6	35.7±0.6	15.5	30
7	39.2±0.7	19.8	27

[a] Average of 4–6 melting points at strand concentrations of 1.5 μM each, 10 mM phosphate buffer, pH 7, and 1 M NaCl, with a heating or cooling rate of 1 °C min^{-1}. [b] Difference to melting point of unmodified duplex. [c] Hyperchromicity at 260 nm.

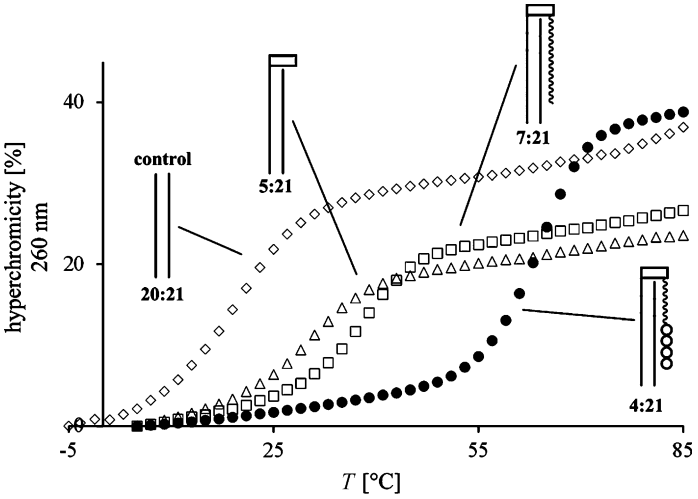


Figure 2. UV-Melting curves of duplexes of strands **4**, **5**, **7**, or **20** with target strand **21**. The cartoons highlight the components of each probe in simplified form.

residue is less than the 2–4 °C increase per β-alanine observed for the duplexes of **1**/**2**/**3**/**4** with **21**, confirming how well the combination of β-alanines and pyrroles cooperate in stabilizing the duplex with the target. Next, we tested the DNA-binding properties of a β-alanine oligomer in free form, without a covalent link to a probe. An NMR titration experiment with free $(\beta\text{-Ala})_6$ and the mixed-sequence DNA hairpin 5′-CCAG-HEG-CTG-3′ (where HEG denotes a hexaethyleneglycol linker) showed that the oligoamide does indeed bind to DNA (Figure 3). Binding affects protons located at terminal residues as well as H2 of the central A residue, located in the minor groove. The signal broadening suggests that there is no static complex, though, but rather a dynamic binding situation when the oligoamide is in its free form.

Finally, we performed exploratory experiments with three-pronged probe **4** to study the level of mismatch discrimination and to gain insights into the range of nucleotides covered by the oligoamide portion. For this, every fourth base of the target strand was "mutated" from an A to a G (mismatched target strands **22**–**25**).

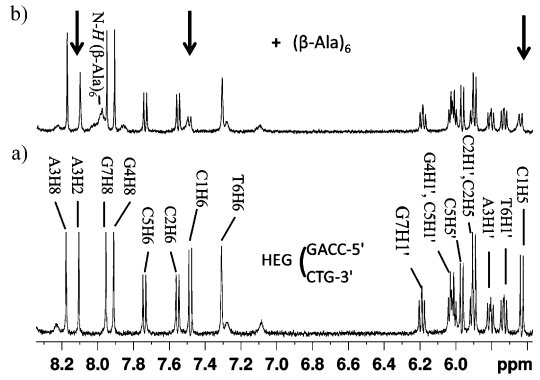


Figure 3. Binding of a β-alanine oligomer alone to DNA. ^{1}H NMR spectrum of DNA hairpin 5′-CCAG-HEG-CTG-3′ in the absence (a) or in the presence of one equivalent of $(\beta\text{Ala})_6$ (b), in H_2O/D_2O (9:1) at 2 mM strand concentration 500 MHz and 20 °C. Arrows show signals most affected by the binding of $(\beta\text{Ala})_6$.

The G:T base combination resulting from this change forms rather stable mismatched base pairs (wobble base pairs)[20] that are difficult to suppress. The results from this A/G scan of the duplex (Figure 4 and Table S1, Supporting Information) showed strong mismatch discrimination at positions 1, 4, and 7 of the target strand of **4**, but little at the terminal residue (position 10) that is outside the range of the polyamides, even in fully extended conformation. When a wobble pair was introduced at the center of the sequence (G4), the loss in affinity of three-pronged probe **4** for the target was so strong that there was a noticeable break-down of cooperativity (Figure 4a). This effect was not observed for the control duplex (Figure 4b).

In conclusion, we note that the combination of Uaq cap, β-alanine oligomer and oligopyrrolamide leads to very high affinity for an all-A/T-target strand. The combination propels the melting point into a region usually reserved for sequences with a significant G/C content. This was unexpected. Individual residues of β-alanine, when incorporated in polyamide minor groove binders, had been reported to promote binding in an extended conformation,[16b] but were not known to be DNA binders in their own right. Neither the cap alone, nor the minor groove binder with a minimal linker, a β-alanine oligomer alone, or a combination of the latter without the linker to the cap lead to the same level of cooperativity and selectivity in binding. As with every probe, a hybridization experiment requires proper choice of incubation temperature. Without this, either false positives (temperature too low) or false negatives (temperature too high) will result. Three-pronged probes are more selective in their sequence recognition and they give a temperature optimum for weakly pairing sequences that is more similar to that of typical mixed sequences, so that massively parallel high fidelity detection of different DNA sequences may be more readily achieved, using the concept of isostable duplexes and universally stringent conditions.[21]

We believe that the TPP probes are a step towards more thorough interrogation of DNA target strands with synthetic

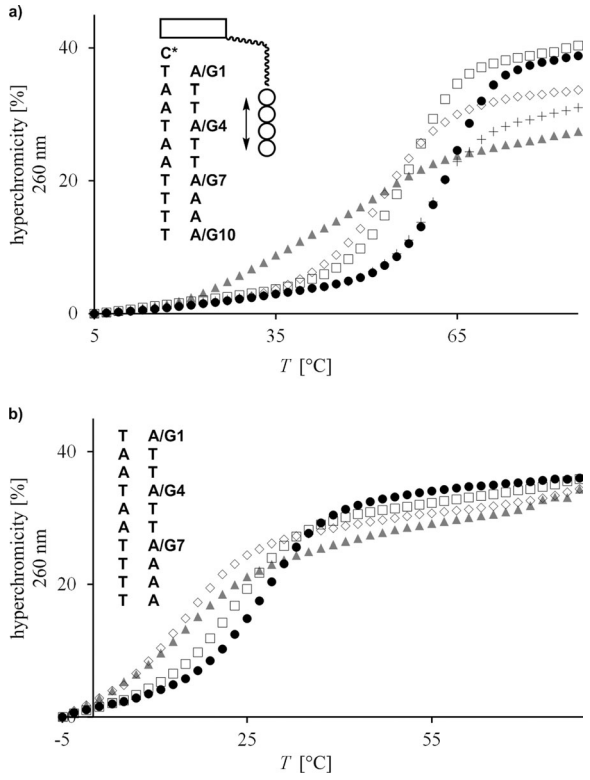


Figure 4. a) "G-Scan" analysis of the binding site of the oligoamide portion of **4**. a) UV-Melting curves of the duplex of **4** with **21**, or **22**–**25** with one mismatched G residue. Conditions: 1.5 μM strands, 10 mM phosphate buffer, pH 7, and 1 M NaCl. b) UV-melting curves of duplexes of control strand **20** with target strands **21**–**24** under the same conditions, except that the strand concentration was 18 μM in order to shift the curves into the observable temperature range. Symbols used: ● perfectly matched sequence, □ G1, ▲ G4, ◇ G7, and + G10.

molecules. Enzymes that catalyze pivotal genetic processes, such as polymerases, frequently bind DNA substrates from several sides, including interactions in the minor groove and from above the "terminus" where the incoming dNTP forms a terminal base pair.[22] Further, proteins at a pivotal point of genetic control, such as p53, show contacts in the major and the minor groove in their complex with DNA.[23] In mimicking proteins, the design of the TPPs builds on the successful optimization of molecular caps for the termini and the exceptionally well developed chemistry of minor groove binders. The extension of our current work to TPPs with MGBs that are specific for other target sequences[15] is a natural next step.

Experimental Section

Melting curves of oligonucleotide duplexes were acquired at 260 nm with a temperature gradient of 1 °C min^{-1}. Synthetic protocols, NMR and mass spectra, additional melting points and data on NMR experiments with free $(\beta\text{-Ala})_6$ can be found in the Supporting Information.

Acknowledgements

The authors are grateful to Drs. B. Claasen and E. Kervio for the acquisition and help with the interpretation of NMR results, H. Griesser for a review of the manuscript and Dr. P. Grünefeld for sharing experimental results and discussions. This work was supported by DFG (grants No. RI 1063/9-1 and RI 1063/13-1).

Keywords: DNA · duplexes · hybridization · minor groove binder · oligonucleotides

[1] J. D. Watson, F. H. Crick, *Nature* **1953**, *171*, 964–967.
[2] a) J. Gralla, D. M. Crothers, *J. Mol. Biol.* **1973**, *78*, 301–306; b) J. SantaLucia, Jr., D. Hicks, *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 415–440; c) W. W. Hadiwikarta, J. C. Walter, J. Hooyberghs, E. Carlon, *Nucleic Acids Res.* **2012**, *40*, e138.
[3] a) K. Siegmund, U. E. Steiner, C. Richert, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 2153–2162; b) U. Plutowski, C. Richert, *Angew. Chem.* **2005**, *117*, 627–631; *Angew. Chem. Int. Ed.* **2005**, *44*, 621–625.
[4] A. Kornberg, T. A. Baker, in *DNA Replication*, 2nd ed., University Science Books, Mill Valley, **2005**, pp. 118–129.
[5] a) K. Hitomi, S. Iwai, J. A. Tainer, *DNA Repair* **2007**, *6*, 410–428; b) J. J. P. Perry, E. Cotner-Gohara, T. Ellenberger, J. A. Tainer, *Curr. Opin. Struct. Biol.* **2010**, *20*, 283–294.
[6] a) B. Vester, J. Wengel, *Biochemistry* **2004**, *43*, 13233–13241.
[7] a) K.-Y. Lin, M. D. Matteucci, *J. Am. Chem. Soc.* **1998**, *120*, 8531–8532; b) F. Seela, G. Becher, *Nucleic Acids Res.* **2001**, *29*, 2069–2078; c) H. Liu, J. Gao, S. R. Lynch, Y. D. Saito, L. Maynard, E. T. Kool, *Science* **2003**, *302*, 868–871.
[8] See e.g.: a) R. W. Wagner, M. D. Matteucci, J. G. Lewis, A. J. Gutierrez, C. Moulds, B. C. Froehler, *Science* **1993**, *260*, 1510–1513; b) J. I. Gyi, D. Gao, G. L. Conn, J. O. Trent, T. Brown, A. N. Lane, *Nucleic Acids Res.* **2003**, *31*, 2683–2693; c) T. Kottysch, C. Ahlborn, F. Brotzel, C. Richert, *Chem. Eur. J.* **2004**, *10*, 4017–4028.
[9] E. T. Kool, *Acc. Chem. Res.* **1998**, *31*, 502–510.
[10] a) I. V. Kutyavin, I. A. Afonina, A. Mills, V. V. Gorn, E. A. Lukhtanov, E. S. Belousov, M. J. Singer, D. K. Walburger, S. G. Lokhov, A. A. Gall, R. Dempcy, M. W. Reed, R. B. Meyer, J. Hedgpeth, *Nucleic Acids Res.* **2000**, *28*, 655–661; b) J. Robles, S. B. Rajur, L. W. McLaughlin, *J. Am. Chem. Soc.* **1996**, *118*, 5820–5821; c) A. N. Sinyakov, S. G. Lokhov, I. V. Kutyavin, H. B. Gamper, R. B. Meyer, *J. Am. Chem. Soc.* **1995**, *117*, 4995–4996; d) E. Kawashima, Y. Nakanishi, Y. Terui, H. Tomitori, K. Kashiwagi, Y. Ohba, K. Kamaike, *Nucleosides Nucleotides Nucleic Acids* **2013**, *32*, 196–205; e) S. B. Rajur, J. Robles, K. Wiederholt, R. G. Kuimelis, L. W. McLaughlin, *J. Org. Chem.* **1997**, *62*, 523–529; f) J. W. Szewczyk, E. E. Baird, P. D. Dervan, *Angew. Chem.* **1996**, *108*, 1596–1598; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1487–1489.
[11] a) C. F. Bleczinski, C. Richert, *J. Am. Chem. Soc.* **1999**, *121*, 10889–10894; b) J. Tuma, C. Richert, *Biochemistry* **2003**, *42*, 8957–8965; c) S. Narayanan, J. Gall, C. Richert, *Nucleic Acids Res.* **2004**, *32*, 2901–2911; d) S. Egetenmeyer, C. Richert, *Chem. Eur. J.* **2011**, *17*, 11813–11827.
[12] a) R. Lucas, E. Vengut-Climent, I. Gómez-Pinto, A. Aviñó, R. Eritja, C. González, J. C. Morales, *Chem. Commun.* **2012**, *48*, 2991–2993; b) J. C. Morales, J. J. Reina, I. Díaz, A. Aviñó, P. M. Nieto, R. Eritja, *Chem. Eur. J.* **2008**, *14*, 7828–7835; c) R. Lucas, I. Gómez-Pinto, A. Aviñó, J. J. Reina, R. Eritja, C. González, J. C. Morales, *J. Am. Chem. Soc.* **2011**, *133*, 1909–1916.
[13] Z. Dogan, R. Paulini, J. A. Rojas Stütz, S. Narayanan, C. Richert, *J. Am. Chem. Soc.* **2004**, *126*, 4762–4763.
[14] A. Patra, C. Richert, *J. Am. Chem. Soc.* **2009**, *131*, 12671–12681.
[15] P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 2215–2235.
[16] a) J. M. Turner, S. E. Swalley, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1998**, *120*, 6219–6226; b) A. R. Urbach, J. J. Love, S. A.

Ross, P. B. Dervan, *J. Mol. Biol.* **2002**, *320*, 55–71; c) V. C. Rucker, C. Melander, P. B. Dervan, *Helv. Chim. Acta* **2003**, *86*, 1839–1851.
[17] D. M. Chenoweth, D. A. Harki, P. B. Dervan, *J. Am. Chem. Soc.* **2009**, *131*, 7175–7181.
[18] T. Hintermann, C. Mathes, D. Seebach, *Eur. J. Org. Chem.* **1998**, 2379–2387.
[19] K. Siegmund, C. Ahlborn, C. Richert, *Nucleosides Nucleotides Nucleic Acids* **2008**, *27*, 376–388.
[20] F. H. C. Crick, *J. Mol. Biol.* **1966**, *19*, 548–555.
[21] C. Ahlborn, K. Siegmund, C. Richert, *J. Am. Chem. Soc.* **2007**, *129*, 15218–15232.
[22] S. Doublie, S. Tabor, A. M. Long, C. C. Richardson, T. Ellenberger, *Nature* **1998**, *391*, 251–258.
[23] Y. Cho, S. Gorina, P. D. Jeffrey, N. P. Pavletich, *Science* **1994**, *265*, 346–355.

Received: July 28, 2013
Published online: October 29, 2013

