DNA Bisintercalators

Design of a Sequence-Specific DNA Bisintercalator**

*Eric J. Fechter, Bogdan Olenyuk, and Peter B. Dervan**

The development of sequence-specific DNA bisintercalators has been an ongoing challenge in the field of bioorganic chemistry and molecular recognition. Since the first proposed model of intercalation by Lerman,[1] the disruption of transcription or replication by helix unwinding and extension has become an attractive strategy for blocking essential gene functions in the field of cancer therapy and antibiotics, but these methods have proved toxic, presumably owing to lack of specificity (Figure 1 a).[2] By the mid-1970s Waring and coworkers had reported the DNA-binding characteristics of the first known bisintercalating natural product, echinomycin[3] (Figure 1 b). This pseudosymmetrical bifuntional molecule is part of a larger class, known as quinoxaline antibiotics, and contains a cyclic octapeptide minor-groove-binding region with two linked chromophores capable of simultaneous intercalation.[4] The minor-groove-binding bicyclic depsipeptide backbone undergoes hydrogen bonding to DNA bases and was thought to impose the modest sequence specificity for poly(dG–dC) sites.[3] Similarly, the related bisintercalator triostin[5] also inhibits DNA replication and RNA synthesis,[6] yet has a slight specificity toward poly(dA–dT) sites.[7] As with other members of the quinoxaline family of antibiotics, the incorporation of modified amino acids presumably provides the relative specificities.

Synthetic bisintercalators were later constructed by linking two heterocycles, such as acridines,[8] methidium,[9] and anthracyclines,[10] with chains of varying lengths to maximize the bracketing of two (or more) base pairs between the intercalator sites (Figure 1 c). In general, these molecules have enhanced affinity for DNA but lack significant sequence specificity. To our knowledge, in the past 30 years, attempts to design bisintercalators with *programmable* sequence specificity have been largely unsuccessful.

We recently reported the synthesis of a hybrid molecule, a hairpin polyamide–acridine conjugate, which enforces two very different modes of DNA binding: groove binding and intercalation.[11] Remarkably, the pyrrole–imidazole polyamide, which prefers to bind B-form DNA, maintains its sequence specificity despite the presence of an adjacent intercalated acridine moiety which extends and unwinds the helix (unwinding angle $\phi = 14$–15°) proximal to the groove-binding ligand.[11]

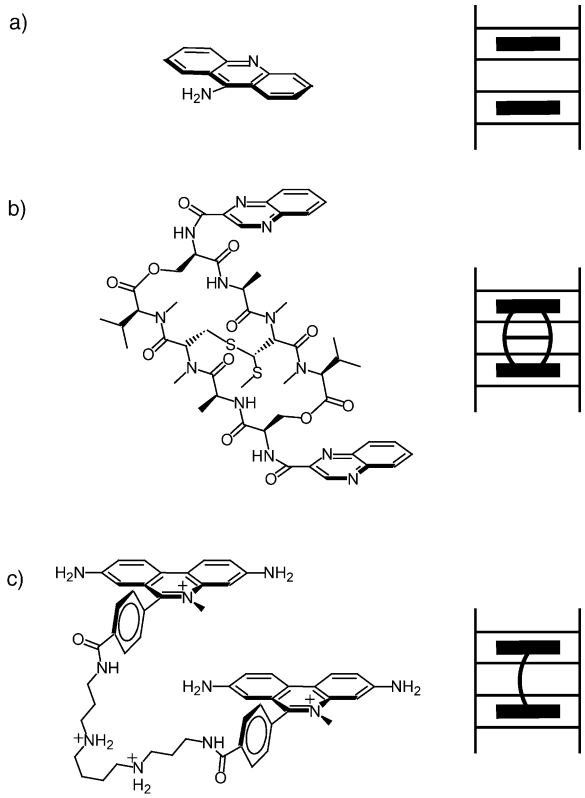


Figure 1. Intercalative models: molecular structures (left) and binding models with the shaded bars depicting the acridine intercalators (right). a) 9-Aminoacridine,[1] b) echinomycin,[3] c) bis(methidium)spermine.[9]

Based on this sequence-specific intercalator lead[12] we explored the synthesis and binding properties of *sequence-specific bisintercalators* (Figure 2). Our design is a symmetric molecule which contains a minor-groove-binding polyamide based on the H-pin motif[13] with an acridine moiety at each C terminus. According to the pairing rules[14] the H-pin core should target the sequence 5′-TGACA-3′ and, based on earlier precedent, the two acridine moieties should unwind DNA by $\geq 30^{\circ}$.

Cross-linked resin **1** was synthesized by loading β-Ala-PAM (PAM = phenylacetamido methyl) resin with activated pyrrole amino acid,[15] subsequent *tert*-butoxycarbonyl (Boc) deprotection and addition of the ring-linked dimeric building block to couple the C termini of the growing polyamide chain on the resin,[13] and final capping by using activated imidazole carboxylate (Scheme 1). Resin-bound H-pin **1** was then subjected to aminolytic cleavage with 2,2′-(ethylenedioxy)-bis(ethylamine) to form H-pin diamine **2**. Following purification, **2** was coupled with 9-chloroacridine under reported conditions[11] to produce the mono- and bisacridine–H-pin conjugates **3** ((ImPyPy-β-Do-Acr)(ImPyPy-β-Do)$(CH_2)_6$) and **4** ((ImPyPy-β-Do-Acr)$_2(CH_2)_6$), respectively.

[*] E. J. Fechter, B. Olenyuk, Prof. P. B. Dervan
Division of Chemistry and Chemical Engineering
164-30, California Institute of Technology
Pasadena, CA 91125 (USA)
Fax: (+1) 626-564-9297
E-mail: dervan@caltech.edu

[**] We thank the National Institutes of Health (GM-27681) for research support, a Research Service Award to E.J.F., and a postdoctoral fellowship to B.O. (F32 GM-19788). We also thank the Ralph M. Parsons Foundation for a predoctoral fellowship to E.J.F.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

 DOI: 10.1002/ange.200454231

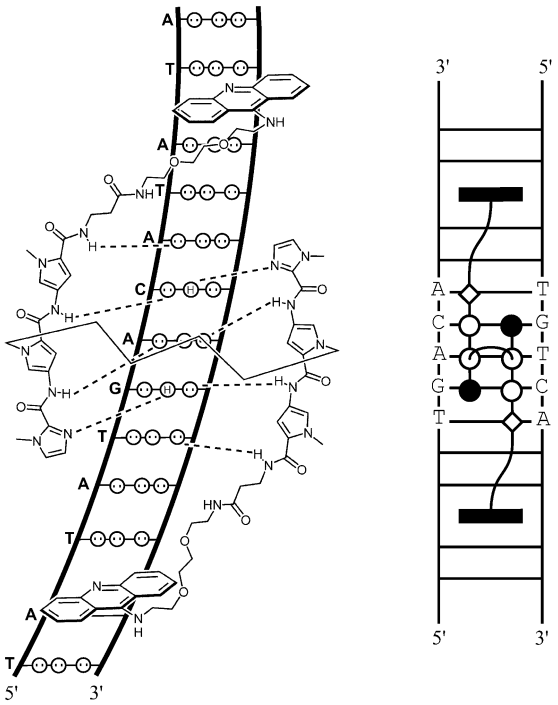
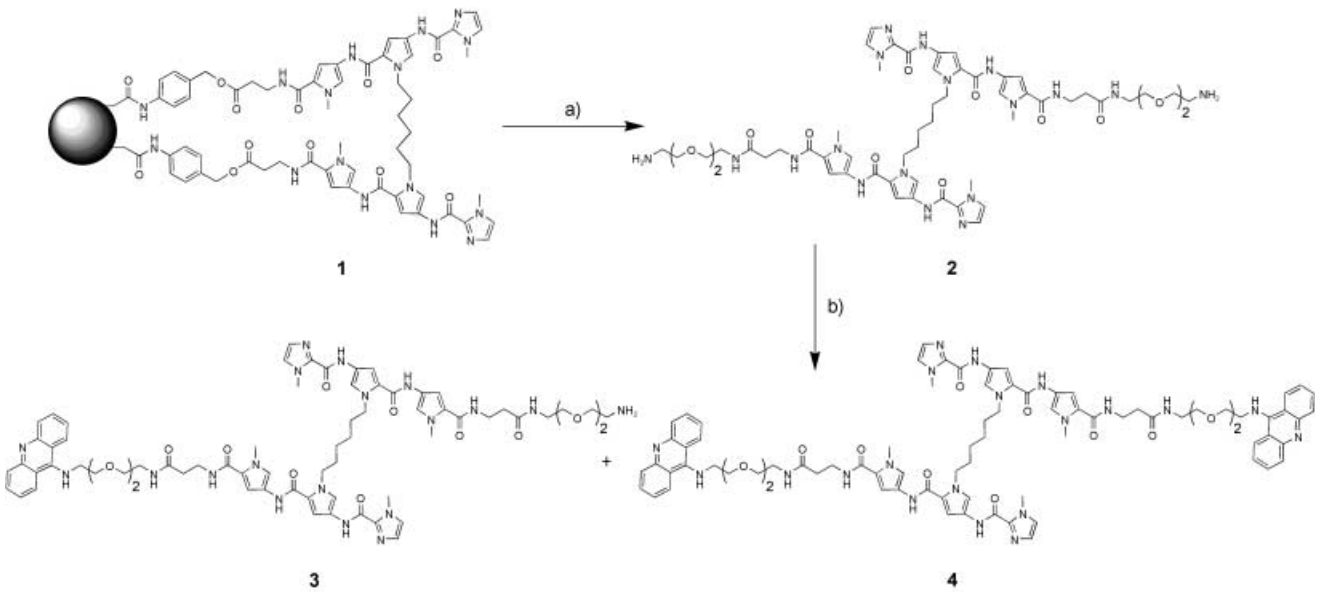


Figure 2. DNA binding model for the symmetric bisintercalating H-pin polyamide–acridine conjugate (ImPyPy-β-Do-Acr)$_2$(CH$_2$)$_6$ bound to the minor groove of 5′-TGACA-3′. Im = imidazole, Py = bridged pyrrole, Py = pyrrole, $\beta = \beta$-alanine, Do = 2,2′-(ethylenedioxy)bis(ethylamine), Acr = acridine. Left: Circles with dots represent the lone pairs of electrons of N3 of purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogen atoms of guanine. Putative hydrogen bonds are illustrated by dotted lines. Right: Solid circles represent imidazoles, open circles represent pyrroles, and diamonds denote β-alanine. According to the pairing rules, Im/Py codes for G·C, Py/Py for A·T or T·A, Py/Im for C·G, and β-alanine for A·T.

The DNA-binding properties of **2**–**4** were investigated by quantitative DNase I footprinting assays[16] (Figure 3). The 5′-^{32}P-labeled PCR-amplified fragment of pEF12 (Figure 3a) contains a match site (A) and a single-bp mismatch site (B) to study both affinity and specificity. The equilibrium binding constants of compounds **2**–**4** for match site A (5′-TGACA-3′) are compiled in Table 1 and their corresponding binding isotherms are shown in Figure 4a. Conjugation of one acridine (**3**) results in a nearly 40-fold increase in binding affinity ($K_a = 1.4 \times 10^9\,\text{M}^{-1}$) over the parent H-pin **2** ($K_a = 3.7 \times 10^7\,\text{M}^{-1}$). The conjugation of a second acridine intercalator (**4**) increases the binding affinity by an additional 10-fold ($K_a = 1.5 \times 10^{10}\,\text{M}^{-1}$), thereby resulting in the bisacridine conjugate **4** having a more than 400-times higher affinity for a DNA match site than its unconjugated counterpart **2**. It is noteworthy that polyamides conjugated to nonintercalating moieties (such as fluorescent dyes and peptides) display *decreased* binding affinities relative to their parent, unconjugated compounds.[17] The steep slopes of the isotherms for **3** and **4** in Figure 4 suggest a more complex DNA binding mode than the expected 1:1 association of parent H-pin **2**. Compounds **2** and **3** show no binding on mismatch site B (5′-TG**G**CA-3′) at concentrations as high as 1 μM. It appears that bisintercalator **4** may have a partial occupation of the mismatch site (B) at the highest concentrations (Figure 3b). Nonetheless, the affinity of H-pin **4** for mismatch site B could not be quantified at concentrations as high as 1 μM, a result indicating a high level of specificity in this series of conjugates.

The DNA-unwinding properties of compounds **3** and **4** were determined from a helical assay developed by Crothers, Zeeman, and colleagues that provides an unwinding angle (ϕ) from sequence-specific interactions.[18] A series of relaxation reactions were carried out by using topoisomerase I (Topo I) on closed-circular pUC19 DNA preequilibrated with varying concentrations of polyamides. The plasmid was then sepa-



Scheme 1. Synthetic scheme for H-pin polyamide–acridine conjugates: a) 2,2′-(ethylenedioxy)bis(ethylamine), 55 °C, 18 h; b) 9-chloroacridine (1.5 equiv), DIPEA, 100 °C (1.5 h). DIPEA = diisopropylethylamine.

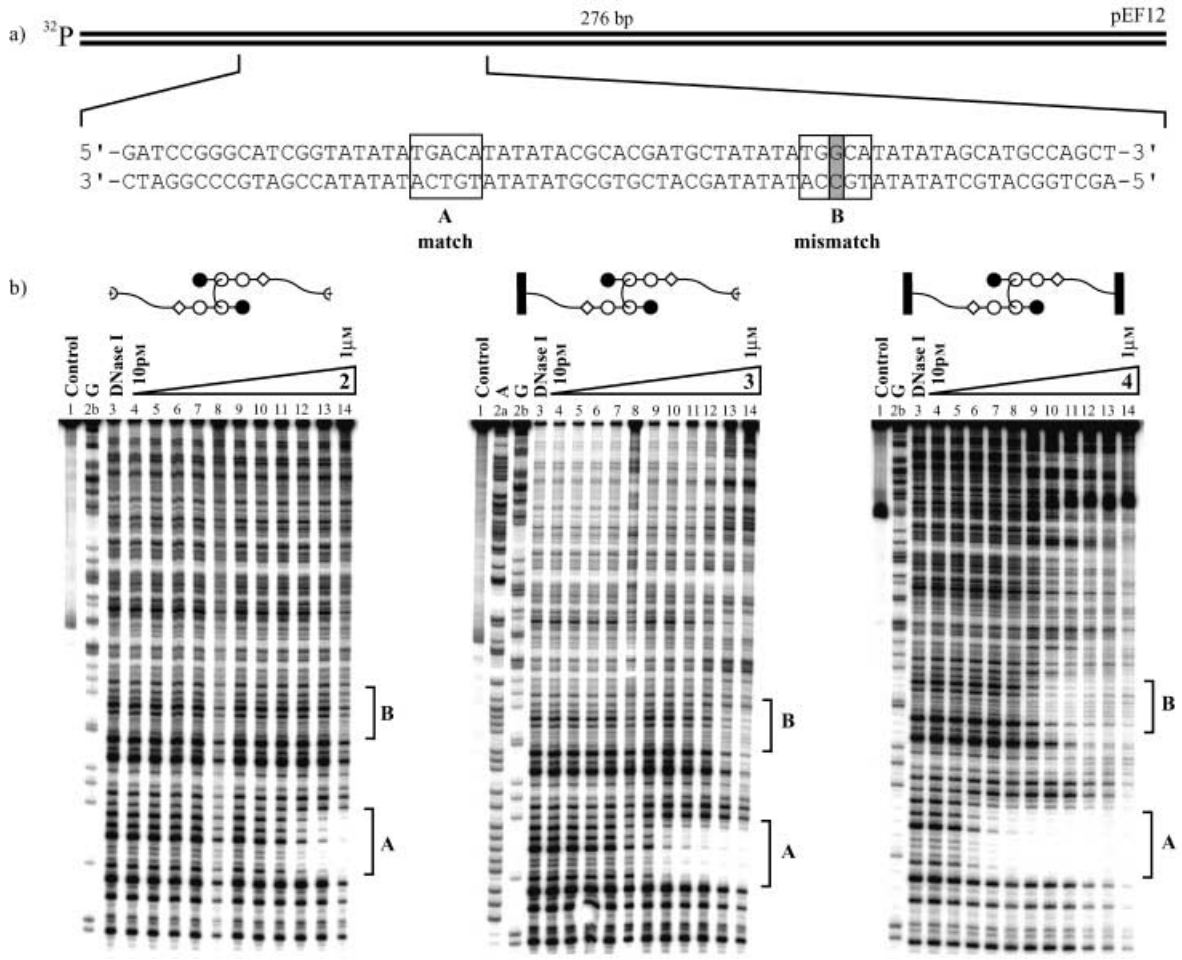


Figure 3. a) Sequence of the synthesized insert from the pEF12 plasmid containing the 5-bp target match site (A) and single-bp mismatch site (B). Target sites are shown in boxes with the mismatch site shaded. b) Quantitative DNase I footprint titration experiments with H-pins **2** (left), **3** (middle), and **4** (right) on the PCR-amplified 5′-^{32}P-labeled fragment from pEF12. Lane 1, intact DNA; lane 2a, A reaction; lane 2b, G reaction; lane 3, DNase I standard; lanes 4–14, DNase I digestion products in the presence of 10, 30, 100, 300 pM; 1, 3, 10, 30, 100, 300 nM; and 1 μM polyamide, respectively.

Table 1: Thermodynamic Data and Unwinding Angles

Polyamide	No. of acridines	K_a[a] [M^{-1}]	ϕ [°]
2	0	3.7×10^{7}	0
3	1	1.4×10^{9}	15
4	2	1.5×10^{10}	34

[a] Association constants, K_a, are the average values obtained from at least three DNase I footprint titration experiments. Assays were performed at 22 °C and pH 7.0 in the presence of tris(hydroxymethyl)-aminomethane·HCl (10 mM), KCl (10 mM), $MgCl_2$ (10 mM), and $CaCl_2$ (5 mM).

rated from the polyamide conjugate by phenol/chloroform extraction and 2D agarose-gel electrophoresis was performed to distinguish the resulting distribution of topoisomers. DNA unwinding would shift the topoisomer distribution toward a more negatively supercoiled population. Indeed, each reaction containing conjugate **4** had a more highly negative distribution of topoisomers than those containing conjugate **3** (see the Supporting Information). Control experiments lacking polyamide resulted in a primarily positive distribution of topoisomers. Mathematical analysis of the topoisomer distributions relative to the controls showed decreasing apparent unwinding angles (ϕ_{ap}) for simultaneously decreasing conjugate and plasmid concentrations (Figure 4b). The actual unwinding angles (ϕ), determined from the ordinate intercepts, are 15° and 34° for the acridine conjugate **3** and bisacridine conjugate **4**, respectively (Figure 4b).

Our results provide strong evidence that the symmetric H-pin–bisacridine conjugate **4** is a sequence-specific bisintercalator, capable of binding discrete sites at subnanomolar concentrations and unwinding DNA by more than 30°. To date, this synthetic molecule exceeds the specificity and binding affinity of any known natural or non-natural bisintercalator. Its ease of synthesis, sequence specificity, and potency of DNA distortion at discrete sites make it an attractive candidate for future biological studies, such as transcription inhibition at specific genes. The design features of combining the programmable H-pin motif with intercala-

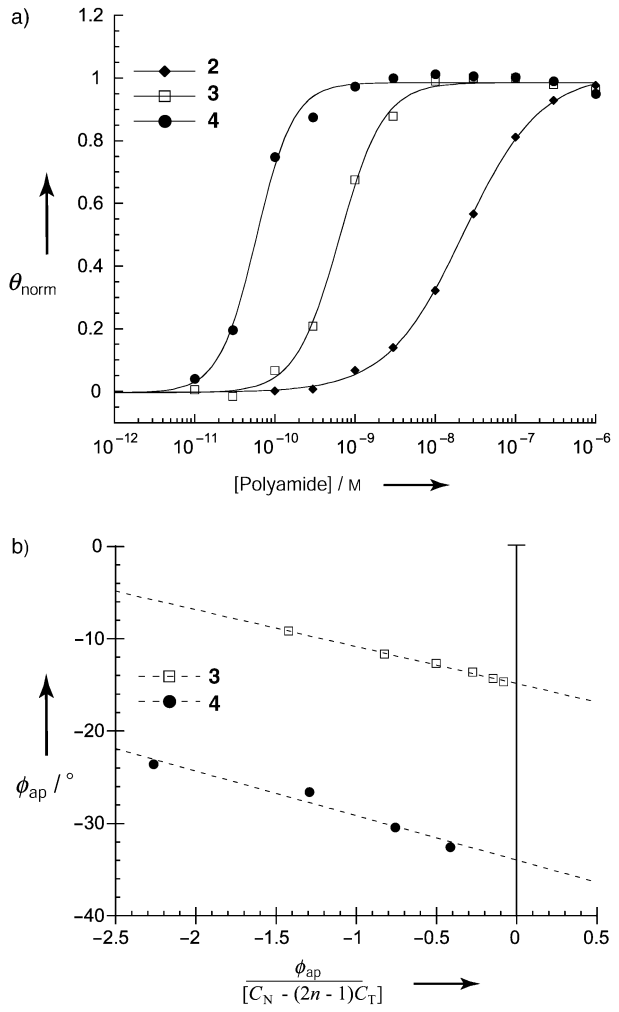


Figure 4. a) Binding isotherms at match sites for H-pins **2**, **3**, and **4**. θ_{norm} points were obtained by using storage phosphor autoradiography and processed by standard methods.[15] Each data point shows the average value obtained from three footprinting experiments. The solid curves are best-fit Langmuir binding titration isotherms obtained from a nonlinear least squares algorithm. b) Binding isotherms for polyamide–acridine conjugates **3** and **4** on pUC19. Each data point was calculated from one set of topoisomer distributions from reactions containing polyamide compared to a control distribution lacking polyamide. Interception of the ordinate yields the unwinding angle (ϕ) per polyamide–acridine conjugate.

tion should allow a large repertoire of discrete DNA sequences to be targeted.

Received: March 10, 2004 [Z54231]

Keywords: conjugation · DNA recognition · intercalation · polyamides

[1] L. S. Lerman, *J. Mol. Biol.* **1961**, *3*, 18–30.
[2] a) W. A. Denny, *Anti-Cancer Drug Des.* **1989**, *4*, 241–263; b) B. C. Baguley, *Anti-Cancer Drug Des.* **1991**, *6*, 1–35.
[3] a) M. J. Waring, L. P. G. Wakelin, *Nature* **1974**, *252*, 653–657; b) L. P. G. Wakelin, M. J. Waring, *Biochem. J.* **1976**, *157*, 721–740; c) R. H. Shafer, M. J. Waring, *Biopolymers* **1980**, *19*, 431–443; d) M. M. Van Dyke, P. B. Dervan, *Science* **1984**, *225*, 1122–1127.
[4] A. Dell, D. H. Williams, H. R. Morris, G. A. Smith, J. Feeney, G. C. K. Roberts, *J. Am. Chem. Soc.* **1975**, *97*, 2497–2502.
[5] a) A. H. J. Wang, G. Ughetto, G. J. Quigley, T. Hakoshima, G. A. van der Marel, J. H. van Boom, A. Rich, *Science* **1984**, *225*, 1115–1121; b) H. Chen, D. J. Patel, *J. Mol. Biol.* **1995**, *246*, 164–179.
[6] D. C. Ward, E. Reich, I. H. Goldberg, *Science* **1965**, *149*, 1259–1263.
[7] J. S. Lee, M. J. Waring, *Biochem. J.* **1978**, *173*, 115–128.
[8] a) J.-B. Le Pecq, M. Le Bret, J. Barbet, B. Roques, *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 2915–2919. b) L. P. G. Wakelin, M. Romanos, T. K. Chen, D. Glaubiger, E. S. Canellakis, M. J. Waring, *Biochemistry* **1978**, *17*, 5057–5063.
[9] M. M. Becker, P. B. Dervan, *J. Am. Chem. Soc.* **1979**, *101*, 3664–3666.
[10] F. Leng, W. Priebe, J. B. Chaires, *Biochemistry* **1998**, *37*, 1743–1753.
[11] E. J. Fechter, P. B. Dervan, *J. Am. Chem. Soc.* **2003**, *125*, 8476–8485.
[12] For an example of a major-groove sequence-specific intercalator, see: C. L. Kielkopf, K. E. Erkkila, B. P. Hudson, J. K. Barton, D. C. Rees, *Nat. Struct. Biol.* **2000**, *7*, 117–121.
[13] B. Olenyuk, C. Jitianu, P. B. Dervan, *J. Am. Chem. Soc.* **2003**, *125*, 4741–4751.
[14] S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan, *Nature* **1998**, *391*, 468–471.
[15] E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1996**, *118*, 6141–6146.
[16] J. W. Trauger, P. B. Dervan, *Methods Enzymol.* **2001**, 340, 450–466.
[17] a) T. P. Best, B. S. Edelson, P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12063–12068; b) A. K. Mapp, A. Z. Ansari, M. Ptashne, P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3930–3935.
[18] a) S. M. Zeeman, K. M. Depew, S. J. Danishefsky, D. M. Crothers, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 4327–4332; b) S. M. Zeeman, D. M. Crothers, *Methods Enzymol.* **2001**, 340, 51–68.

