Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters 14 (2004) 3803–3807
Bioorganic & Medicinal Chemistry Letters

DNA sequence recognition in the minor groove by hairpin pyrrole polyamide–Hoechst 33258 analogue conjugate

Putta Mallikarjuna Reddy, Rachel Dexter and Thomas C. Bruice*

Department of Chemistry and Biochemistry, University of California at Santa Barbara, Santa Barbara, CA 93106, USA

Received 26 February 2004; revised 27 April 2004; accepted 27 April 2004
Available online 28 May 2004

Abstract—A hairpin pyrrole polyamide conjugated to a Hoechst 33258 (Ht) analogue, PyPyPy-γ-PyPyPy-γ-Ht, was synthesized on solid-phase by adaptation of an Fmoc technique using a series of PyBOP/HOBt mediated coupling reactions. Sequence selectivity and complex stabilities were characterized by spectrofluorometric titrations and thermal melting studies. The polyamide of the conjugate was observed to bind in a hairpin motif forming 1:1 conjugate:dsDNA complexes. The conjugate is able to recognize nine contiguous A/T bps, discriminating from the sequences containing fewer than nine contiguous A/T bps.

1. Introduction

The incorrect expression or biochemical function of specific genes, whose protein products are essential for the regulation of cell growth, causes cancer.[1] Drugs capable of recognizing specific sequences in DNA have the potential to specifically control a specific gene's expression, and thereby cure the disease rather than simply treating the symptoms. Studies indicate that the synthetic analogues of the naturally-occurring DNA binding ligand Distamycin A have the capability to sequence-specifically recognize dsDNA by binding in the minor groove.[2–6] Among these analogues, polyamides consisting of *N*-methylpyrrole (Py), *N*-methylimidazole (Im), and other structurally related amino acids have received special attention due to their high affinity, degree of sequence specificity, and relatively small size.[3,7] Specific sequences of DNA can be targeted by designing the hairpin polyamide with a monomer sequence that can assemble side-by-side pairing and discriminate between all four Watson–Crick bps.[7] These polyamides have the potential to modulate gene expression by

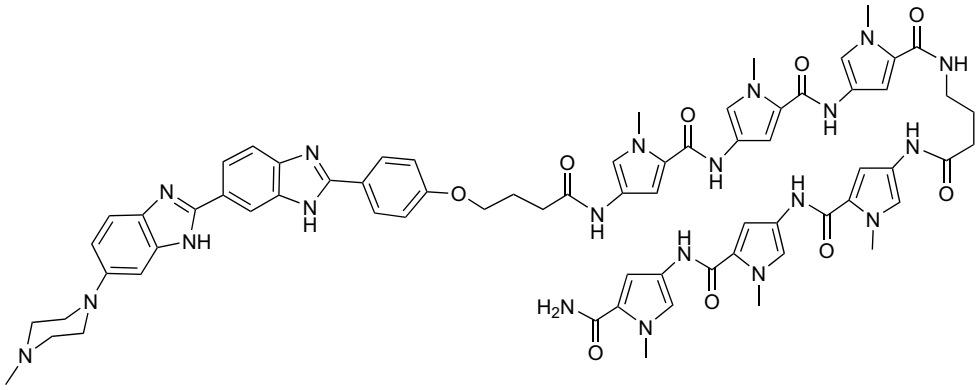


Figure 1. Structure of hairpin pyrrole polyamide–Hoechst 33258 conjugate, Py3-γ-Py3-γ-Ht (**1**).

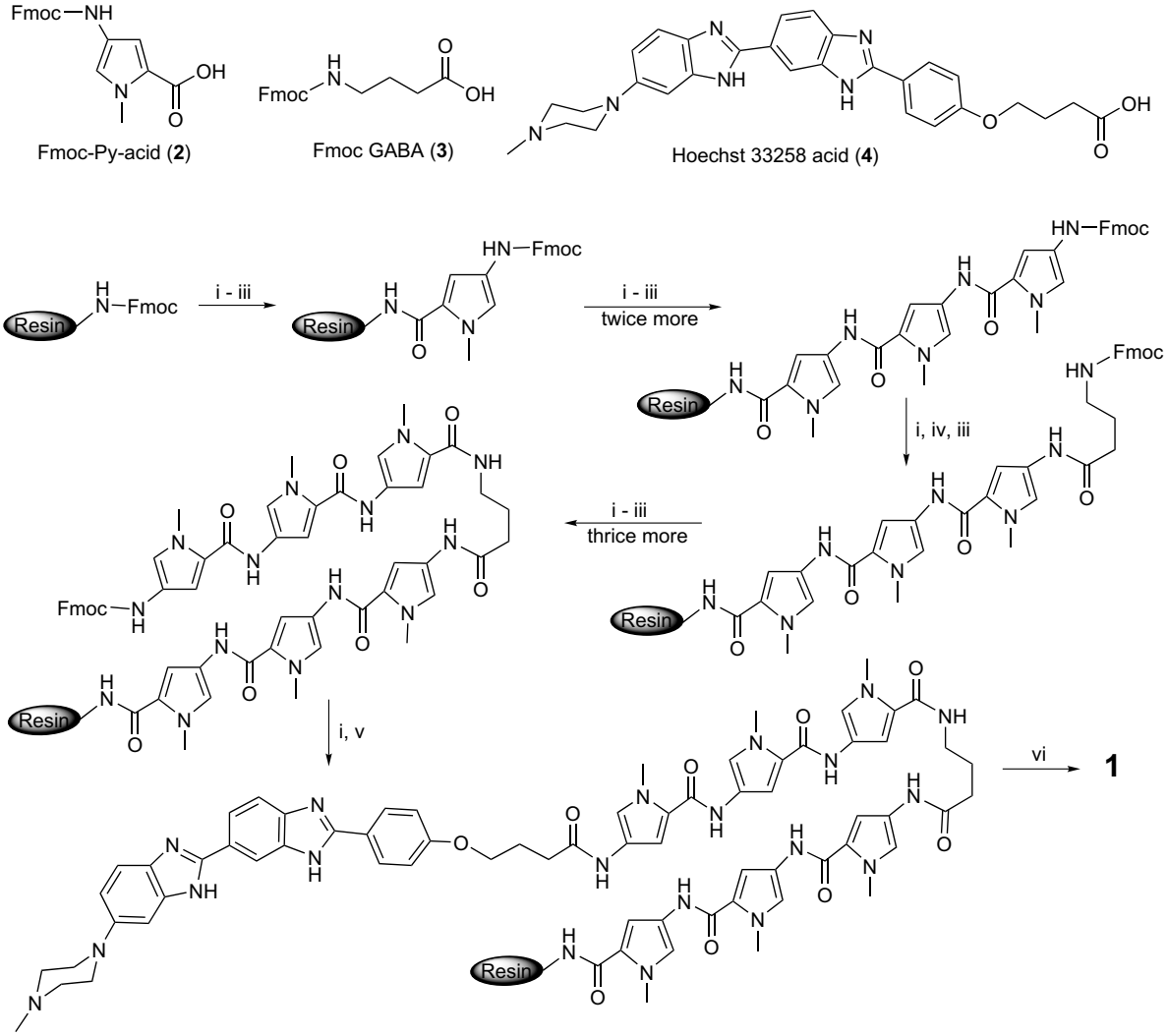
* Corresponding author. Tel.: +1-805-893-2044; fax: +1-805-893-2229; e-mail: tcbruice@bioorganic.ucsb.edu

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.04.089

blocking the activation of transcription factor because their affinity for DNA is on the same order of magnitude as endogenous TFs.[8] However, attempts to inhibit the transcription of endogenous genes in cell lines other than insects or T lymphocytes have met little success, presumably due to poor cellular uptake or nuclear localization of these polyamides.[9]

The bis-benzimidazole fluorophores are known to selectively recognize the A/T-rich minor groove and are capable of crossing the cellular and nuclear membranes.[10] We have taken advantage of these Hoechst ligand characteristics and recently developed novel microgonotropen tripyrrole/imidazole–Hoechst conjugates.[11] These conjugates are capable of passing through the NIH 3T3 cell membrane and inhibiting the DNA–TF complex formation by binding to its nuclear DNA targets.[12] Structure–activity relationships suggest that although the tripyrrole/polyamine functional groups may be largely responsible for the inhibition of TF complexes in cell-free assays, the bis-benzimidazole moiety appears to impart improved cellular uptake and activity. These findings are significant because they represent the first minor groove binders possessing the important characteristics of cell penetration and inhibition of protein expression. However, they lack the clear discrimination of all four Watson–Crick bps, that is, A:T from T:A and G:C from C:G. To achieve the goal of specifically inhibiting the expression of a misregulated protein, minor groove binding agents must bind any predetermined DNA sequence with high affinity, as well as permeate living cells. Although any predetermined DNA sequence can be targeted by designing hairpin polyamides following the established side-by-side amino acid pairing rules, their poor cellular uptake remains a drawback.

Based on the sequence specificity and high affinity of polyamides for the minor groove as well as our previous results, we envisioned that conjugation of a bis-



Scheme 1. Solid-phase synthesis of PyPyPy-γ-PyPyPy-γ-Ht conjugate on commercially available rink amide MBHA resin: Reagents and conditions: (i) *Deprotection*: 20% piperidine/DMF, 10 min; (ii) *Coupling*: Fmoc-Py-acid (**2**), PyBOP, HOBt, DIPEA, DMF, 12 h; (iii) *Capping*: acetic anhydride, TEA, DMF, 10 min; (iv) Fmoc-γ-aminobutyric acid (**3**), PyBOP, HOBt, DIPEA, DMF, 12 h; (v) Hoechst 33258 acid (**4**), PyBOP, HOBt, DIPEA, DMF, 24 h; (vi) *Cleavage*: TFA, TIS, 2 h.

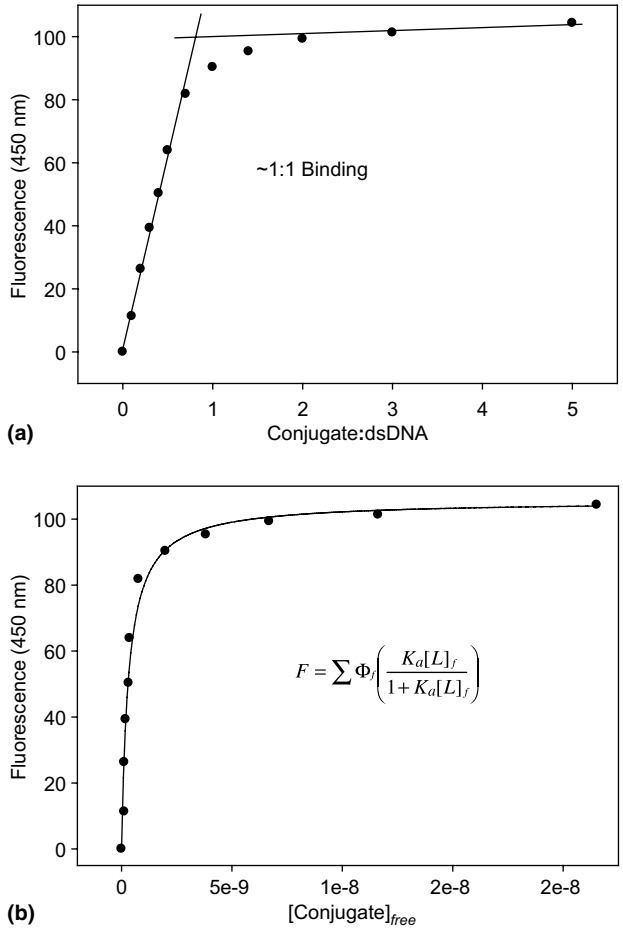


Figure 2. Titration of a 5 nM dsDNA (entry 1) with relatively concentrated conjugate **1**. Plots show (a) fluorescence intensity (arbitrary units) versus conjugate **1**:dsDNA ratio; (b) isothermal binding curve generated from fluorescence intensity versus concentration of free conjugate **1** as calculated by Eq. 3, and data points were fit using Eq. 1.

benzimidazole analogue to polyamides could greatly enhance cellular uptake while maintaining sequence specificity. Furthermore, these conjugates may be

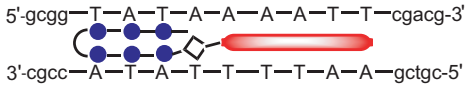


Figure 3. Schematic representation of polyamide–Hoechst 33258 conjugate's binding in hairpin motif in the nine-bp A/T-rich minor groove. Binding site typed in CAPS. The blue circles represent Py, elongated red circle represents Ht and diamond represents linker.

capable of recognizing longer sequences. In order to determine the cellular uptake potential of polyamides upon conjugation with a Hoechst ligand, our initial study focused on the synthesis of hairpin pyrrole polyamide–Hoechst 33258 conjugate **1** (Fig. 1), which is capable of recognizing nine-bp A/T-rich sequence while forming 1:1 conjugate:dsDNA complexes. The fluorescence emission of the conjugate increases greatly upon binding in the minor groove of dsDNA. Thus, spectrofluorometric titrations were employed to determine the conjugate:dsDNA stoichiometries and binding affinities. Thermal denaturation experiments were also performed as an alternative method for the determination of conjugate:dsDNA complex stabilities.

2. Synthesis

The solid-phase synthesis of conjugate PyPyPy-γ-Py-PyPy-γ-Ht (**1**), was accomplished manually in a stepwise manner on rink amide MBHA resin (0.5 mmol/g loading sites) by employing Fmoc technique and a series of PyBOP/HOBt mediated coupling reactions as described in Scheme 1. The monomers Fmoc-Py-acid (**2**)[13] and Hoechst 33258 acid (**4**)[14] were synthesized as reported, and Fmoc-γ-aminobutyric acid (**3**) was purchased from Novabiochem. The Fmoc protection on the resin was removed by treatment with a solution of 20% piperidine in anhydrous DMF. The in situ activation and coupling reaction of Fmoc-Py-acid (**2**) was achieved in 12 h in the presence of HOBt, PyBOP, and DIPEA in anhydrous DMF. After the coupling reaction, any unreacted amine

Table 1. Melting temperatures[a] (T_m in °C) and equilibrium association constants for complexation[b] (K_a) of ligand:dsDNA complexes

Entry	dsDNA	T_m^0	ΔT_m		K_a (M^{-1})
			1	**4**	
1	5′-gcggATATAAAATTcgacg-3′	64	25	5	3.0×10^9
2	5′-gcggTATAAAATTcgacg-3′	63	25	5	2.9×10^9
3	5′-gcgg**C**ATAAAATTcgacg-3′	65	17	5	8.0×10^8
4	5′-gcggT**G**TAAAATTcgacg-3′	64	17	4	8.4×10^8
5	5′-gcggTAT**G**AAATTcgacg-3′	65	16	4	7.7×10^8
6	5′-gcggTATAA**G**ATTcgacg-3′	64	15	2	2.0×10^8
7	5′-gcggTATAAAA**C**TTcgacg-3′	65	14	2	2.4×10^8
8	5′-gcggTATA**GG**AATTcgacg-3′	67	15	3	1.5×10^8
9	5′-gcggTAT**CAG**ATTcgacg-3′	68	12	2	1.2×10^8
10	5′-gcggTA**C**AAA**G**TTcgacg-3′	68	11	1	1.0×10^8
11	5′-gcggT**C**TAAAA**G**Tcgacg-3′	68	10	2	9.1×10^7

[a] All T_m values are average of at least two determinations and standard deviations are ±1 °C. T_m^0 values are T_m values of dsDNA in the absence of ligand, and ΔT_m values are differences in T_m values of dsDNA in the presence and absence of ligand.

[b] The reported K_a constants are the average values obtained from three spectrofluorimetric titration experiments and standard deviations are of ±20%. DNA duplex binding site is typed in CAPS.

sites were blocked by a capping reaction with acetic anhydride and TEA in DMF, rendering them inert toward further chain extension. The terminal Fmoc was deprotected with 20% piperidine in DMF and the coupling yield was determined to be 98% by monitoring the absorbance of deprotected Fmoc at 290 nm. The coupling cycle (coupling/capping/deprotection) was repeated six more times with Fmoc-Py-acid (**2**) or Fmoc-γ-aminobutyric acid (**3**) in the order shown in Scheme 1 to afford the desired hairpin polyamide. The coupling yield for each cycle was found to be 95–100%. The final coupling with the Hoechst 33258 acid (**4**) was achieved in 24 h in the presence of HOBt, PyBOP, and DIPEA in anhydrous DMF. Cleavage of the conjugate from the resin was complete in 2 h using TFA containing 1% of triisopropylsilane (TIS) solution. The crude product was purified by RP-HPLC using a C_8 column with an increasing gradient of acetonitrile in water containing 0.1% TFA. After purification, the product was reconstituted in a minimal amount of methanol and precipitated out of solution by the addition of ether. Product purity was checked by analytical RP-HPLC using the same column and solvent system. ESI/TOF+ mass analysis exhibited the expected peaks at m/z 1327.64 (M+H) and 664.31 (M+2H); calcd 1327.59 (M+H) and 664.30 (M+2H) for $C_{69}H_{74}N_{20}O_9$.

3. Results and discussion

The fluorescence emission of conjugate **1** increases tremendously upon complexation with dsDNA. When excited at 345 nm, conjugate **1**:dsDNA complexes emit a broad fluorescence signal centered at 450 nm, which is consistent with the fluorescence signal emitted by Ht33258:dsDNA complexes at 445 nm.[6,15] Titration of dsDNA with a relatively concentrated solution of conjugate **1** provided a well-defined titration curve that allowed the determination of conjugate:dsDNA complex stoichiometry and equilibrium association constants (K_a). Figure 2 shows a representative plot used for the determination of conjugate **1**:dsDNA stoichiometry and K_a constant. All of the oligomeric duplexes investigated (Table 1) were observed to form 1:1 complexes with conjugate **1**. Our previous study shows that the tripyrrole–Hochst conjugate binds to the nine-bp A/T-rich site in a 2:1 stoichiometry of conjugate:dsDNA complex.[11] In contrast, addition of another tripyrrole unit via γ-aminobutyric acid linker (conjugate **1**) resulted in a 1:1 stoichiometry of conjugate:dsDNA complex (Fig. 2). These results suggest that the polyamide of conjugate **1** is adapting a hairpin conformation in the nine-bp site and binding in a 1:1 stoichiometry as shown in Figure 3. As can be seen from the K_a data in Table 1, conjugate **1** is highly selective ($K_a = 2.9 \times 10^9$ M^{-1}) for the nine-bp A/T-rich binding site 5′-gcggTATAAAATTcgacg-3′ (entry 2, binding site in caps). Little change in K_a was observed when the A/T-rich binding site length was expanded from 9- to 10-bps (5′-gcggATATAAAATTcgacg-3′, entry 1). However, a considerable drop in K_a is observed when a minor groove shorter than nine-bps was present (entry 3). Furthermore, incorporation of a single G/C bp mismatch in the preferred nine-bp A/T-rich binding site resulted in a considerable decrease in binding affinity. A G/C bp mismatch at the 3′-terminus (entries 6 and 7) exhibited a more pronounced effect than the presence of a mismatch at the 5′-terminus (entries 3–5) of dsDNA. Two G/C bp mismatches (entries 8–11) in the binding site causes further decrease in K_a by ~20–30-fold. However, Hoechst 33258 shows no such selectivity.[11]

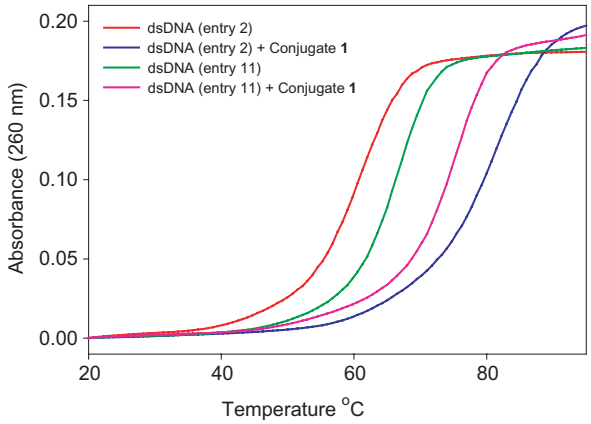


Figure 4. Normalized T_m curves for dsDNA (entries 2 and 11) in the absence and presence of 1 equiv of conjugate **1** in 10 mM $KHPO_4$ buffer pH 7.0 containing 150 mM KCl.

Thermal denaturation experiments were employed as an alternative method for the investigation of conjugate:dsDNA complex stabilities and sequence selectivities. As shown in Table 1, the T_m data strongly correlates with the data acquired via the equilibrium assays. The conjugate **1** form significantly more stable complexes than does Hoechst 33258 acid (**4**) (Table 1). Conjugate **1** forms stable complex ($\Delta T_m = 25$ °C) with the preferred nine-bp dsDNA sequence 5′-gcggTATAAAATTcgacg-3′ (Fig. 4). The stability of the complex suffered greatly upon substitution of a G/C bp mismatch in the nine-bp A/T-rich binding site (Table 1, entries 3–7). In agreement with K_a data, ΔT_m values were affected more strongly (Fig. 4) when two G/C bps are present in the binding site (Table 1, entries 8–11). In contrast, Hoechst 33258 acid (**4**) exhibited fairly similar stability with all dsDNAs investigated (Table 1). Thus, from K_a and ΔT_m values, it is clear that conjugate **1** is able to recognize nine contiguous A/T bps and discriminate with the shorter sequences containing fewer than nine contiguous A/T bps while forming 1:1 conjugate:dsDNA complexes.

4. Conclusions

In conclusion, we have synthesized a hairpin pyrrole polyamide–Hoechst 33258 conjugate and shown that it selectively recognizes a nine-bp A/T-rich sequence. The synthesis of conjugates possessing both pyrrole and

imidazole units to target sequences containing G/C bps, their cellular uptake, as well as their ability to inhibit DNA–TF complex formation is in progress, and will be published separately.

5. Binding constants

Conjugate:dsDNA stoichiometries and binding constants (K_a) were determined by fluorescence titrations. A constant concentration (5 nM) of dsDNA in 10 mM potassium phosphate buffer pH 7.0 containing 150 mM KCl was titrated with aliquots of a relatively concentrated solution of ligand. The fluorescence (excitation at 345 nm and emission at 450 nm) was measured on a Perkin–Elmer LS50B fluorophotometer equipped with a constant temperature water bath set at 25 °C. The point of intersection of the two straight lines generated from pre- and post-saturation data points provided the stoichiometry of ligand:dsDNA binding (1:1, Fig. 2). Equilibrium association constants (K_a) for 1:1 ligand:dsDNA complexes were calculated by generating and fitting isothermal binding curves using Eqs. 1–3.[11,16] Eq. 1 was employed to fit plots of fluorescence versus concentration of unbound ligand, $[L]_f$, where $[L]_f$ is calculated by Eq. 3 (derived from Eq. 2). The derivation and use of Eqs. 1–3 have been previously discussed by our laboratory.[6c,16]

$$F = \sum \Phi_f \left(\frac{K_a[L]_f}{1 + K_a[L]_f} \right) \quad (1)$$

$$F = \sum \Phi_f \frac{[L]_{Bound}}{n[DNA]_T} \quad (2)$$

$$[L]_f = [L]_T - \frac{n[DNA]_T F}{\sum \Phi_f} \quad (3)$$

In Eq. 1, $\sum \Phi_f$ is the total fluorescence intensity upon saturation of dsDNA binding sites with ligand, K_a is the equilibrium association constant for the first binding event, and $[L]_f$ is the concentration of ligand free in solution. In Eq. 2, $[L]_{Bound}$ is the concentration of ligand bound to dsDNA, n is the stoichiometry of binding, and $[DNA]_T$ is the total concentration of duplex DNA in the sample. In Eq. 3 $[L]_T$ is the total ligand concentration added to the constant concentration of dsDNA.

6. T_m experiments

The thermal stability of ligand:dsDNA complexes was investigated by thermal denaturation experiments using a Cary 100 Bio UV–vis spectrophotometer equipped with a temperature programmable cellblock. All T_m experiments for each oligomeric duplex (2 μM), in 10 mM potassium phosphate buffer pH 7.0 containing 150 mM KCl, were carried out in the absence and presence of conjugate **1** (1 equiv) or Hoechst 33258 acid **4** (2 equiv). Data points were taken, between 20 and 95 °C, for every 1 °C with a temperature ramp of 0.5 °C min^{-1}, and T_m temperatures were calculated by first-derivative analysis.

Acknowledgements

This work was supported by a grant from the National Institute of Health (5R37DK09171-36).

References and notes

1. Latchman, D. S. *Gene regulation: a eukaryotic perspective*, 3rd ed.; Stanley Thornes: Cheltenham, 1998.
2. Bailly, C.; Chaires, J. B. *Bioconjugate Chem.* **1998**, *9*, 513.
3. Reddy, B. S. P.; Sondhi, S. M.; Lown, J. W. *Pharmacol. Ther.* **1999**, *84*, 1.
4. Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215.
5. Neidle, S. *Nat. Prod. Rep.* **2001**, *18*, 291.
6. (a) Chiang, S.-Y.; Bruice, T. C.; Azizkhan, J. C.; Gawron, L.; Beerman, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2811; (b) He, G. X.; Browne, K. A.; Blasko, A.; Bruice, T. C. *J. Am. Chem. Soc.* **1994**, *116*, 3716; (c) Browne, K. A.; He, G. X.; Bruice, T. C. *J. Am. Chem. Soc.* **1993**, *115*, 7072.
7. (a) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284; (b) Dervan, P. B.; Burli, R. W. *Curr. Opin. Chem. Biol.* **1999**, *3*, 688.
8. (a) Gottesfeld, J. M.; Neeley, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1997**, *387*, 202; (b) Mapp, A. M.; Ansari, A. Z.; Ptashne, M.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3930.
9. Chiang, S. Y.; Burli, R. W.; Benz, C. C.; Gawron, L.; Scott, G. K.; Dervan, P. B.; Beerman, T. A. *J. Biol. Chem.* **2000**, *275*, 24246.
10. (a) Holmquist, G. *Chromosoma* **1975**, *49*, 333; (b) Arndt-Jovin, D. J.; Jovin, T. M. *Cytometry* **1990**, *11*, 80; (c) Frau, S.; Bernadou, J.; Meunier, B. *Bull. Soc. Chim. Fr.* **1996**, *133*, 1053.
11. (a) Satz, A. L.; Bruice, T. C. *J. Am. Chem. Soc.* **2001**, *123*, 2469; (b) Satz, A. L.; Bruice, T. C. *Bioorg. Med. Chem.* **2002**, *10*, 241; (c) Reddy, P. M.; Jindra, P. T.; Satz, A. L.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 7843.
12. (a) White, C. M.; Satz, A. L.; Bruice, T. C.; Beerman, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10590; (b) White, C. M.; Satz, A. L.; Gawron, L. S.; Bruice, T. C.; Beerman, T. A. *Biochim. Biophys. Acta* **2001**, *1574*, 100.
13. Wurtz, N. R.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Org. Lett.* **2001**, *3*, 1201.
14. Reddy, P. M.; Bruice, T. C. *J. Am. Chem. Soc.* **2004**, *126*, 3736.
15. (a) Loontiens, F. G.; Regenfuss, P.; Zechel, A.; Dumortier, L.; Clegg, R. M. *Biochemistry* **1990**, *29*, 9029; (b) Bostock-Smith, C. E.; Searle, M. S. *Nucl. Acids Res.* **1999**, *27*, 1619; (c) Haq, I.; Ladbury, J. E.; Chowdhry, B. Z.; Jenkins, T. C.; Chaires, J. B. *J. Mol. Biol.* **1997**, *271*, 244; (d) Loontiens, F. G.; McLaughlin, L. W.; Diekmann, S.; Clegg, R. M. *Biochemistry* **1991**, *30*, 182.
16. Satz, A. L.; Bruice, T. C. *Bioorg. Med. Chem.* **2000**, *8*, 1871.