Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry Letters 16 (2006) 6255–6261
Bioorganic & Medicinal Chemistry Letters

DNA sequence recognition in the minor groove by hairpin microgonotropens

Alexandra L. Kahane and Thomas C. Bruice*

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

Received 5 August 2006; revised 1 September 2006; accepted 7 September 2006
Available online 10 October 2006

Abstract—Two novel microgonotropens (MGTs) comprised of hairpin *N*-propylaminepyrrole polyamides linked to a Hoechst 33258 (Ht) analogue (**3** and **4**) were synthesized on solid phase by adopting an Fmoc technique using a series of HOBt mediated coupling reactions. The dsDNA-binding properties of MGTs **3** and **4** were determined by thermal denaturation experiments. Both MGTs were found to be selective for their nine-bp match dsDNA sequence **9** and were less tolerant of G/C bp substitutions in the binding region than linear progenitor MGT **1**. MGT **3** was intolerant of a G/C substitution located in the middle of the binding region and did not bind to sequences **13** and **14**. MGT **4** also did not bind to sequence **13**, and its linker-bound Ht moiety was found to be more sensitive to a G/C substitution in the Ht-binding target, as demonstrated by the lack of binding to sequence **16**.

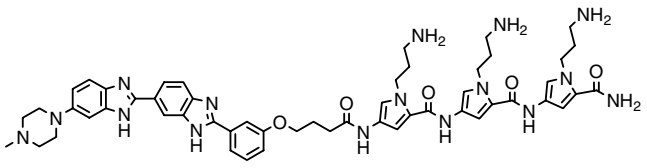
Studies indicate that low molecular weight minor groove-binding agents may bind to specific sequences of dsDNA and influence gene expression by inhibiting the formation of key transcription factor (TF):DNA complexes in a target promoter region, thus impeding the binding of RNA polymerases.[1–7] The development of DNA-binding ligands capable of recognizing long sequences with high affinity and improved sequence specificity is essential to control a specific gene's expression, thus curing the disease rather than simply treating the symptoms. Dervan has reported that synthetic polyamides containing *N*-methylpyrrole (Py) and/or *N*-methylimidazole (Im) amino acids have been shown to bind the minor groove of DNA with sequence specificities comparable to those of natural DNA-binding proteins.[2]

Our interest in the control of gene expression via inhibition of TF binding has led to the development of microgonotropens (MGTs).[6,8] An MGT is comprised of a DNA minor groove-binding moiety attached to a basic amine side chain ($pK_a \geqslant 9$) capable of reaching into the major groove and electrostatically interacting with the acidic phosphodiester backbone. Like *N*-methylpyrrole polyamides, the MGT class of compounds forms hydrogen bonds in the minor groove, but in addition, the basic side chain also extends up and interacts electrostatically with the phosphate backbone of DNA.[9] The basic tails endow the MGTs with the ability to bend DNA, which helps them to more effectively compete with transcription factors for binding to the target sites.[7,10] These attributes result in increased ligand-binding affinity and enhanced inhibition of transcription factor binding in cell-free assays.[11] However, most MGTs were unable to inhibit endogenous gene expression, presumably due to poor cellular uptake of these ligands and their inability to achieve nuclear localization.[5]

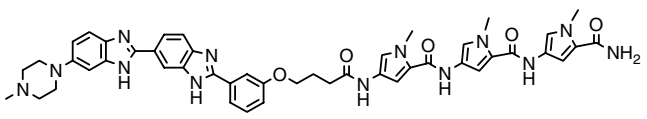
In 2001, we reported the synthesis[12] and biological evaluation[13] of *N*-propylaminepyrrole (prPy)/Hoechst (Ht) conjugate **1** (Fig. 1), the first MGT capable of inhibiting gene expression in whole cells, as well as cell-free assays. Conjugate **1** was shown to bind DNA at subpicomolar concentrations and was at least three orders of magnitude more potent than its *N*-methylpyrrole/Ht analogue (**2**) at inhibiting TF binding to the c-fos serum response element (SRE) in cell-free assays. Unlike any other MGT, **1** also inhibited endogenous c-fos expression in NIH 3T3 cells at micromolar concentrations. These findings are significant because they represent the first minor groove-binding agent possessing the important characteristics of cell penetration and endogenous inhibition of protein expression.

Keywords: Hairpin polyamide-Hoechst 33258 conjugates.
* Corresponding author. Tel.: +805 893 2044; fax: +805 893 2229; e-mail: tcbruice@chem.ucsb.edu

0960-894X/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2006.09.032



1



2

Figure 1. Structure of MGT **1** and its *N*-methylpyrrole:Ht analogue **2**.

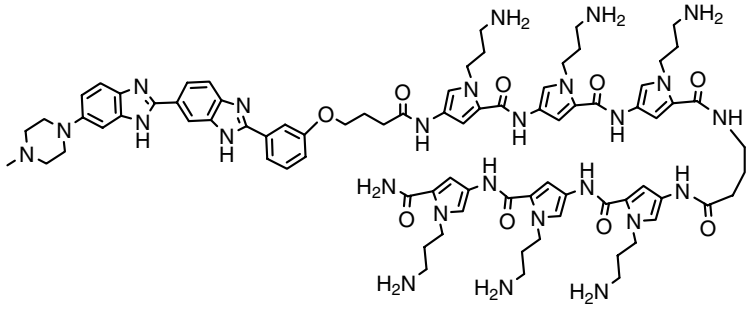
Structure–reactivity relationships[13] suggest that although the tripyrrole–polyamide functionality of **1** may be largely responsible for its inhibition of TF complexes in cell-free assays, its Ht moiety appears to enable cellular uptake and biological activity in whole cells. Thus, subsequent MGTs should contain both of these features. Recently, we have published a series of papers describing the synthesis and DNA-binding affinities of hairpin-shaped *N*-methylpyrrole–imidazole–Ht conjugates.[14–16] These ligands were formed by the head-to-tail linkage of *N*-methylpyrrole–imidazole polyamides with γ-aminobutyric acid (γ) to provide hairpin polyamides that mimic the 2:1 side-by-side anti-parallel binding of the unlinked polyamides.[17] In general, the binding affinity of hairpin structures is 100-fold enhanced relative to that of the unlinked polyamides.[17–23]

In an effort to develop novel MGTs with the desirable characteristics of high specificity, longer-binding sites, and increased probability of cellular uptake, a new class of prPy:Ht minor groove-binding agents has been designed (Fig. 2). These new MGTs combine the prPy and Ht moieties essential for endogenous c-fos expression inhibition with the hairpin scaffold which provides increased binding affinity and 1:1 binding stoichiometries. MGT **3** is based on the common hairpin framework formed by the simple head-to-tail linkage of three prPy polyamides, a γ linker, followed by three more prPy polyamides and conjugation with Ht.

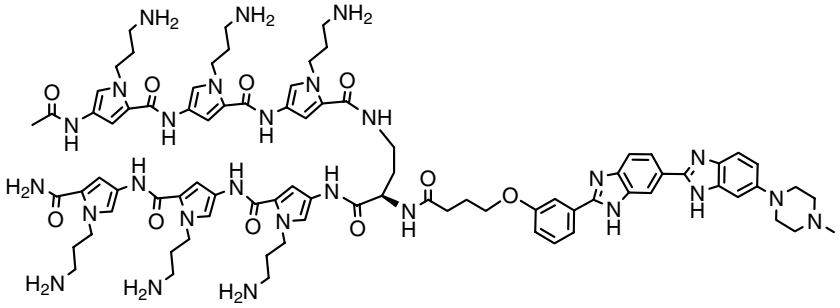
Recent studies have shown that incorporation of bulky groups on the linker-turn forces a polyamide to adopt a hairpin motif instead of an extended motif.[15,21,24–26] The linkage of the Hoechst ligand at the hairpin turn appears to force the polyamide to adopt the hairpin motif irrespective of minor groove length and provide side-by-side amino acid pairings,[15] which is a prerequisite for sequence specific recognition. Incorporation of a bulky group on the linker-turn forms a new stereocenter, and both computational and binding affinity studies have found that (*R*)-configured amine γ-turn substituents produce superior DNA minor groove-binding agents.[15,25] MGT **4** incorporates these findings into our current application. The synthesis of novel MGTs **3** and **4**, and their dsDNA-binding properties are discussed herein.

The solid-phase syntheses of conjugates **3** and **4** were accomplished manually in a stepwise manner on rink amide MBHA resin (100–200 mesh, 0.5 mmol/g loading sites) by employing Fmoc technique with a series of HOBt mediated coupling reactions as described in Schemes 1 and 2. The Fmoc–prPy–OPfp (**5**)[12,27] and Hoechst 33258 acid (**7**)[28] building blocks were synthesized as reported, γ-hydroxybutyric acid linker (**6**) was purchased from Novabiochem, and masked 2,4-diaminobutyric acid derivative (Fmoc-D-Dab(ivDde)-OH) (**8**) was purchased from Chem-Impex. The solid-phase synthesis of MGT **3** is shown in Scheme 1. Coupling reactions for **5** were accomplished using 3 equiv of **5**, 5 equiv of HOBt, and 10 equiv of DIPEA in anhydrous DMF, and were run for 24 h. High coupling yields (95–100%) were measured by absorption at 290 nm of the dibenzofulvene produced from the Fmoc deprotection after treating with 20% piperidine–DMF solution. Coupling reactions for **6** were accomplished using 5 equiv of **6**, 7 equiv PyBOP, 7 equiv HOBt, and 15 equiv DIPEA in anhydrous DMF, and were run 24 h with coupling yields of 95%. After each coupling, unreacted terminal amines were capped with a DMF solution of acetic anhydride and TEA. Coupling reactions for **7** were accomplished employing 3 equiv of **7**, 6 equiv of HOBt and PyBOP, and 10 equiv of DIPEA, and were run for 48 h. Resin cleavage and concurrent removal of the Boc protecting groups were achieved in 2–4 h using TFA containing 1% TIS. All final products synthesized via SPS were purified via preparative HPLC (silica, reverse phase, C8 column) with an increasing gradient of acetonitrile in 0.1% aq TFA solution and lyophilized to dryness. Product purity was checked by analytical RP-HPLC using the same column and solvent system.[29]

The solid-phase synthesis of MGT **4** was accomplished in a similar fashion and is shown in Scheme 2. The coupling of Fmoc–prPy–OPfp (**5**) was achieved after 24 h in the presence of HOBt and DIPEA in anhydrous DMF. After the coupling reaction, any unreacted amine sites were capped by acetylation. The Fmoc protection on the *N*-propylaminepyrrole was removed and the coupling cycle (coupling/capping/deprotection) was repeated two more times with Fmoc–prPy–OPfp (**5**) before introducing orthogonally protected 2,4-diaminobutyric acid (Fmoc-D-Dab(ivDde)-OH) (**8**). The Fmoc protection of the linker α-amino group was removed and coupled with Hoechst 33258 acid (**7**). After the capping reaction, the ivDde protection on the γ-amino group was removed using 5% hydrazine in DMF and the coupling cycle was repeated three more times with Fmoc–prPy–OPfP (**5**). The Fmoc on the last pyrrole was removed and the amine was capped to obtain conjugate **4**. The coupling yield in each cycle was found to be 95–99%, as determined from the absorbance of dibenzofulvene from the deprotected Fmoc. The conjugate was cleaved from the resin in 2 h using 1% TIS in TFA, during which time, the Boc groups were concurrently removed. The cleaved product was purified as 2described above.[30]

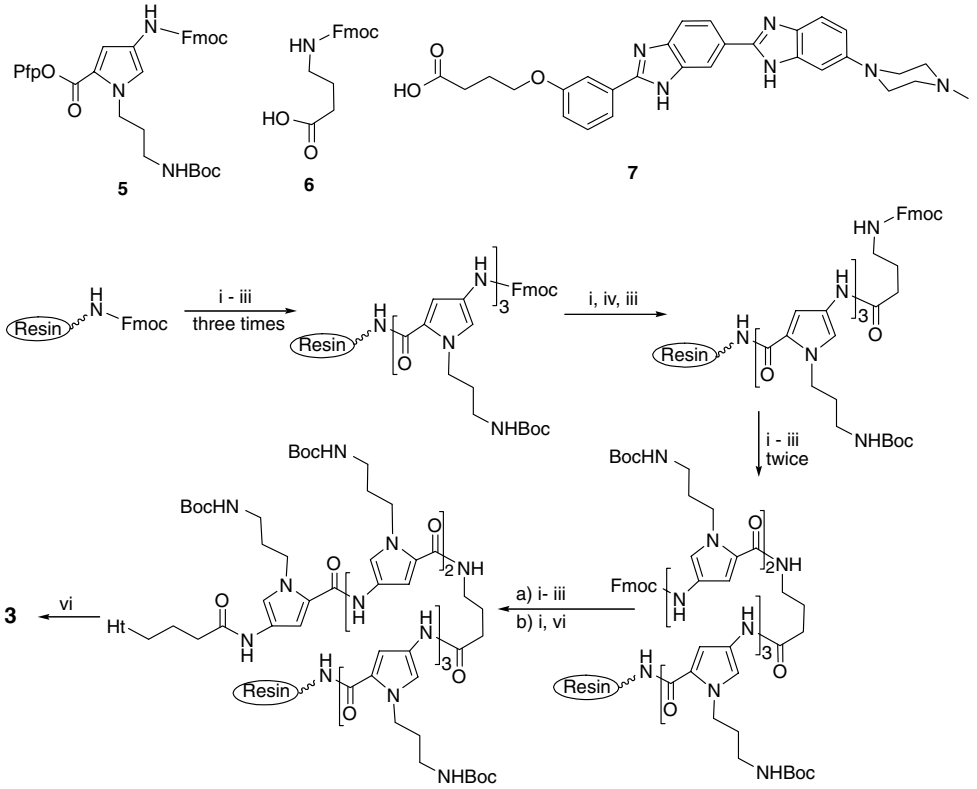


3

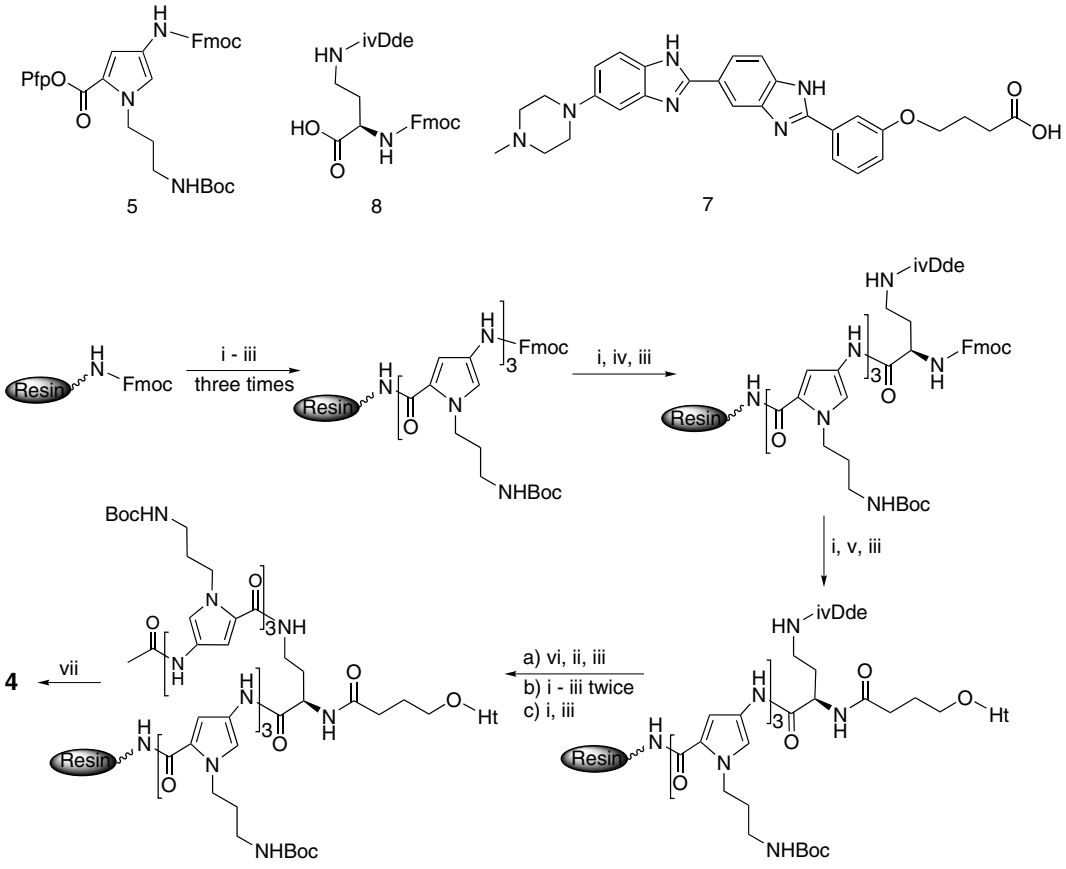


4

Figure 2. Hairpin *N*-propylaminepyrrole:Ht MGTs **3** and **4**.



Scheme 1. Solid phase synthesis of MGT **3**. Reagents and conditions: (i) *Deprotection:* 20% piperidine/DMF, 15 min; (ii) *Coupling:* Fmoc-prPy-OPfp (**5**), HOBt, DIPEA, DMF, 24 h; (iii) *Capping:* acetic anhydride, TEA, DMF, 10 min; (iv) Fmoc-γ-aminobutyric acid (**6**), PyBOP, HOBt, DIEA, DMF, 24 h; (v) Hoechst 33258 acid (**7**), PyBOP, HOBt, DIPEA, DMF, 48 h; (vi) *Cleavage*: TFA, TIS, 2h. Ht = Hoechst 33258.



Scheme 2. Solid phase synthesis of MGT **4**. Reagents and conditions: (i) *Deprotection:* 20% piperidine/DMF, 15 min; (ii) *Coupling:* Fmoc-prPy-OPfp (**5**), HOBt, DIPEA, DMF, 24 h; (iii) *Capping:* acetic anhydride, TEA, DMF, 10 min; (iv) Fmoc-D-Dab(ivDde)-OH (**8**), PyBOP, HOBt, DIEA, DMF, 24 h; (v) Hoechst 33258 acid (**7**), PyBOP, HOBt, DIPEA, DMF, 48 h; (vi) *Deprotection*: 5% hydrazine in DMF (2 × 5 min); (vii) *Cleavage:* TFA, TIS, 2 h. Ht = Hoechst 33258.

Thermal denaturation experiments were employed to investigate the dsDNA:MGT complex stabilities and sequence selectivities of each MGT (Table 1).[31] The sequence selectivity of each ligand was determined by investigating the binding affinities to the 18-bp dsDNA by gradually changing the position of a G/C base pair (bp) substitution in the optimal (or ''match'') nine-bp binding site found in dsDNA **9**. The thermodynamic data for the melting transitions of dsDNA and ligand-bound dsDNA were calculated using the method of Marky and Breslauer.[32]

A noteworthy testament to the remarkable DNA-binding affinity of MGTs **3** and **4** was the need to perform the thermal denaturation studies at a decreased ionic strength ($\mu = 0.03$). Typically, these experiments are performed close to physiological ionic strength ($\mu = 0.17$), however, under these conditions, the T_m values of the ligand-bound DNA were too close to the solution boiling point to be accurately determined (Fig. 3). A decreased μ lowers the T_m of dsDNA, but the binding of **3** and **4** to dsDNA is strengthened. Performing the thermal denaturation studies at a lower ionic strength decreased the observed dsDNA:MGT T_m values thus ensuring accurate determination and free energy calculations. To allow accurate binding affinity comparisons, MGT **1** was reexamined under these conditions.

The largest ΔT_m for each ligand is expected to occur upon its binding in the minor groove to its match sequence, dsDNA **9** (Fig. 4). Indeed, MGTs **1**, **3**, and **4** each bind specifically to **9**, indicated by displaying their largest ΔT_m values (Table 1) and most favorable $\Delta\Delta G$ values (Table 2). As the G/C substitution is moved into the binding region, significant drops in ΔT_m and $\Delta\Delta G$ values are observed. MGT **3** demonstrates the lowest tolerance for a G/C bp in the binding region. MGT **3** does not bind to dsDNA sequences **13** and **14** which have the G/C substitution located in the center of the binding region. MGT **4** also does not bind to **13**, but the **4**/**14** complex has a small ΔT_m of 8 °C. Although when bound to sequence **14**, **1**, and **4** show a decrease in ΔT_m of 9° and 18°, respectively, they do not possess the superior selectivity for their target sequence that **3** does.

Previous studies conducted in our laboratory[14,16] have suggested that the dsDNA-binding affinity of hairpin Ht-polyamides is driven by both the Ht recognizing its target sequence (AATT), and the polyamide interactions within the DNA minor groove. Dependence of the binding affinity of Ht-polyamides on the presence of the Ht target sequence is usually evident from the significantly smaller values of ΔT_m for dsDNA sequences which lack the AATT sequence (i.e., **15**–**17**). Indeed,

Table 1. Melting temperatures[a]

dsDNA	T_m°	T_m			ΔT_m		
		1	**3**	**4**	**1**	**3**	**4**
5′- gcggTATAAAATTcgacg-3′ (**9**)	**49**	**74**	**74**	**75**	**25**	**25**	**26**
5′-gcgg**C**ATAAAATTcgacg-3′ (**10**)	52	77	76	75	25	24	23
5′-gcgg**T**GTAAAATTcgacg-3′ (**11**)	52	72	73	73	20	21	21
5′-gcggTA**C**AAAATTcgacg-3′ (**12**)	52	71	72	74	19	20	22
5′-gcggTAT**G**AAATTcgacg-3′ (**13**)	51	61	50	51	10	0	0
5′-gcggTATA**G**AATTcgacg-3′ (**14**)	50	66	49	58	16	0	8
5′-gcggTATAA**G**ATTcgacg-3′ (**15**)	50	63	63	70	13	13	20
5′-gcggTATAAAA**C**Tcgacg-3′ (**16**)	53	70	71	52	17	18	0
5′-gcggTATAAAAT**C**cgacg-3′ (**17**)	50	75	73	73	25	23	23

[a] All T_m values are the average of at least three determinations and standard deviations are ±1 °C. T_m° values are melting temperature values of dsDNA in the absence of ligand, and ΔT_m values are differences in melting temperature values of dsDNA in the presence and absence of ligand (T_m in °C). The binding region of each sequence is shown capitalized.

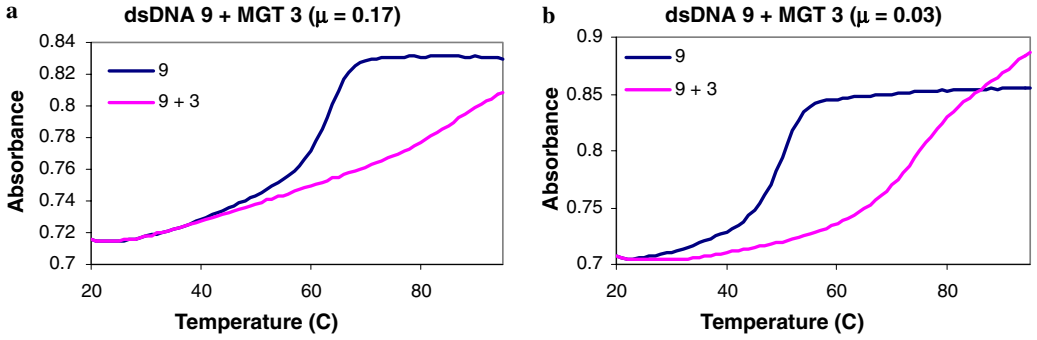


Figure 3. (a) dsDNA **9** + **3**, $\mu = 0.17$ (b) dsDNA **9** + **3**, $\mu = 0.03$.

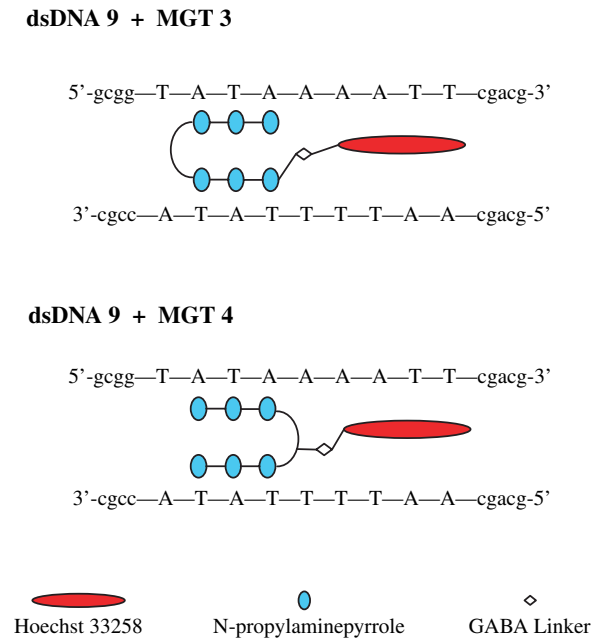


Figure 4. Schematic representation of MGTs **3** and **4** bound to their match sequence. The binding regions of the dsDNA sequences are shown capitalized.

the smallest ΔT_m values are usually observed when the location of the G/C bp substitution interrupts the Ht target sequence. However, the smallest ΔT_m values for MGTs **1** and **3** are observed when the G/C bp is located in the portion of the dsDNA sequence bound by the polyamide moiety. In contrast, MGT **4** shows no binding to **16**, indicating that it is more sensitive to the G/C bp located in the middle of the Ht target sequence than is the binding of MGTs **1** and **3**. Compared to binding with match sequence **9**, although a smaller ΔT_m is observed for each MGT binding to **15** and **16**, little decrease in ΔT_m is observed upon binding with **17**. With the exception of the **4/16** complex, this implies that the complexation of the pyrrole polyamides and subsequent interaction of the propylamine side chains with the phosphate backbone of DNA imparts more stability than binding of the Ht moiety in the minor groove.

In conclusion, in an effort to develop MGTs capable of recognizing long sequences with high affinity, improved sequence specificity, and the increased probability of cellular uptake, prPy:Ht conjugates **3** and **4** were synthesized and evaluated via thermal denaturation experiments. Both MGTs were found to be selective for their match dsDNA sequence **9**, as larger increases in the ΔT_m and the $\Delta\Delta G$ values were observed when MGTs **3** and **4** were interacting with sequence **9** compared to the values observed when each MGT interacted with sequences containing a G/C substitution. When bound to their match sequence **9**, the resultant complexes exhibited T_m values 25–26 °C higher than unbound dsDNA **9**, demonstrating the high binding affinity of MGTs **3** and **4**. Although

Table 2. Free energy calculations[a]

dsDNA	ΔG (kcal/mol)	$\Delta G'$ (kcal/mol)			$\Delta\Delta G$ (kcal/mol)			ΔT_m		
		1	**3**	**4**	**1**	**3**	**4**	**1**	**3**	**4**
9	−18	−22	−22	−24	−4	−4	−6	25	25	26
10	−19	−23	−24	−23	−4	−5	−4	25	24	23
11	−18	−21	−22	−24	−3	−4	−6	20	21	21
12	−19	−20	−21	−24	−1	−2	−5	19	20	22
13	−20	−15	—	—	5	—	—	10	0	0
14	−18	−20	—	−19	−2	—	−1	16	0	8
15	−19	−18	−19	−23	1	0	−4	13	13	20
16	−21	−20	−20	—	1	1	—	17	18	0
17	−18	−22	−21	−23	−4	−3	−5	25	23	23

[a] ΔG values represent dsDNA in the absence of ligand, $\Delta G'$ values represent dsDNA in the presence of ligand, and $\Delta\Delta G$ represents the differences in the changes of free energy of dsDNA in the presence and absence of ligand. $\Delta G'$ and $\Delta\Delta G$ values were not calculated for non-binding ($\Delta T_m = 0$) ligand–dsDNA combinations.

hairpin-shaped MGTs **3** and **4** exhibited binding affinities similar to linear MGT **1**, **3**, and **4** were found to bind with higher specificities. MGT **3** was intolerant of a G/C substitution located in the middle of the nine-bp binding region and did not bind to sequences **13** and **14**. Decreased ΔT_m values 1–12 °C lower than that of the **3/9** complex were observed when the G/C substitution was located elsewhere in the binding region. MGT **4** also did not bind to sequence **13**, and its linker-bound Ht moiety was found to be more sensitive to a G/C substitution in the Ht-binding target, as demonstrated by the lack of binding to sequence **16**. Decreased ΔT_m values of 3–18 °C lower than that of the **4/9** complex were observed when the G/C substitution was located elsewhere in the binding region. MGTs **3** and **4** are promising candidates for gene suppression due to their high-binding affinity, increased specificity, and the presence of the Ht moiety. Biological evaluations of these agents are currently underway.

Acknowledgment

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

References and notes

1. Reddy, B. S. P.; Sondhi, S. M.; Lown, J. W. *Pharmacol. Ther.* **1999**, *84*, 1.
2. Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215.
3. Bailly, C.; Chaires, J. B. *Bioconjugate Chem.* **1998**, *9*, 513.
4. Neidle, S. *Nat. Prod. Rep.* **2001**, *18*, 291.
5. Chiang, S.-Y.; Bruice, T. C.; Azizkhan, J. C.; Gawron, L.; Beerman, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2811.
6. He, G. X.; Browne, K. A.; Blasko, A.; Bruice, T. C. *J. Am. Chem. Soc.* **1994**, *116*, 3716.
7. Browne, K. A.; He, G. X.; Bruice, T. C. *J. Am. Chem. Soc.* **1993**, *115*, 7072.
8. He, G. X.; Browne, K. A.; Groppe, J. C.; Blasko, A.; Mei, H. Y.; Bruice, T. C. *J. Am. Chem. Soc.* **1993**, *115*, 7061.
9. Blasko, A.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10018.
10. Bruice, T. C.; Mei, H. Y.; He, G. X.; Lopez, V. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 1700.
11. White, C. M.; Satz, A. L.; Gawron, L. S.; Bruice, T. C.; Beerman, T. A. *Biochim. Biophys. Acta* **2001**, *1574*, 100.
12. Satz, A. L.; Bruice, T. C. *Bioorg. Med. Chem.* **2002**, *10*, 241.
13. White, C. M.; Satz, A. L.; Bruice, T. C.; Beerman, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10590.
14. Reddy, P. M.; Dexter, R.; Bruice, T. C. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3803.
15. Reddy, P. M.; Toporowski, J. W.; Kahane, A. L.; Bruice, T. C. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 5531.
16. Correa, B. J.; Canzio, D.; Kahane, A. L.; Reddy, P. M.; Bruice, T. C. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3745.
17. de Clairac, R. P. L.; Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **1997**, *119*, 7909.
18. Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983.
19. Parks, M. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6147.
20. Parks, M. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6153.
21. Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1996**, *3*, 369.
22. Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 8198.
23. Pilch, D. S.; Poklar, N.; Gelfand, C. A.; Law, S. M.; Breslaur, K. J.; Baird, E. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8306.
24. Woods, C. R.; Ishii, T.; Wu, B.; Bair, K. W.; Boger, D. L. *J. Am. Chem. Soc.* **2002**, *124*, 2148.
25. Herman, D. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, *120*, 1382.
26. Herman, D. M.; Baird, E. E.; Dervan, P. B. *Chem. Eur. J.* **1999**, *5*, 975.
27. Xue, T.; Browne, K. A.; Bruice, T. C. *Bioconjugate Chem.* **1995**, *6*, 82.
28. Reddy, P. M.; Bruice, T. C. *J. Am. Chem. Soc.* **2004**, *126*, 3736.
29. **3**: ESI/TOF+ MS *m/z* $(M+3H)^{3+}$ calcd 529.28, obsd 529.28.
30. **4**: ESI/TOF+ MS *m/z* $(M+3H)^{3+}$ calcd 548.64, obsd 548.65.
31. UV/vis spectra were acquired on a Cary 100 Bio UV–vis spectrophotometer equipped with a temperature programmable cellblock. All thermomelting temperature (T_m) experiments were carried out using 10 mM potassium phosphate buffer, pH 7.0, containing 10 mM KCl ($\mu = 0.03$). The concentration of each oligomeric duplex was 2 μM and a series of T_m curves were acquired

(dsDNA + no ligand; dsDNA + **1** or **3** or **4**) using 2 equiv of ligand. Data points were taken every 1 °C from 20 to 95 °C with a temperature ramp of 0.5 °C/min.

32. Marky, L. A.; Breslauer, K. J. *Biopolymers* **1987**, *26*, 1601.