Available online at www.sciencedirect.com
SCIENCE DIRECT
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 16 (2006) 3745–3750

DNA sequence recognition by Hoechst 33258 conjugates of hairpin pyrrole/imidazole polyamides

Bryan J. Correa, Daniele Canzio, Alexandra L. Kahane, Putta Mallikarjuna Reddy and Thomas C. Bruice*

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

Received 7 March 2006; accepted 19 April 2006
Available online 8 May 2006

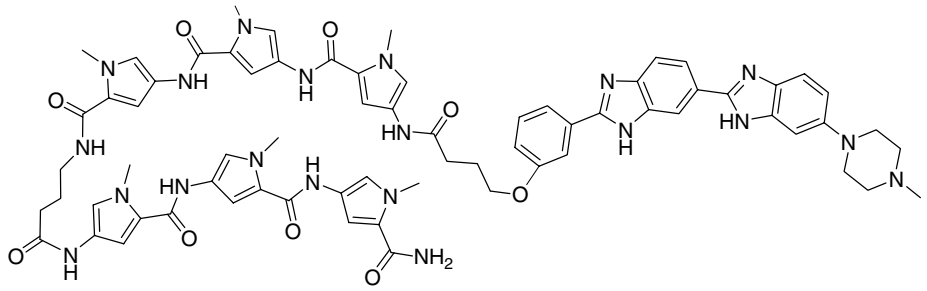
Abstract—A series of hairpin pyrrole/imidazole polyamides linked to a Hoechst 33258 (Ht) analogue (**5**–**7**) were synthesized on solid-phase by adopting an Fmoc technique using a series of PyBOP/HOBt mediated coupling reactions. The dsDNA binding properties of Ht-polyamides **5**–**7** were determined by thermal denaturation experiments. Hairpin Ht-polyamides **5**–**7** bound to dsDNA sequences **16** and **18** show ΔT_m values that are 14–18 degrees higher than linear Ht-polyamides bound to the same sequences. All three Ht-polyamides were found to be selective for their 9-bp match dsDNA sequences, supporting a relative stronger interaction of an Im/Py anti-parallel dimer with an appropriately positioned G/C bp rather than sequences containing only A/T bps. In addition, Ht-polyamides **5** and **7** showed a 20-fold preference for a properly placed G/C bp over a C/G bp, while **6** showed a 10-fold preference.

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 and coworkers report that synthetic polyamides consisting of *N*-methylpyrrole (Py) and *N*-methylimidazole (Im) amino acids bind with sequence specificities comparable to those of natural DNA binding proteins.[2] Furthermore, head-to-tail linkage of such polyamides with γ-aminobutyric acid (γ) provides hairpin polyamides that mimic the 2:1 side-by-side anti-parallel binding of the unlinked polyamides.[8] The binding affinity of these hairpin structures is 100-fold enhanced relative to that of the unlinked polyamides.[8–14] However, attempts to inhibit the transcription of endogenous genes in cell lines have met with little success, presumably due to poor cellular uptake of these polyamides and their inability to achieve nuclear localization.[15]

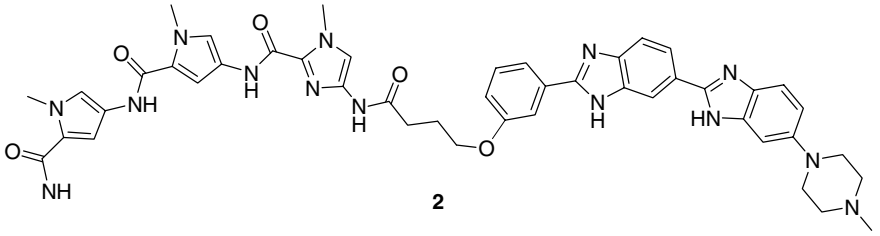
Recent studies have shown that the linkage of polyamides to minor groove binding bis-benzimidazole (biBenz) dyes can recognize longer sequences of DNA and also enhance cellular permeability.[16–20] Linkage of the biBenz dye Hoechst 33258 (Ht) fluorophore to the six-ring hairpin pyrrole polyamide at the N-terminus forms a Ht-polyamide (**1**, Fig. 1) that recognizes a 9-bp A/T-rich site with high affinity and good selectivity.[21] This Ht-polyamide binds to the 9-bp site with 1:1 stoichiometry with the polyamide moiety adopting a hairpin motif. However, the exclusively pyrrole constitution of this Ht-polyamide limits its applications to DNA targets containing only A/T base pairs (bp). To achieve the goal of specifically inhibiting the expression of a misregulated protein, minor groove binding agents must be developed with the ability to recognize G/C bp within their otherwise A/T-rich binding sites. A first step in that direction called for the synthesis of imidazole-containing molecules, which have been previously shown to tolerate, and possibly prefer, a G/C bp.[1–4] The exocyclic amine of guanine presents, on the minor groove edge of a G/C bp, a free hydrogen that can become involved in a hydrogen bond. Crystal structures exploring the hydrogen bond array necessary for specific recognition of the

Keywords: Hairpin polyamide-Hoechst 33258 conjugates; Minor groove binding agents.
* Corresponding author. Tel.: +1 805 893 2044; fax: +1 805 893 2229; e-mail: tcbruice@chem.ucsb.edu

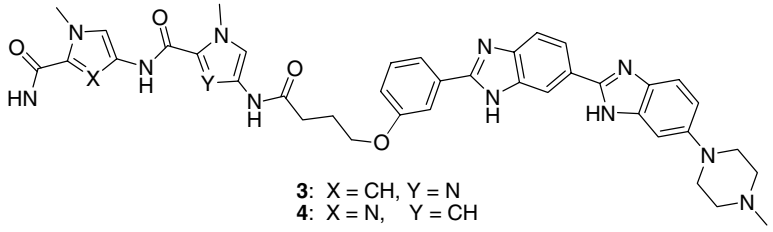
0960-894X/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2006.04.047



1



2



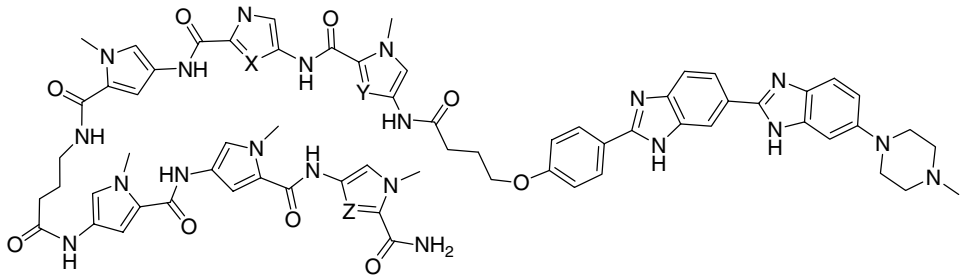
3: X = CH, Y = N
4: X = N, Y = CH

Figure 1. Structures of pyrrole, imidazole Hoechst 33258-polyamides **1**–**4**.

DNA-helix by pyrrole/imidazole polyamides have been reported, providing direct confirmation of the formation of a favorable hydrogen bond between the imidazole N3 and the exocyclic amine of guanine.[22]

In 2003, we reported the synthesis and study of pyrrole/imidazole substituted Ht-polyamides **2**–**4**, with a γ-hydroxybutyric acid linker, capable of recognizing a G/C bp in the 9-bp specific sequences.[18] These conjugates formed 1:1 ligand/dsDNA complexes, binding at subnanomolar concentrations. Typically, the Im residues in polyamides that bind to dsDNA in a 1:1 stoichiometry do not distinguish G/C from A/T, and thus bind to all 4 bps indiscriminately, but with high affinity.[23] However in this case, the Im residues in these Ht-polyamides were found to distinguish G/C from A/T by ~8-fold magnitude.

In an effort to develop novel minor groove binding agents with the desirable characteristics of high specificity, longer binding sites, increased binding affinity and increased probability of cellular uptake, a new class of Ht-polypeptide dsDNA minor groove binding agents has been designed. Replacement of the first pyrrole in **1** with an imidazole residue produces Ht-polyamide **5** (Fig. 2). By varying the position of the imidazole substitution, a series of imidazole-containing hairpin polyamides (**5**–**7**) was created. The sequence selectivity



5: X = CH, Y = N, Z = CH
6: X = N, Y = CH, Z = CH
7: X = CH, Y = CH, Z = N

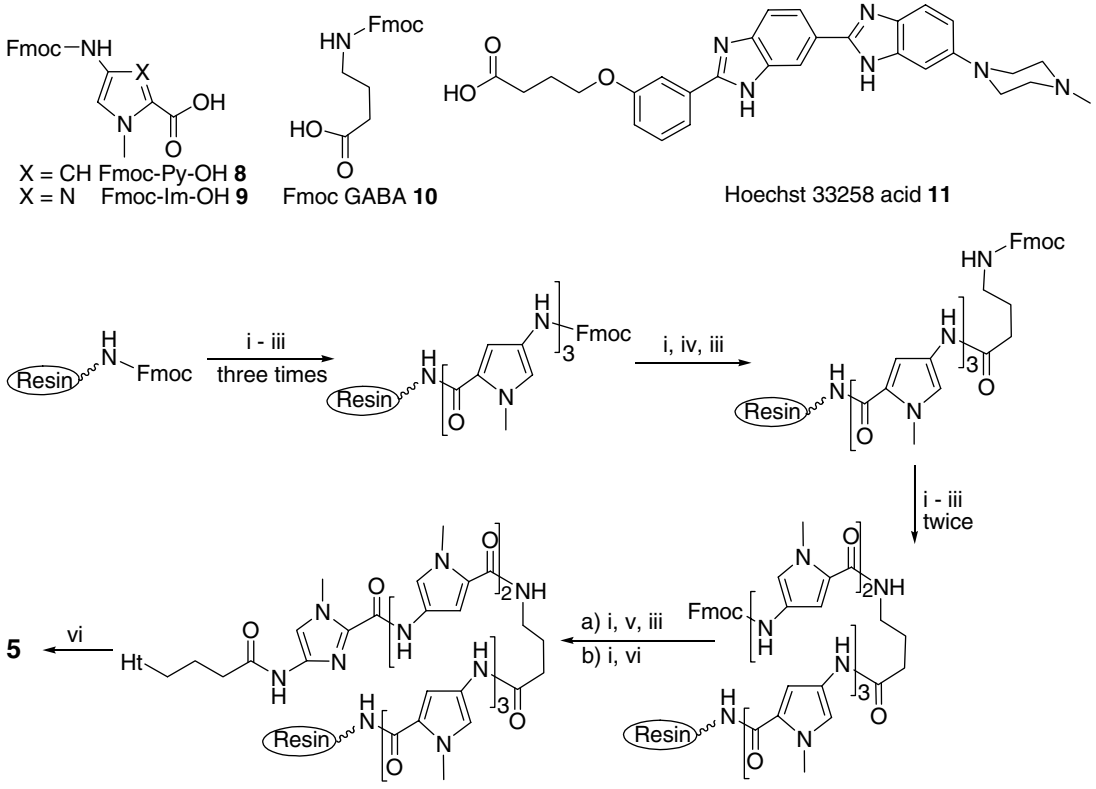
Figure 2. Hairpin pyrrole, imidazole Ht-polyamides **5**–**7**.

of each polyamide was investigated by thermal denaturation experiments. These novel hairpin pyrrole/imidazole Ht-polyamides maintain the sequence selectivity demonstrated by shorter imidazole-containing conjugates, while the hairpin structure imparts increased dsDNA binding affinities.

The solid-phase syntheses of conjugates **5**–**7** were accomplished manually in a stepwise manner on rink amide MBHA resin (0.5 mmol/g loading sites) by employing Fmoc technique with a series of PyBOP/HOBt mediated coupling reactions as described in Scheme 1. The Fmoc-Py-OH (**8**), Fmoc-Im-OH (**9**),[24] and Hoechst 33258 acid (**11**)[25] building blocks were synthesized as reported, and γ-aminobutyric acid linker (**10**) was purchased from Novabiochem. The solid-phase synthesis of the conjugates is shown in Scheme 1. SPS synthesis was accomplished using rink amide MBHA resin (100–200 mesh) and standard Fmoc techniques.[26] Coupling reactions for **8** or **9** were accomplished using 2.5–3 equiv of **8** or **9**, 3.75–6 equiv of HOBt, 3.75–6 equiv of PyBOP, and 10 equiv of DIPEA in anhyd DMF and were run for 12–24 h. High coupling yields (70–100%) were measured by absorption at 290 nm of deprotected Fmoc residue after treating with 20% piperidine/DMF solution. After each coupling, unreacted terminal amines were capped with a DMF solution of acetic anhydride and TEA. Coupling reactions for **11** were accomplished employing 3 equiv of **11**, 6 equiv of HOBt and PyBOP, and 10 equiv of DIPEA, and were run for 48 h. Resin cleavage was achieved in 2–4 h using TFA containing 2.5% TIS. All final products synthesized via SPS were purified via HPLC (silica, reverse phase, C8 column) with an increasing gradient of acetonitrile in 0.1% aq TFA solution and lyophilized to dryness. 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.[27]

Thermal denaturation experiments were employed to investigate the dsDNA/ligand complex stabilities and sequence selectivities of each conjugate (Table 1).[28] The sequence selectivity of each ligand was determined by investigating the binding affinities to the 18-bp dsDNA by gradually changing the G/C position in the 9-bp binding site. The thermodynamic data for the melting transitions of dsDNA and ligand-bound dsDNA were calculated using the method of Marky and Breslauer.[29]

Previous studies conducted in our laboratory[21] 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 **5**–**7** on the presence of the Ht target sequence is evident from the smaller values of ΔT_m for dsDNA sequences **20**–**24**, which each lack the AATT sequence. Amongst the dsDNA sequences



Scheme 1. Reagents and conditions. (i) Deprotection: 20% piperidine/DMF, 15 min; (ii) coupling: Fmoc-Py-OH (**8**), PyBOP, HOBt, DIPEA, DMF, 12 h; (iii) capping: acetic anhydride, TEA, DMF, 10 min; (iv) Fmoc-γ-aminobutyric acid (**10**), PyBOP, HOBt, DIPEA, DMF, 12 h; (v) Fmoc-Im-OH (**9**), PyBOP, HOBt, DIPEA, DMF, 12 h; (vi) Hoechst 33258 acid (**11**), PyBOP, HOBt, DIPEA, DMF, 24 h; (vi) cleavage: TFA, TIS, 2 h. Ht, Hoechst 33258.

Table 1. Melting temperatures[a] (T_m in °C)

dsDNA	T_m^0	T_m			ΔT_m		
		7	**5**	**6**	**7**	**5**	**6**
5′-gcggTATAAAATTcgacg-3′ (**13**)	62	80	81	78	18	19	16
5′-gcgg**C**ATAAAATTcgacg-3′ (**14**)	65	80	80	78	15	15	13
5′-gcgg**T**GTAAAATTcgacg-3′ (**15**)	65	82	80	79	17	15	14
5′-gcggTA**C**AAAATTcgacg-3′ (**16**)	64	80	81	**89**	16	17	**25**
5′-gcggTA**G**AAAATTcgacg-3′ (**17**)	64	78	83	86	14	19	22
5′-gcggTAT**G**AAATTcgacg-3′ (**18**)	64	**84**	87	79	**20**	23	15
5′-gcggTAT**C**AAATTcgacg-3′ (**19**)	64	83	**89**	82	19	**25**	18
5′-gcggTATAA**G**ATTcgacg-3′ (**20**)	64	79	78	78	15	14	14
5′-gcggTATAAAA**C**Tcgacg-3′ (**21**)	65	78	78	77	13	13	12
5′-gcggTATAAAAT**C**cgacg-3′ (**22**)	64	79	79	79	15	15	15
5′-gcggTAT**CAG**ATTcgacg-3′ (**23**)	68	82	83	83	14	15	15
5′-gcgg**TC**TAAAA**G**Tcgacg-3′ (**24**)	69	77	77	79	8	8	10
5′-gcggTATA**GG**AATTcgacg-3′ (**25**)	67	84	84	78	17	17	13

[a] The binding region of each sequence is shown capitalized. All T_m values are the average of at least three determinations and standard deviations are ±1 °C. T_m^0 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.

containing the required Ht target sequence, the largest ΔT_m for each ligand is expected to occur upon its binding in the minor groove to its 'match' sequence (Fig. 3).

Conjugate **5** was found to bind specifically to its match sequence, dsDNA **19**. Substitution of the G/C bp with an A/T bp (dsDNA **13**) resulted in a decrease in ΔT_m, as well as $\Delta\Delta G$ (Table 2). Both the ΔT_m and the $\Delta\Delta G$ for conjugate **5** bound to **19** are 6 °C and 6 kcal/mol, respectively, greater than conjugate **5** bound to **13**, supporting that the substitution of G/C for A/T at position 8 contributes to the binding affinity of the polyamide to its target dsDNA.

Ht-polyamide **6** was found to be selective for dsDNA **16**. Contribution of the imidazole on sequence recognition is shown by both the large ΔT_m and $\Delta\Delta G$ for bound **16**. Again, substitution of a G/C bp with an A/T bp at position 7 (dsDNA **13**) resulted in a decrease in ΔT_m of 8 °C (Fig. 4) and in $\Delta\Delta G$ of about 8 kcal/mol (Table 2).

Ht-polyamide **7** was found to bind highly specifically to the 9-bp binding site of **18**, as demonstrated by the ΔT_m of 20 °C and the $\Delta\Delta G$ of 12 kcal/mol (Table 2). The energy minimized complex of **7** bound to dsDNA **18** is shown in Figure 5. The ligand was hand placed in the minor groove based on PDB files,[30,31] energy minimization and docking functions were performed by Sybyl,[32] and the molecular graphics were created using

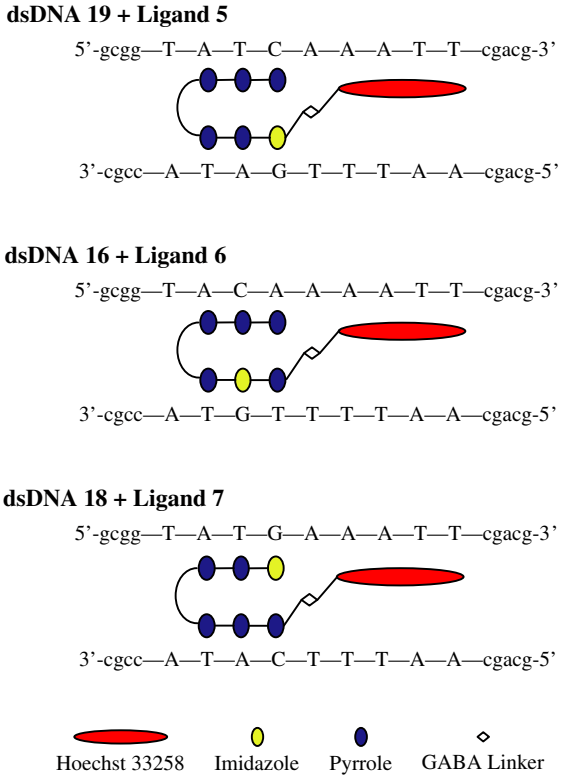


Figure 3. Ligands **5**–**7** bound to their match sequences. The binding regions of the dsDNA sequences are shown capitalized.

Table 2. Free energy calculations[a]

dsDNA	ΔG (kcal/mol)	$\Delta G'$ (kcal/mol)			$\Delta\Delta G$ (kcal/mol)			ΔT_m		
		7	**5**	**6**	**7**	**5**	**6**	**7**	**5**	**6**
13	7	12	16	17	5	9	10	18	19	16
16	7	12	14	**25**	5	7	**18**	16	17	**25**
17	7	17	20	24	10	13	17	15	19	22
18	7	**19**	20	15	**12**	13	8	**20**	23	15
19	7	17	**22**	20	10	**15**	13	19	**25**	18

[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.

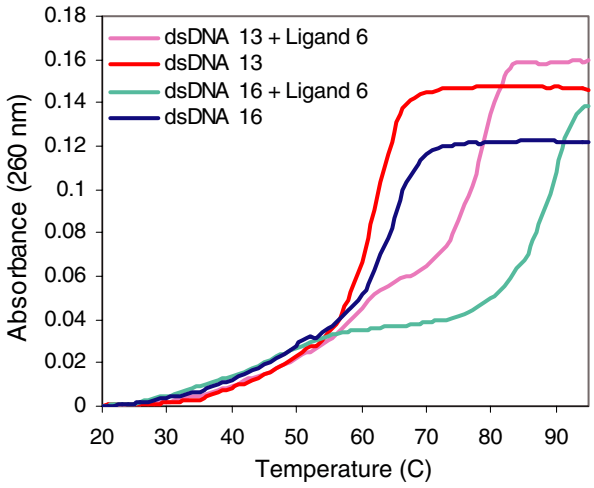


Figure 4. Binding preference for G/C bp over A/T bp.

PyMOL.[33] The strong binding of polyamides **5**–**7** to their respective match sequences relies on the interaction between the Im/Py anti-parallel dimer and the G/C bp.

Previous studies have suggested that the discrimination of the Im/Py anti-parallel dimer for a G/C bp over a C/G bp relies on the strand-specific directionality of the hydrogen bond formed between the exocyclic NH_2 of guanine and the N3 of the imidazole.[22,34] Although the guanine N2 is approximately in the center of the minor groove, the angle between the exocyclic NH_2 of guanine and the N3 of the imidazole is different depending on the location of the Im in respect to the guanine: the angle has been calculated to be ~160° when the Im is on the guanine side, while ~107° when the Im is on the cytosine side.[22]

Indeed, conjugates **5**–**7** each demonstrate a preference for an appropriately positioned G/C bp over a C/G bp. Ht-polyamide **7** does not bind as well to dsDNA **19**, which presents a C/G bp at position 8, but successfully recognized the presence of a properly placed G/C bp in dsDNA **18** (Fig. 6). The binding of ligand **7** has a 2 kcal/mol preference for a G/C bp over a C/G bp. Moreover, a 10-fold better binding of Ht-polyamide **6** with dsDNA **16** was observed compared to that of **17**. Ht-polyamide **5** showed an increase in binding affinity toward the dsDNA **19** over **18** demonstrating the Im G/C preference over a C/G bp (Fig. 6). These findings support the formation of a favorable H-bond between the imidazole and the guanine, indicating a stronger interaction between an Im/Py anti-parallel dimer and a

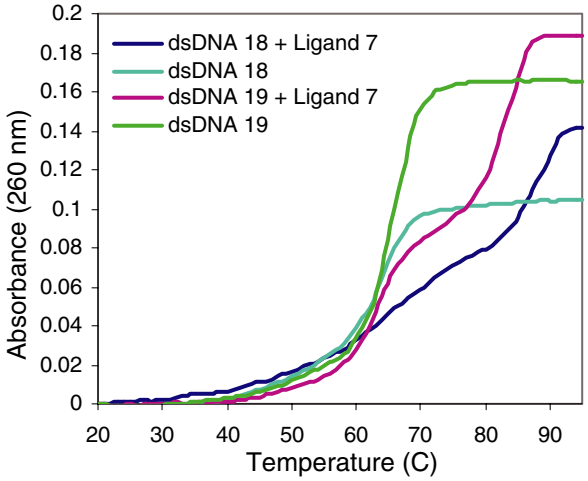


Figure 6. Binding preference for G/C bp over a C/G bp.

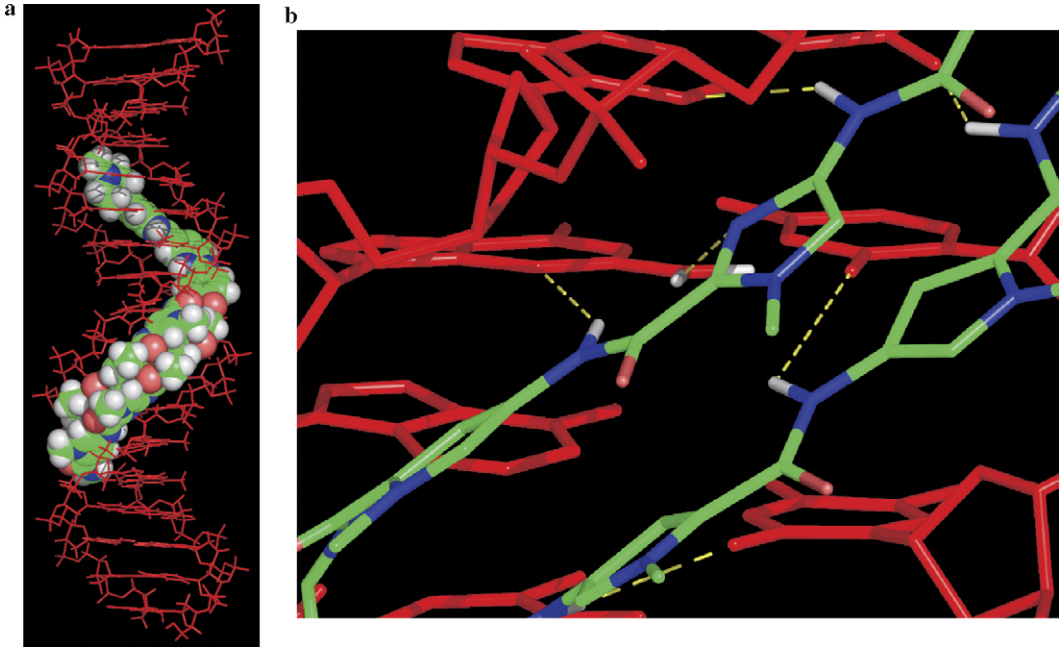


Figure 5. (a) Energy minimized all atom representation of **7** bound to the minor groove of **18**. (b) Heavy atom representation with selected hydrogens shown in white depicting the Im-G interaction present.

G/C bp rather than with a C/G bp, as discussed in previous studies.[22,23,34]

The ΔT_m for sequences**14**–**15** bound to conjugate **5** are all in agreement with the ΔT_m for the same sequences bound to conjugates **6** and **7** showing how the stability of the complex is affected upon substitution of a G/C bp mismatch within the 9-bp A/T-rich binding site. These results confirm a study previously conducted in our laboratory focused on the hairpin pyrrole Ht-polyamide **1**, which was found to be capable of recognizing a 9-bp sequence comprised of only A/T bp.[21]

In an effort to develop novel DNA binding ligands capable of recognizing long sequences with high affinity and improved sequence specificity, Ht-polyamides **5**–**7** were synthesized and evaluated via thermal denaturation experiments. In comparison to the linear Ht-polyamides **2**–**4** bound to dsDNA sequences **16** and **18**,[18] Ht-polyamides **5**–**7** have ΔT_m values that are 14–18 degrees larger, indicating their much improved binding affinities. All three Ht-polyamides were found to be selective for their match dsDNA sequences, supporting a relative stronger interaction of an Im/Py anti-parallel dimer with an appropriately positioned G/C bp rather than sequences containing only A/T bps. Larger increases in ΔT_m and the $\Delta\Delta G$ were observed when conjugates **5**, **6**, and **7** were interacting with dsDNA sequences **19**, **16**, and **18**, respectively, compared to the values observed when each Ht-polyamide interacted with the exclusively A/T bp-containing dsDNA **13**. Ht-polyamides **5** and **7** showed a 20-fold preference for a properly placed G/C bp over a C/G bp, while **6** showed a 10-fold preference. These results suggest that the energetic preference for a stronger Im–G hydrogen bond may be responsible for the observed G/C over C/G selectivity, as previously discussed.[22,23,34] Molecular dynamics experiments are currently underway to examine this hypothesis.

Acknowledgments

This work was supported by a grant from the National Institutes of Health (5R37DK09171). We thank Joseph W. Toporowski for performing the Sybyl molecular modeling calculations.

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. de Clairac, R. P. L.; Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **1997**, *119*, 7909.
9. Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983.
10. Parks, M. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6147.
11. Parks, M. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6153.
12. Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1996**, *3*, 369.
13. Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 8198.
14. 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.
15. 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.
16. Satz, A. L.; Bruice, T. C. *J. Am. Chem. Soc.* **2001**, *123*, 2469.
17. Satz, A. L.; Bruice, T. C. *Bioorg. Med. Chem.* **2002**, *10*, 241.
18. Reddy, P. M.; Jindra, P. T.; Satz, A. L.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 7843.
19. White, C. M.; Satz, A. L.; Bruice, T. C.; Beerman, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10590.
20. White, C. M.; Satz, A. L.; Gawron, L. S.; Bruice, T. C.; Beerman, T. A. *Biochim. Biophys. Acta* **2001**, *1574*, 100.
21. Reddy, P. M.; Dexter, R.; Bruice, T. C. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3803.
22. Kielkopf, C. L.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Nat. Struct. Biol.* **1998**, *5*, 104.
23. Urbach, A. R.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4343.
24. Wurtz, N. R.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Org. Lett.* **2001**, *3*, 1201.
25. Reddy, P. M.; Bruice, T. C. *J. Am. Chem. Soc.* **2004**, *126*, 3736.
26. Fields, G. B.; Noble, R. L. *Int. J. Pept. Protein Res.* **1990**, *35*, 161.
27. Compound **5**: ESI/TOF+ MS *m/z* $(M+2H)^{2+}$ calcd 664.81, obsd 664.83. Compound **6**: ESI/TOF+ MS *m/z* $(M+2H)^{2+}$ calcd 664.81, obsd 664.77. Compound **7**: ESI/TOF+ MS *m/z* $(M+H)^{+}$ calcd 1328.59, obsd 1328.56.
28. 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 150 mM KCl ($\mu = 0.17$). The concentration of each oligomeric duplex was 2 μM and a series of T_m curves were acquired (dsDNA + no ligand; dsDNA + **5** or **6** or **7**) 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.
29. Marky, L. A.; Breslauer, K. J. *Biopolymers* **1987**, *26*, 1601.
30. Mitra, S. N.; Wahl, M. C.; Sundaralingam, M. *Acta. Crystallogr., Sect. D* **1999**, *55*, 602.
31. Vega, M. C.; Saez, I. G.; Aymami, J.; Eritja, R.; van der Marel, G. A.; van Boom, J. H.; Rich, A.; Coll, M. *Eur. J. Biochem.* **1994**, *222*, 721.
32. Sybyl, version 6.8, 2001 Tripos, St. Louis, MO.
33. Delano, W.L. 'The PyMOL Graphics System.' version 0.97, 2004, Delano Scientific LLC, San Carlos, CA, USA. <http://www.pymol.org>.
34. White, S.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1997**, *4*, 569.