Published in final edited form as:
J Am Chem Soc. 2011 May 11; 133(18): 6926–6929. doi:10.1021/ja201397e.

A Bifacial Nucleoside as a Surrogate for both T and A in Duplex DNA

Dongwon Shin and **Yitzhak Tor***
Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093-0358

Abstract

A new functional bifacial nucleoside derived from 7-aminopyrimido[4,5-*d*]pyrimidine-2,4(1*H*,3*H*)-dione, a Janus-type nucleobase, has been synthesized and incorporated into DNA oligonucleotides. The nucleobase, having self-complementary H-bonding faces, mimics both T and A, and engages in the corresponding Watson-Crick-like base pairs, forming stable duplexes.

The native heterocycles found in DNA and RNA and their complementary H-bonding patterns have inspired diverse approaches to the self-assembly of higher structures. In addition to the highly specific Watson Crick pairing between pyrimidines and purines, higher aggregates, in particular G quadruplexes, have stimulated the design and fabrication of numerous self-organizing assemblies.[1–4] Of unique significance among such biomimetic approaches are Janus type heterocycles, which possess self-complementary H-bonding faces.[5,6] Such molecules are capable of self-encoded assembly into higher architectures, in most cases without the need for exogenous ligands or metal ions.[7,8] Intriguingly, this concept, motivated by nucleic acids structure and replication, has rarely been implemented in devising novel nucleosides, nucleotides and oligonucleotides.[9] Tautomerically stable dual-faced nucleosides, capable of self-pairing as well as stable and predictable pairing with other nucleobases, would provide insight into the fundamental forces governing duplex assembly and stability as well as facilitate biophysical approaches to SNP and DNA adduct detection.[10]

When considering Janus-type nucleosides and nucleotides, the heterocycle's glycosylation position and its distance to the H-bonding face have to be optimal, to facilitate seamless integration into double stranded oligonucleosides. Ideally, a single glycosylation position should serve as an anchor point, presenting both a pyrimidine face and a purine face upon simple rotation from the *syn* to the *anti* conformation. Here we disclose the design and synthesis of a new functional bifacial nucleoside derived from 7-aminopyrimido[4,5-*d*]pyrimidine-2,4(1*H*,3*H*)-dione (**BF**), as well as its incorporation into oligonucleotides and the biophysical characterization of the modified duplexes. Nucleobase **BF** possesses self-complementary H-bonding faces (Figure 1), designed after thymine (red) and 2-aminopurine (blue).[11] As such, it can mimic both T and A and can, in principle, engage in the corresponding Watson-Crick-like base pairs by simply rotating around its glycosidic bond (Figure 1). An extensive thermodynamic evaluation of modified oligonucleotides supports the bifacial behavior of this Janus-type nucleoside in diverse sequence contexts.

ytor@ucsd.edu.

Supporting Information Available: Synthetic details, thermal denaturation measurements, van't Hoff analyses, X-ray crystallography, MALDI-TOF MS spectra and CD spectra. This information is available free of charge via the Internet at http://pubs.acs.org.

The synthesis of the bifacial nucleoside and the necessary building block for oligonucleotide synthesis commenced with the treatment of 6-aminouracil (**1**) with phosphorus oxychloride in DMF under the Vilsmeier reaction conditions (Scheme 1).[12] Copper (I) iodide catalyzed glycosylation of **2**, followed by deprotection, provided the desired regioisomer **3** as a mixture of the two anomeric nucleosides (**3α** + **3β**). Protection of the exocyclic amine to give the corresponding dimethylformamidine, was followed by separation of the β- and α-anomers of **4**. Tritylation of the 5′-hydroxyl group in **4β** and phosphitylation of the 3′-hydroxyl group provided phosphoramdite **5**, suitable for automated oligonucleotide synthesis.[12]

The stereochemical assignment of the epimeric **4α** and **4β** was supported by NOE experiments and cemented by a crystal structure of **3α**, obtained after removal of the formamidine protecting group (Figure 2).[13] Its crystal packing shows, in addition to the nucleoside's absolute configuration, the intermolecular ribbon-like network of H bonds. As expected of a Janus heterocycle, two complementary hydrogen bonds are formed between the thymine-like face and the neighboring 2-aminopurine-like face. Importantly, the N–N and the N–O distances observed (2.820 and 2.992 Å, respectively, Figure 2) are essentially identical to the corresponding distances seen for Watson–Crick A:T pairs in B-DNA oligonucleotides (2.82–2.84 and 2.95–2.97 Å, respectively).[14] This suggests that the bifacial nucleoside can indeed engage both its faces in complementary hydrogen bonding, serving as a surrogate for both A and T.

Two complementary non-palindromic 17-mer oligonucleotides **6** and **11**, placing the bifacial nucleoside at a central position in two different sequence contexts (surrounded by either pyrimidines or purines), were synthesized (Figure 3).[12] Each was hybridized to four oligonucleotides (**6** to **12**–**15** and **11** to **7**–**10**, respectively), generating eight distinct duplexes with BF paired to all possible nucleobases. For comparison, the perfectly matched native oligonucleotide duplexes having an A:T or T:A base pairs in the same central position (**7**·**15** and **10**·**12**, respectively) were also examined. Thermal denaturation experiments in two different ionic strengths (Figure 4 and Table S1) and van't Hoff analyses of the melting curves (Table 1) provided insight into the behavior of the new nucleoside within duplex DNA.[15]

Focusing on the central variable position, the following T_m hierarchy, in order of decreasing thermal stability, was observed for **6** in 100 mM of NaCl: BF:T > BF:A > BF:BF ≥ BF:C > BF:G. Importantly, duplexes containing BF:T and BF:A (**6**·**15** and **6**·**12**, respectively) exhibit high stability, comparable to native unmodified duplexes (Table 1). Furthermore, these duplexes are significantly more stable than the corresponding mismatched duplexes, containing BF:C and BF:G (**6**·**13** and **6**·**14**, respectively), supporting the proposed ability of BF to pair with A and T. The slightly lower stability observed for a BF:T pair compared to a native duplex containing an A:T pair in the central position (ΔT_m = −1.3 °C) is not unexpected, as similar destabilization had previously been observed for 2AP:T when compared to A:T.[16] Interestingly, BF:T (**6**·**15**, Table 1) appears to provides stronger pairing than BF:A (**6**·**12**, Table 1), suggesting that expulsion of the 2AP face of the bifacial nucleoside into the major groove is energetically more costly (see below). The identity of the pairing face in the mismatched duplexes is harder to decipher. The reasonable stability of the BF:C containing duplex compared to those containing A:C and T:C (ΔT_m = +2.8 and 4.1 °C, respectively) suggests, however, that the 2AP face of BF may be H bonded, forming a wobble pair, similarly to observations made with 2AP.[17]

Thermodynamic parameters, derived from van't Hoff analyses of the thermal denaturation curves, yielded the following order of decreasing duplex stability at 37 °C: A:T > BF:T ≥ T:A > BF:A > BF:C > BF:BF > BF:G (Table 1).[15] This trend highlights the ability of BF to

stably pair with both T and A. In fact, the difference in free energy (ΔΔG°) between A:T and BF:T (**7·15** and **6·15**, respectively) is only 0.8 kcal/mol, which may be accounted for by the difference observed for A:T vs. 2AP:T base pairs (ΔΔG° = 1.1 kcal/mol).[18] As the same order to duplex stability is obtained by comparing the extracted ΔH° values, we conclude that the major contribution to base pair stability of BF-containing duplexes is enthalpic. We further note that the range of ΔH° values obtained for duplexes **6·11**–**6·15** (Table 1), with the possible exclusion of the BF:G-containing duplex **6·14**, suggest a similar number of H bonding interactions.[19]

To examine the context dependency of the modification with BF, oligonucleotide **11** was hybridized to the corresponding complementary and mismatch sequences (**7**–**10**). For this "reverse" sequence, where BF is embedded between two dC instead of dG residues, thermal denaturation measurements yielded the following order of decreasing stability: T:A > A:T > T:BF > A:BF > BF:BF > G:BF > C:BF (Table 1). With the exception of C:BF, this trend matches the order seen in the set of modified duplexes discussed above.[20] In general, however, lower T_m values are observed when the bifacial nucleoside **BF** is embedded between two pyrimidines (dC) when compared to two purines (dG). Better base stacking within the G-BF-G stretch (**6·12**–**6·15**, Table 1) likely results in more stable duplexes compared to the corresponding C-BF-C duplexes (**7·11**–**10·11**, Table 1). These observations are consistent with reports showing inferior stacking interactions in C-2AP-C compared to G-2AP-G, and with calculated thermal stabilities of stacked DNA doublets, which predict that C-2AP is significantly less stable than G-2AP when paired with G-T and C-T, respectively.[17b,21]

Taken together, we conclude that the designed bifacial nucleoside **3** performs as anticipated, generating stable base pairs with both T and A within duplex DNA. While pairing BF with T appears to be slightly favored over pairing with A, both these matched pairs are more stable than the corresponding mismatches with G and C. Intriguingly, a duplex containing a BF:BF pair, although more stable that mismatched duplexes, shows lower stability than "perfect" duplexes containing BF:A or BF:T pairs. This is consistent with the stability differences observed between the BF:A and BF:T pairs, as in a BF:BF pair one of the nucleobases inevitably projects the 2AP face into the major groove. It remains to be seen whether or not homopolymeric duplexes containing nucleoside **3** would generate alternate or random arrangements of the nucleobase BF.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the National Institutes of Health for their generous support (GM 069773), Andro C. Rios for his input, and Mary Noé for her assistance with MALDI experiments.

References

1. (a) Whitesides GM, Simanek EE, Mathias JP, Seto CT, Chin DN, Mammen M, Gordon DM. Acc Chem Res. 1995; 28:37–44.(b) Lawrence DS, Jiang T, Levett M. Chem Rev. 1995; 95:2229–2260. (c) Prins LJ, Reinhoudt DN, Timmerman P. Angew Chem, Int Ed. 2001; 40:2383–2426.
2. (a) Bonazzi S, Demorais MM, Gottarelli G, Mariani P, Spada GP. Angew Chem, Int Ed Engl. 1993; 32:248–250.(b) Ciuchi F, Dinicola G, Franz H, Gottarelli G, Mariani P, Bossi MGP, Spada GP. J Am Chem Soc. 1994; 116:7064–7071.(c) Gottarelli G, Mezzina E, Spada GP, Carsughi F, DiNicola G, Mariani P, Sabatucci A, Bonazzi S. Helv Chim Acta. 1996; 79:220–234.(d) Chaput JC, Switzer C. P Natl Acad Sci USA. 1999; 96:10614–10619.(e) Sakai N, Kamikawa Y, Nishii M, Matsuoka T,

Kato T, Matile S. J Am Chem Soc. 2006; 128:2218–2219. [PubMed: 16478168] (f) Davis JT, Spada GP. Chem Soc Rev. 2007; 36:296–313. [PubMed: 17264931] (g) Gonzalez-Rodriguez D, Janssen PGA, Martin-Rapun R, De Cat I, De Feyter S, Schenning APHJ, Meijer EW. J Am Chem Soc. 2010; 132:4710–4719. [PubMed: 20039610] (h) Linder J, Garner TP, Williams HEL, Searle MS, Moody CJ. J Am Chem Soc. 2011; 133:1044–1051. [PubMed: 21162526]

3. (a) Burrows AD. Supramolecular Assembly Via Hydrogen Bonds I. 2004; 108:55–95.(b) Reinhoudt DN, Crego-Calama M. Science. 2002; 295:2403–2407. [PubMed: 11923525] (c) Mathias JP, Simanek EE, Zerkowski JA, Seto CT, Whitesides GM. J Am Chem Soc. 1994; 116:4316–4325.
4. (a) Fournier JH, Maris T, Wuest JD, Guo WZ, Galoppini E. J Am Chem Soc. 2003; 125:1002–1006. [PubMed: 12537499] (b) Laliberte D, Maris T, Wuest JD. J Org Chem. 2004; 69:1776–1787. [PubMed: 15058918] (c) Hosseini MW. Acc Chem Res. 2005; 38:313–323. [PubMed: 15835878]
5. (a) Marsh A, Nolen EG, Gardinier KM, Lehn JM. Tetrahedron Lett. 1994; 35:397–400.(b) Branda N, Kurz G, Lehn JM. Chem Commun. 1996:2443–2444.
6. (a) Corbin PS, Lawless LJ, Li ZT, Ma YG, Witmer MJ, Zimmerman SC. P Natl Acad Sci USA. 2002; 99:5099–5104.(b) Mayer MF, Nakashima S, Zimmerman SC. Org Lett. 2005; 7:3005–3008. [PubMed: 15987191]
7. (a) Asadi A, Patrick BO, Perrin DM. J Org Chem. 2007; 72:466–475. [PubMed: 17221963] (b) Asadi A, Patrick BO, Perrin DM. J Am Chem Soc. 2008; 130:12860–12861. [PubMed: 18767852]
8. (a) Mascal M, Hext NM, Warmuth R, Arnall-Culliford JR, Moore MH, Turkenburg JP. J Org Chem. 1999; 64:8479–8484.(b) Yang Y, Xue M, Xiang JF, Chen CF. J Am Chem Soc. 2009; 131:12657–12663. [PubMed: 19685881] (c) Zhao JW, Veliath E, Kim S, Gaffney BL, Jones RA. Nucleos Nucleot Nucl. 2009; 28:352–378.(d) Arrachart G, Carcel C, Trens P, Moreau JJE, Man MWC. Chem-Eur J. 2009; 15:6279–6288.(e) Borzsonyi G, Johnson RS, Myles AJ, Cho JY, Yamazaki T, Beingessner RL, Kovalenko A, Fenniri H. Chem Commun. 2010; 46:6527–6529.
9. (a) Hirano T, Kuroda K, Kodama H, Kataoka M, Hayakawa Y. Lett Org Chem. 2007; 4:530–534.(b) Hirano T, Kuroda K, Kataoka M, Hayakawa Y. Org Biomol Chem. 2009; 7:2905–2911. [PubMed: 19582300]
10. Gong JC, Sturla SJ. J Am Chem Soc. 2007; 129:4882–4883. [PubMed: 17402738]
11. To maximize “decoupling” of the two H-bonding faces, we selected to use a U and 2AP substructures, instead of U and A. In the latter heterocycle {(5-aminopyrimidino[4,5-*d*]pyrimidine-2,4(1*H*,3*H*)-dione}, the carbonyl at the 4 position would likely be interacting with the exocyclic NH_2 group.
12. See SI for additional detailed experiments and analytical data.
13. See SI for NOE experiments and X-ray crystal structures.
14. Tereshko V, Minasov G, Egli M. J Am Chem Soc. 1999; 121:6970–6970.(b) Kielkopf CL, Ding S, Kuhn P, Rees DC. J Mol Biol. 2000; 296:787–801. [PubMed: 10677281] (c) Wang F, DeMuro NE, Elmquist CE, Stover JS, Rizzo CJ, Stone MP. J Am Chem Soc. 2006; 128:10085–10095. [PubMed: 16881637] (d) Locasale JW, Napoli AA, Chen SF, Berman HM, Lawson CL. J Mol Biol. 2009; 386:1054–1065. [PubMed: 19244617]
15. Due to the minimal differences between the duplexes studied, we used a two-state approximation to extract thermodynamic parameters from the optical melting curves (and assumed no difference in heat capacity for these states).[12] For potential disparities and agreement between optical melting curves and calorimetry, see: (a) Albergo DD, Marky LA, Breslauer KJ, Turner DH. Biochemistry. 1981; 20:1409–1413. [PubMed: 6261793] (b) Petersheim M, Turner DH. Biochemistry. 1983; 22:256–263. [PubMed: 6824629] (c) Erie D, Sinha N, Olson W, Jones R, Breslauer K. Biochemistry. 1987; 26:7150–7159. [PubMed: 3427065]
16. (a) Eritja R, Kaplan BE, Mhaskar D, Sowers LC, Petruska J, Goodman MF. Nucleic Acids Res. 1986; 14:5869–5884. [PubMed: 3737416] (b) Mclaughlin LW, Leong T, Benseler F, Piel N. Nucleic Acids Res. 1988; 16:5631–5644. [PubMed: 2838824] (c) Law SM, Eritja R, Goodman MF, Breslauer KJ. Biochemistry. 1996; 35:12329–12337. [PubMed: 8823167]
17. (a) Sowers LC, Fazakerley GV, Eritja R, Kaplan BE, Goodman MF. P Natl Acad Sci, USA. 1986; 83:5434–5438.(b) Fagan PA, Fabrega C, Eritja R, Goodman MF, Wemmer DE. Biochemistry. 1996; 35:4026–4033. [PubMed: 8672436]

18. (a) Clayton LK, Goodman MF, Branscomb EW, Galas DJ. J Biol Chem. 1979; 254:1902–1912. [PubMed: 422561] (b) Galas DJ, Branscomb EW. J Mol Biol. 1978; 124:653–687. [PubMed: 712851]
19. CD spectra (Figure S5) suggest an unperturbed B-form duplex for the modified oligonucleotides.
20. The large decrease in a duplex stability observed for a C:BF mismatch (**8**·**11**, Table 1), is associated with significantly smaller −ΔDH° (70 kcal/mol), compared to other pairs (> 100 kcal/mol). Together with the substantially lower −ΔS° value, it is likely that the C:BF mismatch is unstructured.
21. Petruska J, Goodman MF. J Biol Chem. 1985; 260:7533–7539. [PubMed: 3158658]

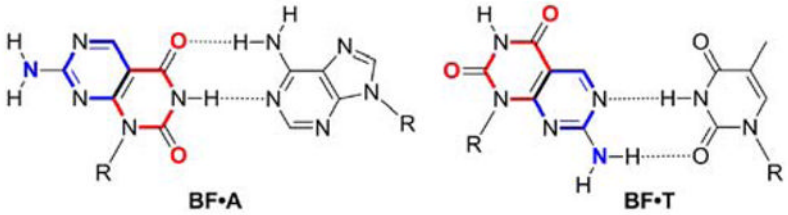


Figure 1.
A BiFacial nucleoside (BF) and the pairing of its uracil face (red) with A and its 2-aminopurine face (blue) with T.

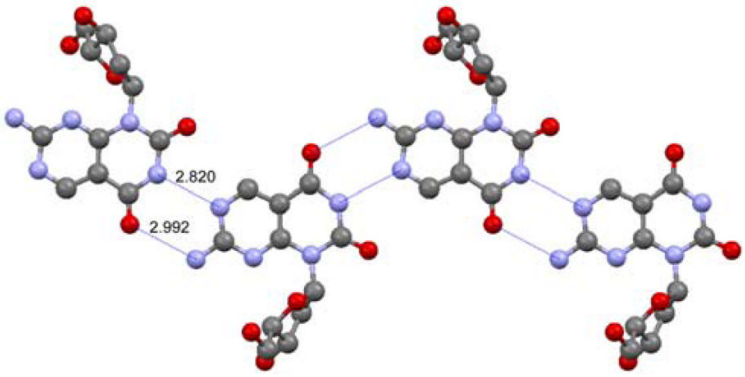


Figure 2.
X-Ray crystal structure of **3α** highlighting the ribbon-like network of H-bonds and the pairing of the complementary faces.

6	5'-GAG CGA TGBF GTA GCG AG-3'
7	5'-GAG CGA TGA GTA GCG AG-3'
8	5'-GAG CGA TGC GTA GCG AG-3'
9	5'-GAG CGA TGG GTA GCG AG-3'
10	5'-GAG CGA TGT GTA GCG AG-3'
11	3'-CTC GCT ACBF CAT CGC TC-5'
12	3'-CTC GCT ACA CAT CGC TC-5'
13	3'-CTC GCT ACC CAT CGC TC-5'
14	3'-CTC GCT ACG CAT CGC TC-5'
15	3'-CTC GCT ACT CAT CGC TC-5'

Figure 3.
Oligonucleotide sequences containing **3β** (BF) and their complements.

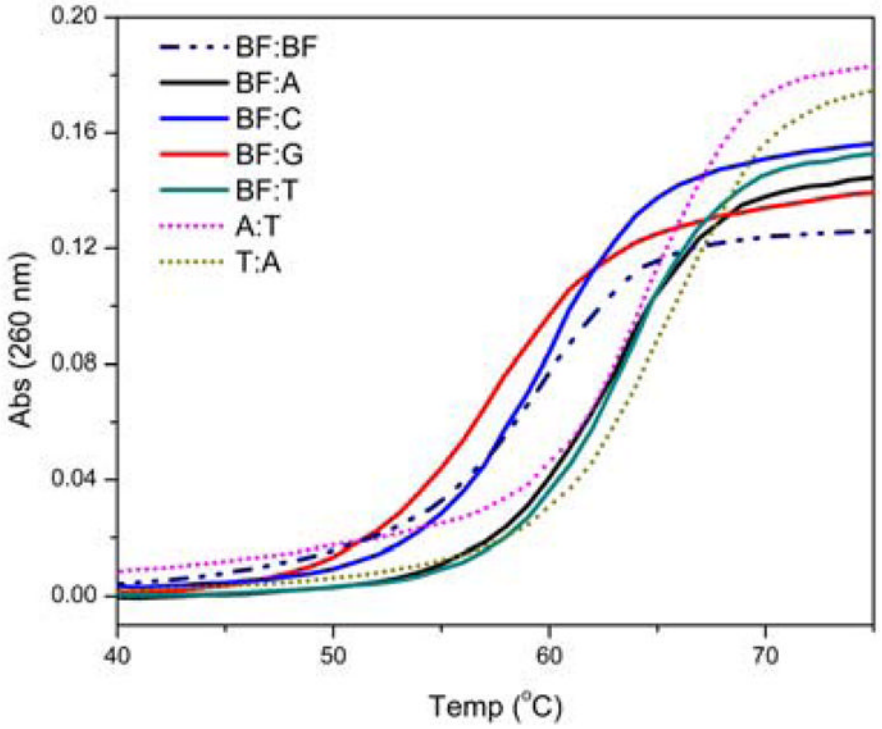
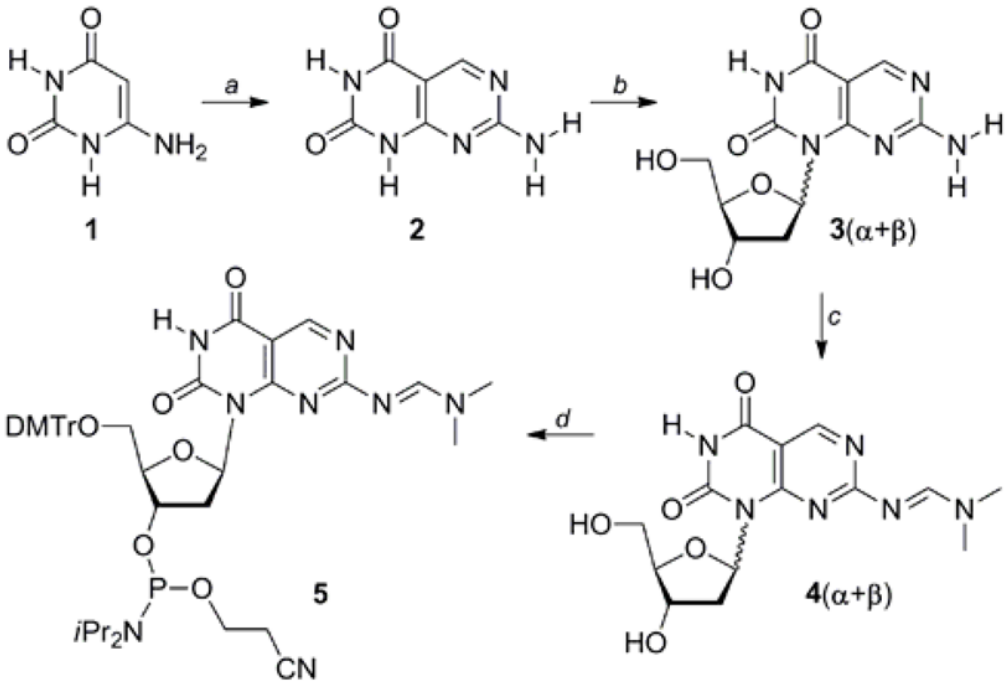


Figure 4.
Selected denaturation profiles for duplexes containing oligonucleotide **6** determined by correlating absorbance at 260 nm vs. temperature. Conditions as given in Table 1.



Scheme 1.

Synthesis of heterocycle **2**, ribonucleoside **3**, and the corresponding Phosphoramidite **5.**[a]

[a] *Reagents and conditions*: (a) (i) $POCl_3$, DMF, 20 °C, 74 %; (ii) guanidine·HCl, NaOEt, EtOH, reflux, (iii) diglyme, 160 °C, 5 d, 93 %; (b) (i) *N*,*O*-bis(trimethylsilyl)acetamide, CH_3CN, RT; (ii) 1-chloro-2-deoxy-3,5-di-*O*-toluoyl-α-D-ribofuranose, CuI, 54 %; (iii) NH_3, MeOH, 80 °C, 89 %; (d) *N*,*N*-dimethylformamide dimethyl acetal, pyridine, 54 % (**4β**), 43 % (**4α**); (d) (i) DMTrCl, Py, 61 %; (ii) *i*Pr_2NEt, (*i*Pr_2N)P(Cl)OCH_2CH_2CN, CH_2Cl_2, 30 %.

Table 1

Thermodynamic data for duplexes.[a]

5′-GAG CGA TGX GTA GCG AG-3′ 3′-CTC GCT ACY CAT CGC TC-5′					
No.	**X:Y**	**Tm (°C)**[b]	**$-\Delta G°_{37}$**[c, d]	**$-\Delta H°$**[d]	**$-\Delta S°$**[e]
6·11	BF:BF	59.4 (± 0.2)	15.8	110.0	303.0
6·12	BF:A	62.2 (± 0.3)	18.0	128.0	355.0
6·13	BF:C	59.3 (± 0.1)	16.9	127.0	355.0
6·14	BF:G	56.5 (± 0.1)	14.9	109.1	304.0
6·15	BF:T	63.2 (± 0.3)	19.1	137.0	380.0
7·11	A:BF	60.4 (± 0.7)	15.9	106.9	294.0
8·11	C:BF	54.0 (± 0.2)	12.0	69.9	187.0
9·11	G:BF	56.4 (± 0.2)	14.5	103.3	286.0
10·11	T:BF	62.1 (± 0.2)	16.7	111.0	304.0
7·15	A:T	64.5 (± 0.2)	19.9	141.9	393.0
10·12	T:A	65.0 (± 0.0)	19.0	128.1	352.0
8·12	C:A	52.6 (± 0.2)	12.6	87.5	241.0
9·12	G:A	61.5 (± 0.1)	17.4	123.2	341.0
8·15	C:T	50.0 (± 0.0)	11.5	76.9	211.0
9·15	G:T	59.2 (± 0.1)	16.3	118.8	330.0
7·13	A:C	56.5 (± 0.2)	14.5	105.5	293.0
7·14	A:G	58.3 (± 0.1)	15.5	110.3	306.0
10·13	T:C	55.2 (± 0.1)	13.8	98.4	273.0
10·14	T:G	59.6 (± 0.1)	16.0	112.0	310.0

[a] All samples contained 2.5 μM each strand of DNA, 20 mM Na cacodylate, pH 7.0, 0.5 mM EDTA and 100 mM NaCl;

[b] errors reflect standard deviation derived from two individual experiments;

[c] ± <0.3;

[d] unit is kcal/mol;

[e] unit is cal/mol.