Towards A DNA-Like Duplex without Hydrogen-Bonded Base Pairs**

Gérald Mathis and Jürg Hunziker*

The creation of alternative base pairs to expand the genetic alphabet has recently been the focus of intensive research.[1–13] Initial work was based on the systematic permutation of hydrogen bond donor and acceptor arrangements within the pyrimidine and purine nuclei.[1–3] In subsequent studies the Watson–Crick complementarity paradigm was abandoned altogether.[4–9] It was shown that aromatic entities devoid of hydrogen bonding functional groups can satisfy at least part of the criteria for replicable base pairs, that is, thermodynamic stability of the resulting duplex, orthogonality to natural nucleobases, and enzymatic processability. Such nonpolar base pair isosteres are held together by hydrophobic dispersion forces which act predominantly along the base stack but less so at the interface between the bases. In an attempt to orient the base–base attraction perpendicular to the helix axis, chelating nucleosides have been developed.[10–13] However, such metal-mediated base pairs can hardly be amenable to enzymatic replication.

In the search for alternative base pair design principles we were particularly intrigued by the interaction between benzene and hexafluorobenzene.[14–18] The two compounds, both liquids at room temperature, form a solid aggregate[19,20] which is characterized not only by π stacks of alternating components but by the lateral alternation between hexafluorobenzene and benzene rings as well (Figure 1). The inverse quadrupolar moments of benzene and hexafluorobenzene lead to edge-to-edge attractive intermolecular forces which could be utilized in designing novel base pairs, such as the one formed between C-nucleosides **1** and **2** (Scheme 1). Here we report on the synthesis and the promising pairing properties of oligonucleotides containing **1** and **2**.

The synthesis of phenyl-2-deoxyriboside **2** and its incorporation into oligonucleotides has been described.[21] The novel pentafluorophenyl congener **1** was prepared from 2-deoxyribonolactone following a literature strategy (Scheme 2).[22] Commercially available bromopentafluorobenzene (**4**) was lithiated and then the silyl-protected 2-deoxyribonolactone **3**[22] was added. The resulting hemiacetal was reduced in situ with Et_3SiH to give an inseparable 9:1 mixture of the two anomers **5** (β) and **6** (α) in low yield. These compounds could be separated by chromatography after deprotection of the deoxyribose moiety and reprotection as 5′-dimethoxytrityl ethers. At this stage the configuration of the two isomers could be assigned by ^{1}H NMR NOE experiments.[23] The pentafluorophenyl-β-D-riboside phosphoramidite **8** was then obtained by standard phosphitylation.[24] Oligonucleotides containing **1** and **2** (Tables 1 and 2) were prepared using a regular solid-phase DNA synthesis protocol.[24,25]

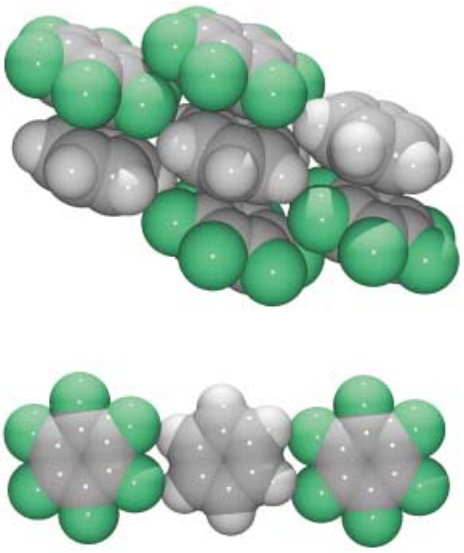
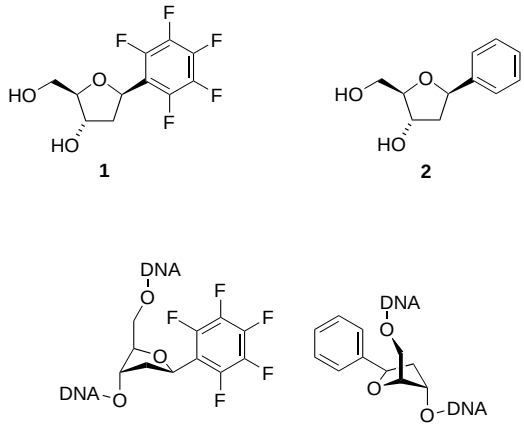


Figure 1. Details of the X-ray crystal structure of the solid 1:1 complex formed from benzene and hexafluorobenzene (grey: carbon, white: hydrogen, green: fluorine). The side view (top) shows the stacks of alternating composition. The aromatic molecules are tilted with respect to the stack axis, which leads to lateral alternation within the plane of individual benzene or hexafluorobenzene molecules (bottom).



Scheme 1. Structures of the pentafluorophenyl-β-D-deoxyriboside (F^5, **1**) and phenyl-β-D-deoxyriboside (P, **2**) used to investigate the stability of F^5–P base pairs in oligonucleotide duplexes.

When incorporated into a 10-mer duplex, a single pentafluorophenyl–phenyl pair (F^5–P) leads to a strong decrease in duplex stability, as judged by the melting temperature T_m (Table 1). The stability of the novel base pair is intermediate to that of the relatively well-tolerated mismatched pairs A–G and G–G as well as the wobble pair G–T and all other more deleterious mismatches between natural bases. The interaction between the phenyl and the pentafluorophenyl units is stronger than those with any of the natural bases. Unfortunately, there is no or only a slight discrimination between F^5–P and the corresponding self-pairs P–P and F^5–F^5 in this sequence context.

Moving to a self-complementary 10-mer duplex with several F^5–P base pairs dramatically changes the picture.

[*] Priv.-Doz. Dr. J. Hunziker, Dipl.-Chem. G. Mathis
Department of Chemistry and Biochemistry
University of Bern
Freiestrasse 3, 3012 Bern (Switzerland)
Fax: (+41) 31-631-8057
E-mail: juerg.hunziker@ioc.unibe.ch

[**] This work was supported by the Swiss National Science Foundation (grant no. 20-61730.00).

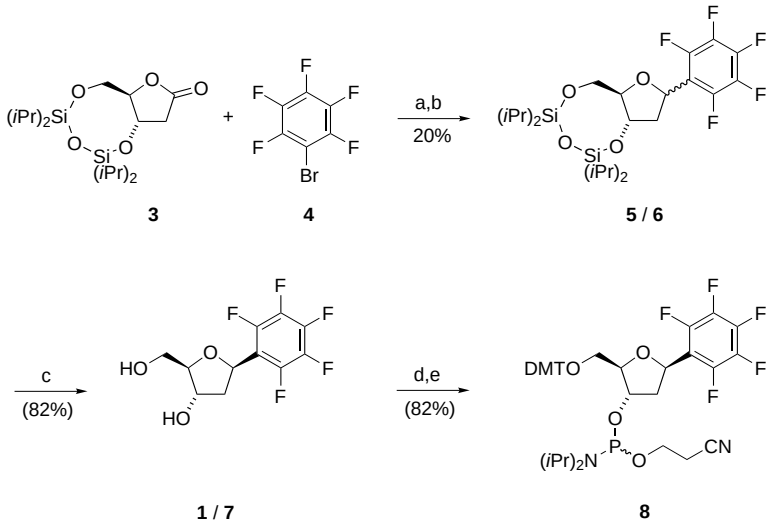
  0044-8249/02/4117-3203 $ 20.00+.50/0

Table 1. The T_m values [°C] determined from UV melting experiments for the 10-mer DNA duplex **9** with varying bases in positions **X** and **Y**.[a]

5′-CTGA**X**TCGAC-3′
3′-GACT**Y**AGCTG-5′ **9**

		d(CTGA**X**TCGAC)					
		A	T	G	C	F^5	P
d(GTCGA**Y**TCAG)	A	29.0	**44.3**	36.2	28.1	25.0	25.1
	T	**45.0**	27.1	35.5	27.1	22.0	21.1
	G	35.6	36.1	38.9	**49.0**	25.0	23.0
	C	29.8	28.0	**50.0**	25.1	24.1	24.1
	F^5	25.0	22.0	28.7	21.1	29.9	***30.6***
	P	25.1	21.0	27.2	17.6	***30.0***	30.0

[a] Duplex concentration: 4 μM; buffer: 10 mM NaH_2PO_4, 150 mM NaCl, pH 7.0; detection at 260 nm, T_m = mean value of three melting curves, heating rate = 0.5 °C min^{-1}.



Scheme 2. a) Compound **4** (2.0 equiv), BuLi (1.8 equiv), Et_2O, −78 °C, 1 h, then **3** (1.0 equiv) in Et_2O, −78 °C, 2.5 h; b) Et_3SiH, BF_3·Et_2O, CH_2Cl_2, −78 °C, 12 h; c) Bu_4NF, THF, RT, 2 h; d) DMT-Cl, *N*,*N*-dimethylaminopyridine, pyridine, RT, 3 h; e) [(*i*Pr$_2$N)(NCCH$_2$CH$_2$O)P]Cl, *i*Pr$_2$NEt, THF, RT, 1 h. DMT = 4,4′-dimethoxytrityl.

With two such pairs in a central position (duplex $\mathbf{12}_2$) there is still a decrease in thermal stability (ΔT_m/mod. −4.0 K compared to A–T, Table 2), however, to a considerably lower degree than in the case of a single substitution (ΔT_m/mod. −15 K). Introducing four consecutive F^5–P pairs ($\mathbf{13}_2$) leads to a duplex which is more stable than its natural counterpart with four A–T pairs (ΔT_m/mod. +1.5 K). The beneficial mutual interaction is unique to the phenyl and pentafluorophenyl residues in this system. The duplexes $\mathbf{15}_2$ and $\mathbf{16}_2$ with four F^5–G and F^5–A pairs, respectively, are destabilized in comparison.

The differences in duplex stability of sequences with one, two, or four consecutive F^5–P base pairs may reflect the altered geometry of this artificial base pair. The contact between phenyl and pentafluorophenyl in an optimal orientation likely requires a backbone distance shorter than the 11 Å of a standard pyrimidine–purine base pair. The backbone may only adapt to this if several F^5–P pairs are arranged in a row.[26]

Measuring the thermal stability of 10-mer duplexes containing more than four F^5–P pairs meets technical hurdles. The UV absorption of C-nucleosides **1** and **2** is weak over the wavelength range of the melting experiments. In essence, only the destacking of natural bases is detected, as can be seen from the drop in hyperchromicity in duplexes $\mathbf{12}_2$ (14.5 %, Table 2), $\mathbf{13}_2$ (8 %), $\mathbf{18}_2$ (7.5 %), and $\mathbf{21}_2$ (−1 %). In addition, duplex $\mathbf{18}_2$ with six alternating F^5–P base pairs does not show a sigmoidal curve indicative of a cooperative transition.

Table 2. The T_m values [°C] determined from UV melting experiments for the 10-mer oligonucleotides **10**–**21** with several F^5–P pairs.[a]

Oligonucleotide	T_m [°C]	ΔT_m/mod. [K]	Hyperchromicity
d(CTGATATCAG) (**10**)	38.0	–	27 %
d(CTGAGCTCAG) (**11**)	48.5	+5.3	23 %
d(CTGA**F^5**PTCAG) (**12**)	30.0[b]	−4.0	14.5 %
d(CTG**F^{5}PF^{5}P**CAG) (**13**)	44.0	+1.5	8 %
d(CTG**F^{5}F^{5}PP**CAG) (**14**)	44.7[b]	+1.7	8 %
d(CTG**F^5**G**F^5**GCAG) (**15**)	35.8	−1.1	6 %
d(CTG**F^5**A**F^5**ACAG) (**16**)	37.7	−0.2	6.5 %
d(C**F^5**G**P**AT**F^{5}C P**G) (**17**)	–[c]	–	2.5 %
d(CT**PF^{5}PF^{5}PF^5**AG) (**18**)	–[d]	–	7.5 %
d(CT**F^{5}F^{5}F^{5}PPP**AG) (**19**)	–[d]	–	5 %
d(C**F^{5}PF^5**GC**PF^{5}P**G) (**20**)	23.1[b]	−2.5	9 %
d(C**F^{5}PF^{5}PF^{5}PF^{5}P**G) (**21**)	–[c]	–	−1 %

[a] Duplex concentration: 4 μM; buffer: 10 mM NaH_2PO_4, 150 mM NaCl, pH 7.0; detection at 260 nm, T_m = mean value of three melting curves, heating rate = 0.5 K min^{-1}. [b] Monomolecular hairpin at 4 μM. [c] No melting transition observed. [d] No sigmoidal (cooperative) melting transition.

The strength and the directionality of the interaction between the phenyl and pentafluorophenyl residues is revealed in concentration variation experiments, which also yield the thermodynamic parameters of duplex formation (Table 3). A single F^5–P base pair in the duplex d(CTGA**P**TCGAC)·d(GTCGA**F^5**TCAG) (**24**·**25**) leads to a decrease in the enthalpy of duplex formation compared to the unmodified reference, which cannot be compensated by a more favorable pairing entropy. The situation is different in the self-complementary duplex d(CTG**F^{5}PF^{5}P**CAG)$_2$ ($\mathbf{13}$)$_2$. The pairing enthalpy in this system is almost 40 % higher than for the reference d(CTGATATCAG)$_2$ ($\mathbf{10}$)$_2$, which can most likely be attributed to the extraordinarily strong stacking interactions between alternating phenyl and pentafluorophenyl bases. Support for this notion stems from the following observations. The oligonucleotide d(CTG**F^{5}PF^{5}P**CAG) (**13**) forms a duplex at concentrations above 3 μM; below this threshold a monomolecular structure—presumably a hairpin—is predominant

  0044-8249/02/4117-3204 $ 20.00+.50/0 *Angew. Chem. Int. Ed.* **2002**, *41*, No. 17

Table 3. Thermodynamic data for 10-mer DNA duplexes containing F^5–P base pairs and the corresponding references.[a]

Oligonucleotide	$\Delta H°$ [kcal mol^{-1}]	$\Delta S°$ [cal mol^{-1} K]	$\Delta G°$ (25 °C) [kcal mol^{-1}]
d(CTGAATCGAC)·d(GTCGATTCAG) (**22**·**23**)	−77.6	−213	−13.8
d(CTGA**P**TCGAC)·d(GTCGA**F**^{5}TCAG) (**24**·**25**)	−61.5	−175	−9.1
d(CTGATATCAG)$_2$ (**10**)$_2$	−67.5	−192	−10.1
d(CTGAGCTCAG)$_2$ (**11**)$_2$	−82.3	−231	−13.3
d(CTG**F**^{5}**PF**^{5}**P**CAG)$_2$ (**13**)$_2$	−94.3	−273	−13.0

[a] Duplex concentration: 1–32 μM; buffer: 10 mM NaH_2PO_4, 150 mM NaCl, pH 7.0.

(Figure 2). In contrast, the sequences d(CTG**F**^{5}**F**^{5}**PP**CAG) (**14**) as well as d(CTGA**F**^{5}**P**TCAG) (**12**) prefer a monomolecular structure with constant T_m up to much higher concentrations (> 10 μM). In the case of **13** the central portion of the oligonucleotide seems to be rigidified by the alternating F^5-P-F^5-P stack, which is absent or smaller in **12** and **14**.[27] The intramolecular stacking interaction between the phenyl and pentafluorophenyl residues is therefore more stabilizing than the lateral, intermolecular one. This is not surprising given the different sizes of the contacting surfaces in the two dimensions.

In conclusion, complementary charge distribution as in the pentafluorophenyl–phenyl C-nucleosides **1** and **2** represents a novel design principle for artificial base pairs. The results from this study highlight the importance of favorable intrastrand stacking interactions in the thermodynamic stabilization of oligonucleotide duplexes. On the other hand, interstrand stacking has recently been exploited in a base pair formed between two bipyridine residues.[13] A combination of these two features could lead to orthogonal, non-hydrogen bonded, non-shape complementary base pairs. Experiments towards this end as well as attempts to replicate the pentafluorophenyl–phenyl pair by polymerases are currently under way.

Received: March 18, 2002 [Z18923]

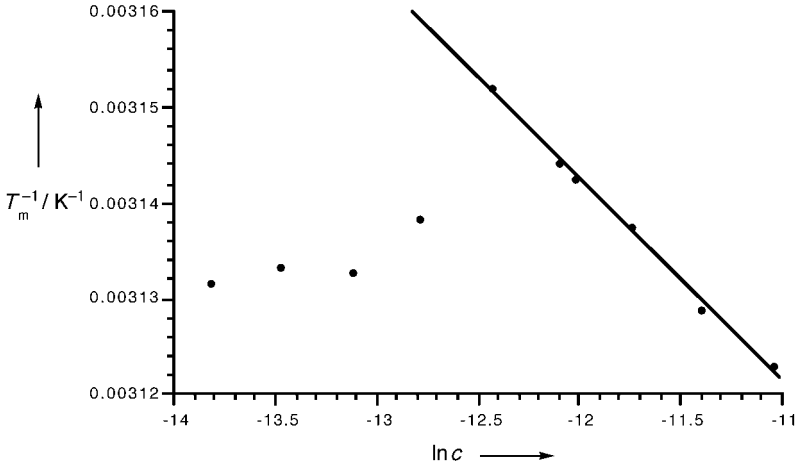


Figure 2. Van't Hoff plot of $1/T_m$ vs. ln(c) (c in mol L^{-1}) for oligonucleotide d(CTG**F**^{5}**PF**^{5}**P**CAG) (**13**). Sequence **13** with four alternating F^5–P base pairs in the middle forms duplexes at concentrations above 3 μM (sloped region, linear regression indicated). In contrast, sequences d(CTGA**F**^{5}**P**TCAG) (**12**) and d(CTG**F**^{5}**F**^{5}**PP**CAG) (**14**) form a duplex only at much higher concentrations.

[1] J. A. Piccirilli, T. Krauch, S. E. Moroney, S. A. Benner, *Nature* **1990**, *343*, 33 – 37.
[2] J. D. Bain, C. Switzer, A. R. Chamberlin, S. A. Benner, *Nature* **1992**, *356*, 537 – 539.
[3] J. J. Voegel, S. A. Benner, *Helv. Chim. Acta* **1996**, *79*, 1863 – 1880; J. J. Voegel, S. A. Benner, *Helv. Chim. Acta* **1996**, *79*, 1881 – 1898.
[4] T. J. Matray, E. T. Kool, *J. Am. Chem. Soc.* **1998**, *120*, 6191 – 6192.
[5] R. S. Coleman, M. L. Madaras, *J. Org. Chem.* **1998**, *63*, 5700 – 5703.
[6] K. M. Guckian, B. A. Schweitzer, R. X.-F. Ren, C. J. Sheils, D. C. Tahmassebi, E. T. Kool, *J. Am. Chem. Soc.* **2000**, *122*, 2213 – 2222.
[7] M. Berger, A. K. Ogawa, D. L. McMinn, Y. Wu, P. G. Schultz, F. E. Romesberg, *Angew. Chem.* **2000**, *112*, 3069 – 3071; *Angew. Chem. Int. Ed.* **2000**, *39*, 2940 – 2942.
[8] E. Lee Tae, Y. Wu, G. Xia, P. G. Schultz, F. E. Romesberg, *J. Am. Chem. Soc.* **2001**, *123*, 7439 – 7440.
[9] I. Singh, W. Hecker, A. K. Prasad, V. S. Parmar, O. Seitz, *Chem. Commun.* **2002**, 500 – 501.
[10] M. Shionoya, K. Tanaka, *Bull. Chem. Soc. Jpn.* **2000**, *73*, 1945 – 1954.
[11] E. Meggers, P. L. Holland, W. B. Tolman, F. E. Romesberg, P. G. Schultz, *J. Am. Chem. Soc.* **2000**, *122*, 10714 – 10715.
[12] H. Weizman, Y. Tor, *J. Am. Chem. Soc.* **2001**, *123*, 3375 – 3376.
[13] C. Brotschi, A. Häberli, C. J. Leumann, *Angew. Chem.* **2001**, *113*, 3101 – 3103; *Angew. Chem. Int. Ed.* **2001**, *40*, 3012 – 3014.
[14] F. Cozzi, F. Ponzini, R. Annunziata, M. Cinquini, J. S. Siegel, *Angew. Chem.* **1995**, *107*, 1092 – 1094; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1019 – 1020.
[15] R. E. Gaillard, J. F. Stoddart, A. J. P. White, B. J. Williams, D. J. Williams, *J. Org. Chem.* **1996**, *61*, 4504 – 4505.
[16] G. W. Coates, A. R. Dunn, L. M. Henling, D. A. Dougherty, R. H. Grubbs, *Angew. Chem.* **1997**, *109*, 290 – 293; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 248 – 251.
[17] F. Ponzini, R. Zagha, K. Hardcastle, J. S. Siegel, *Angew. Chem.* **2000**, *112*, 2413 – 2415; *Angew. Chem. Int. Ed.* **2000**, *39*, 2323 – 2325.
[18] C.-Y. Kim, P. P. Chandra, A. Jain, D. W. Christianson, *J. Am. Chem. Soc.* **2001**, *123*, 9620 – 9627.
[19] C. R. Patrick, G. S. Prosser, *Nature* **1960**, *187*, 1021.
[20] J. H. Williams, J. K. Cockcroft, A. N. Fitch, *Angew. Chem.* **1992**, *104*, 1666 – 1669; *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 1655 – 1657.
[21] T. A. Millican, G. A. Mock, M. A. Chauncey, T. P. Patel, M. A. W. Eaton, J. Gunning, S. D. Cutbush, S. Neidle, J. Mann, *Nucleic Acids Res.* **1984**, *12*, 7435 – 7453.
[22] U. Wichai, S. Wosky, *Org. Lett.* **1999**, *1*, 1173 – 1176.
[23] The tritylated pentafluorophenyl-β-D-deoxyriboside was deprotected under acidic conditions ($CHCl_2COOH$, pyrrole, CH_2Cl_2, RT, 20 min) to give anomerically pure **1**. ^{1}H NMR NOE (400 MHz, [D_6]DMSO): irradiation of the signal for H-C(1′) (5.30 ppm) leads to enhancement of the signal for H-C(4′) (3.75 ppm), and vice versa.
[24] B. A. Conolly in *Oligonucleotides and Analogues: A Practical Approach* (Ed.: F. Eckstein), Oxford University Press, Oxford, **1991**, pp. 151 – 183.
[25] The identity of the oligonucleotides synthesized was confirmed by ESI mass spectrometry.
[26] Substituting the phenyl part of a F^5–P base pair by a naphthalene or indole unit could possibly resolve this problem. Experiments with the corresponding modified oligonucleotides are under way.
[27] Experiments with the non-self-complementary 10-mer duplexes d(CTGF^{5}PF^{5}PGAC)·d(GTCF^{5}PF^{5}PCAG) and d(CTGF^{5}F^{5}F^{5}F^{5}GAC)·d(GTCPPPPCAG) were inconclusive since they showed melting temperatures below 10 °C.

  0044-8249/02/4117-3205 $ 20.00+.50/0