Bipyridyl- and Biphenyl-DNA: A Recognition Motif Based on Interstrand Aromatic Stacking

Christine Brotschi, Gérald Mathis, and Christian J. Leumann*[a]

Abstract: The synthesis and incorporation into oligonucleotides of C-nucleosides containing the two aromatic, non-hydrogen-bonding nucleobase substitutes biphenyl (I) and bipyridyl (Y) are described. Their homo- and hetero-recognition properties in different sequential arrangements were then investigated via UV-melting curve analysis, gel mobility assays, CD- and NMR spectroscopy. An NMR analysis of a dodecamer duplex containing one biphenyl pair in the center, as well as CD data on duplexes with multiple insertions provide further evidence for the zipper-like interstrand stacking motif that we proposed earlier based on molecular modeling. UV-thermal melting experiments with duplexes containing one to up to seven I- or Y base pairs revealed a constant increase in T_m in the case of I and a constant decrease for Y. Mixed I/Y base pairs lead to stabilities in between the homoseries. Insertion of alternating I/abasic site- or Y/abasic site pairs strongly decreases the thermal stability of duplexes. Asymmetric distribution of I- or Y residues on either strand of the duplex were also investigated in this context. Duplexes with three natural base pairs at both ends and 50% of I pairs in the center are still readily formed, while duplexes with blunt ended I pairs tend to aggregate unspecifically. Duplexes with one natural overhang at the end of a I–I base pair tract can both aggregate or form ordered duplexes, depending on the nature of the natural bases in the overhang.

Keywords: bioorganic chemistry · DNA recognition · hydrophobic effect · nucleosides

Introduction

Hydrogen-bonding and stacking interactions between nucleobases are the major components of the noncovalent forces that stabilize the DNA and RNA double helix.[1,2] The relative contribution of each to the stability has been a matter of debate since the discovery of the double helix. About a decade ago, Kool and co-workers started to investigate stacking interactions in detail with oligodeoxynucleotide duplexes containing nonpolar nucleobase substitutes as dangling ends.[3,4] Factors such as hydrophobicity (log *P* values), polarizability, dipole moment, surface area and stacking area were discussed as contributors to the observed enhanced thermodynamic stability.[4–6] Moreover, base substitutes which mimic the shape of the natural bases but lack their hydrogen-bonding capabilities (isosters) were used as tools for studying the fidelity and mechanistic aspects of DNA-polymerase activity. It was shown that such isosters, although destabilizing a DNA duplex, can code for each other with high precision in DNA polymerase-based primer-template extension reactions.[7–13]

The stabilities of such unnatural base pairs are strongly dependent on the nature of the aromatic unit and vary greatly from distinctly less stable to more stable compared with a natural base pair. One of the first examples of a stabilizing pair lacking hydrogen bonds was the pyrene/abasic site pair.[6,12] Investigations towards the expansion of the genetic alphabet from the Schultz and Romesberg group lead to the production of a whole series of such stable, unnatural self-pairs and cross-pairs.[14–22]

In an effort to decipher the role of interstrand stacking interactions in DNA duplex stabilization with C-nucleosides containing nonfunctionalized aromatic units (Figure 1) we investigated bipyridyl self-pairs and found that two bipyridyl C-nucleoside residues Y were able to recognize each other within a DNA duplex with equal affinity as a G–C base pair.[23] Based on molecular modeling we proposed a structural motif in which the distal pyridyl rings stack upon each other with the interstrand stacking interaction as the main driving force for the observed stability.

[a] Dipl.-Chem. C. Brotschi, Dipl.-Chem. G. Mathis, Prof. C. J. Leumann
Department of Chemistry and Biochemistry
University of Bern
Freiestrasse 3, 3012 Bern (Switzerland)
Fax: (+41) 31-631-3422
E-mail: leumann@ioc.unibe.ch

 DOI: 10.1002/chem.200400858

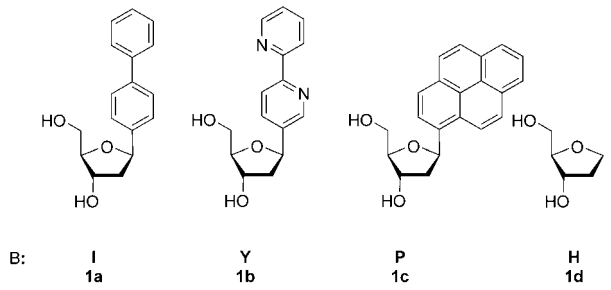


Figure 1. Chemical structures of the hydrophobic nucleosides used in this study.

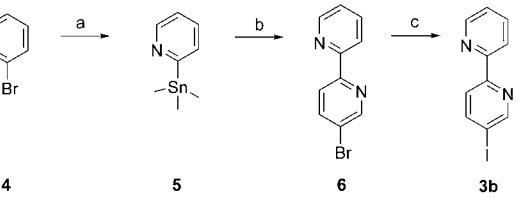
Recently we extended our investigations also on pairs of the simple biphenyl unit I and found an increase in duplex stability of 3–4 K per I pair in duplexes containing up to four of such base pairs. We proposed again a zipper-like interstrand stacking motif of the distal biphenyl rings as the most reasonable duplex structure.[24]

Here we report on the synthesis and further investigation of I and Y as hydrophobic base substitutes in various DNA sequence contexts. We provide NMR-structural and CD-spectroscopic evidence supporting the zipper-like interstrand stacking motif and show scope and limitations of this recognition motif as known so far.

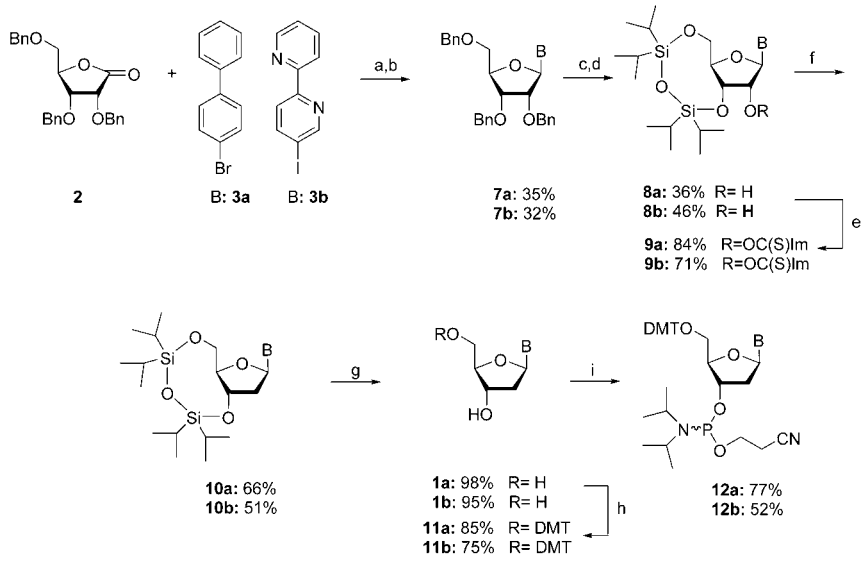
Results

Synthesis of monomers: Of the hydrophobic nucleoside analogues used in this work, **1a** was already known and was incorporated into DNA as a tool to study enzymatic base methylation in DNA.[25,26] However, its synthesis was different from the one described here. The analogue **1c** was obtained along the same lines as described.[6] Our syntheses of the C-nucleosides **1a** and **1b** followed established routes in C-glycoside chemistry[27,28] (Scheme 1) and started with 2,3,5-tri-*O*-benzyl-D-ribonolactone **2**[29] as the common deoxyribose precursor to which the lithiated 4-bromo-2,2-biphenyl **3a** or 5-iodo-2,2-bipyridine **3b**[30,31] were added. While **3a** is commercially available, **3b** had to be prepared in three steps by stannylation of 2-bromopyridine **4** (→**5**),[32] followed by a $[Pd(PPh_3)_4]$ catalyzed Stille coupling with 2,5-dibromopyridine and a subsequent copper(I)-catalyzed halogen-exchange reaction (Scheme 2). The halogen exchange was necessary as, in our hands, the intermediate 5-bromo-2,2-bipyridine **6**[32] failed to react to the coupled product **7b** by the standard procedure.

Lithiation of **3a** or **3b** followed by addition to lactone **2** resulted in the intermediate formation of the corresponding hemiacetals, which upon reduction with $Et_3SiH/BF_3{\cdot}Et_2O$ afforded selectively the β-isomeric C-glycosides **7a** and **7b** in moderate but sufficient yield (ca. 35%). The assignment of the anomeric configuration in **7a** and **7b** was confirmed by ^{1}H NMR NOE spectroscopy (see Experimental Section). After removal of the benzyl protecting groups with BBr_3 the 3′- and 5′-hydroxyl groups were selectively protected with 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane to give **8a** and **8b**, again in moderate yield most likely due to side reactions at the benzylic, pseudo-anomeric center C(1′). However, scrambling of the configuration at C(1′) during benzyl-deprotection could be excluded again by ^{1}H NMR NOE evidence at the stage of the tritylated com-



Scheme 2. Synthesis of 5-iodo-2,2-bipyridine. a) **4** (1 equiv), *n*BuLi (1 equiv), Et_2O, −78 °C then Me_3SnCl (1 equiv) in THF, 93%; b) 2,5 dibromopyridine (1.1 equiv), $[Pd(PPh_3)_4]$, *m*-xylene, reflux, 86%; c) CuI (0.05 equiv), NaI (2 equiv), *trans-N,N′*-dimethyl-1,2-cyclohexanediamine (0.1 equiv), dioxane, 110 °C, 94%.



Scheme 1. a) **3a** or **3b** (1.0 equiv), *n*BuLi (1.0 equiv), THF, −78 °C, then **2** (0.9 equiv), −78 °C to RT; b) Et_3SiH, (5 equiv), CH_2Cl_2, −78 °C to RT; c) BBr_3 (3.4 equiv), CH_2Cl_2, −78 °C; d) 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane (1.0 equiv), pyridine, RT; e) 1,1′-thiocarbonyldiimidazole (1.2 equiv), CH_3CN, RT; f) Bu_3SnH (1.6 equiv), AIBN (1.5 equiv), toluene, 80 °C; g) $NEt_3{\cdot}3HF$ (10.0 equiv), THF, RT; h) 4,4′-dimethoxytrityl (DMT) chloride (1.2 equiv), pyridine, RT; i) iPr$_2$NEt (3 equiv), $[(i\text{Pr}_2\text{N})(\text{NCCH}_2\text{CH}_2\text{O})\text{P}]\text{Cl}$ (1.5 equiv), THF, RT. AIBN = α,α′-azobisisobutyronitrile.

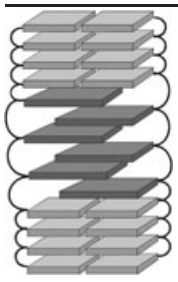
pounds **11a** and the TIPS-protected intermediate **8b**. Deoxygenation of the 2′-hydroxy group in **8a** and **8b** by means of a Barton–McCombie reduction lead to **10a** and **10b**. Desilylation with (HF)·NEt_3 gave the unprotected deoxynucleoside analogues **1a** and **1b** that were subsequently converted into the corresponding phosphoramidite building blocks **12a** and **12b** via tritylation (4,4′-dimethoxytrityl chloride) followed by phosphitylation in typical yields for these transformations.

Synthesis of oligonucleotides: Oligonucleotides were prepared according to standard protocols for automated DNA synthesis on a 1 μmol scale in the trityl-off mode. The coupling time for the modified phosphoramidites was extended to 6–10 min, and ethyl thio-1*H*-tetrazole was used as activator. Typical coupling yields for the modified building blocks, as deduced from the trityl assay, were in the range of 98%. After assembly, the oligonucleotides were detached from the solid support and deprotected in concentrated aqueous ammonia (12–18 h at 55 °C) and were purified by HPLC. All oligonucleotides were routinely characterized by ESI^- mass spectrometry. The synthesis and analytical data for all modified oligonucleotides used in this study are summarized in Table 6 (see Experimental Section).

We then examined the structural and thermal melting properties of duplexes by UV melting curve analysis, gel mobility experiments, CD and NMR spectroscopy.

Pairing properties of duplexes containing I or Y base pairs in opposing positions ("stretched backbone" series): First, we examined the stability of non self-complementary oligonucleotide duplexes containing one or multiple aromatic base pairs (I or Y) in the center of the sequence. The sequences were designed to determine not only homo base pair but also hetero base pair formation. The thermal stabilities of the corresponding duplexes were determined by UV-melting curve analysis, and the corresponding T_m data are summarized in Table 1.

Table 1. Duplex sequence information and a schematic representation of the interstrand-stacking model, as well as corresponding T_m values.

5′-GATGAC(X)$_n$GCTAG 3′-CTACTG(Z)$_n$CGATC	n	T_m[a,d] [°C] X=I Z=I	T_m[b,d] [°C] X=Y Z=Y	T_m[a,d] [°C] X=Y Z=I
	0		45.0	
	1	42.5	48.0	45.7
	2	46.9	42.9	45.3
	3	49.9	40.7	46.0
	4	53.2	39.4	46.0
	5	55.5	37.5	45.2
	6	56.2	35.0	42.5
	7	57.0	n.d.[c]	n.d.[c]

[a] $c=1.2$ μM duplex, 10 mM NaH_2PO_4, 0.15 M NaCl, pH 7.0. [b] $c=1.2$ μM duplex, see Table 2, 10 mM NaH_2PO_4, 0.15 M NaCl,1 mM EDTA, pH 7.0. [c] Not determined. [d] Estimated error in $T_m=\pm0.5$ °C.

In the case of I all melting profiles show a single, highly cooperative transition. For melting experiments with the base Y 1 mM EDTA was added to the buffer in order to suppress metal chelation by the bipyridyl ligand. Also in these cases only single transitions were observed. Examples of representative melting curves are depicted in Figure 2. The steeper base lines before and after the transition in melting experiments with bipyridyl oligonucleotides were found to be due to the EDTA in the buffer.

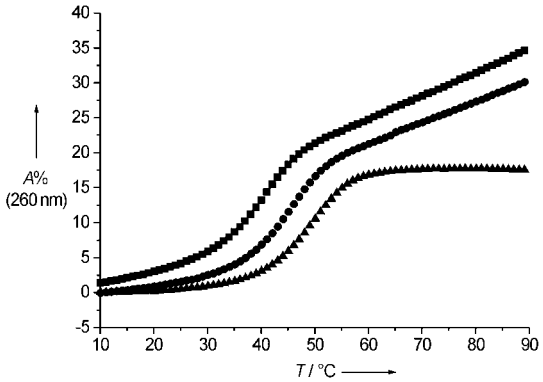


Figure 2. UV-melting curves (260 nm) of duplexes containing I–I pairs (▲, $n=3$), Y–Y pairs (■, $n=3$) and I–Y mixed pairs (●, $n=3$).

The differences in T_m upon incorporation of a growing number of I-, Y- or mixed base pairs are illustrated in Figure 3. Incorporation of one I base pair leads to a drop in T_m by 2.5 K relative to the unmodified duplex while additional consecutive I base pairs lead to an increase in T_m of initially 4.4 K per base pair to 0.7 K for the incorporation of the seventh base pair. On the other hand incorporation of one Y base pair in the given sequence leads to a duplex that is by 3 K more stable than the corresponding unmodified duplex. However, contrary to the I series, any additional Y base pair decreases duplex stability. The mixed, I–Y series produces duplexes with almost invariant T_m relative to the number of aromatic base pairs, with a tendency to lower T_m for $n=6$. Thus the duplex stability of the mixed series lies in between that of the two homoaromatic series.

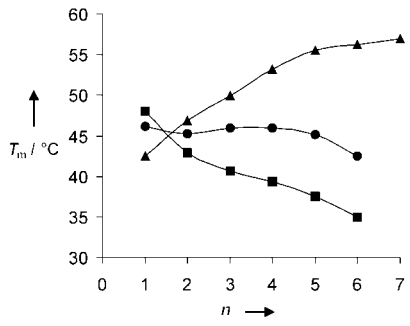


Figure 3. Graphical representation of the stabilities of duplexes from Table 1. I–I (▲, $n=1$–7), Y–Y (■, $n=1$–6) and I–Y (●, $n=1$–6).

The higher stability of the duplex containing one Y base pair compared to that with one I base pair ($\Delta T_m = 5.5$ K) seems to be in inverted order to the tendency of duplex stabilization for the two aromatic units. This observation becomes plausible if one considers the stacking interactions of the aromatic residues with the next natural base pair. Given the fact that the bipyridyl system is intrinsically planar, it is expected to perturb its nearest neighbor natural base pair less than a biphenyl unit, which is intrinsically nonplanar.

NMR analysis of a duplex containing two opposing I units: In order to get preliminary structural information we measured ^{1}H NMR spectra of the duplex containing one I base pair (Table 1, $n=1$) at different temperatures. The proton signals arising from the biphenylic units in the duplex can easily be identified by their chemical shift (6.5–7.0 ppm), lying in between the signals of the nonexchangeable nucleobase protons and the anomeric sugar protons. Upon thermal denaturation of the duplex the biphenyl signals shift in a cooperative way towards lower field and end up at 7.1–7.3 ppm in the single stranded state (Figure 4, shaded area). This shift implies considerable loss of ring current effects upon denaturation which is in agreement with the biphenyl units stacking on each other in the duplex. Chemical shift arguments have recently also been used as a criterion of intercalation in a duplex containing a pyrene pair by others.[33] The T_m that can be deduced from the cooperative shift of the biphenyl protons versus T is at about 55 °C and thus about 13 K higher than that measured by UV-melting curves. This difference is expected and can be explained by the 1000-fold higher oligonucleotide concentration in the NMR sample. Thus the highly cooperative deshielding of the biphenyl protons with a T_m in the expected range fully supports the interstrand stacking model. A detailed structural analysis of this duplex by 2D-NMR techniques is currently underway and will be reported soon.

CD spectra: A CD-spectroscopic investigation of selected duplexes from Table 1 containing an increasing number of I- or Y base pairs was performed and compared with the unmodified duplex. The CD traces for I duplexes are compared in Figure 5a, and those of the Y duplexes in Figure 5b. In the case of the I duplexes, a shift of the negative maximum towards 250 nm with increasing numbers of I units is observed. At the same time the positive maximum shifts from 275 nm in the unmodified duplex to about 270 nm and grows in intensity as a function of the number of I base pairs. This is due to the contribution of the biphenyl chromophores, which have their maximum absorbance at about 250 nm. The continuous shift of the CD spectrum upon increasing the number of I base pairs without a sudden major change in its shape is indicative for highly ordered helical structures and disproves hydrophobic collapse of the central biphenylic core part.

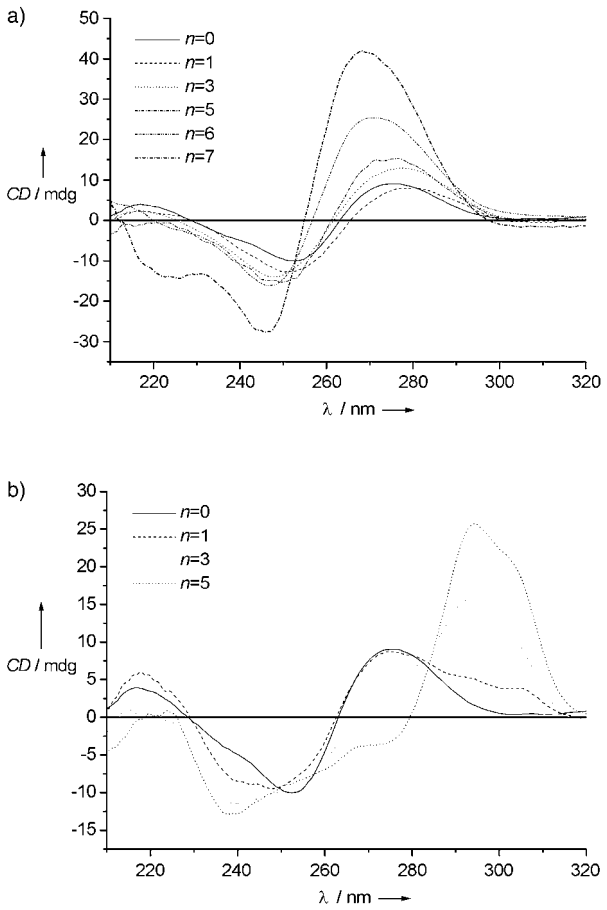


Figure 5. CD spectra of selected duplexes from Table 1. Top: base I, $n=$ 0,1,3,5,6,7; $c=3.6$ μM in 10 mM NaH_2PO_4, 0.15 M NaCl, pH 7, $T=20$ °C; bottom: base Y, $n=0,1,3,5$. $c=3.6$ μM in 10 mM NaH_2PO_4, 0.15 M NaCl, 1 mM EDTA, pH 7, $T=20$ °C.

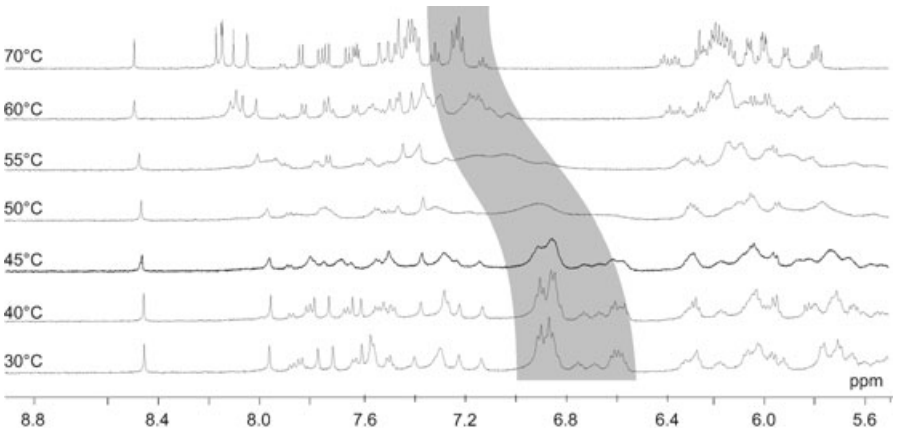
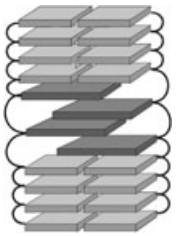


Figure 4. Temperature dependent ^{1}H NMR spectra (500 MHz) of the DNA duplex 5′d(GATGACIGCTAG)-d(CTAGCIGTCATC) ($c=1$ mM, 10 mM NaH_2PO_4, 150 mM NaCl, pH 7, 100 % D_2O).

In the case of the Y duplexes a similar picture emerges as n is increased from 1 to 5 (Figure 5b). The minimum ellipticity at about 255 nm in the unmodified duplex is shifted towards shorter wavelengths (about 240 nm) as the number of Y residues is increased. At the same time the maximum ellipticity shifts from about 275 to about 295 nm and gains intensity. This again is most likely due to the bipyridyl chromophore becoming increasingly CD active. Interestingly an isodichroic point arises at 284 nm. From this we conclude again that formation of consecutive Y base pairs occurs in a highly ordered manner and leads to duplexes of one structural family, as in the case of the I base pairs. Thus also the CD spectra are in support of an interstrand stacking model.

Permutational analysis of the thermal stability of duplexes with two consecutive I and Y base pairs: For the duplexes with two consecutive aromatic units we evaluated the stability of all eight possible arrangements of I and Y (Table 2, entries 1–8). The T_m values of the six mixed duplexes are lying in between the T_m values of the fully Y-substituted duplex (42.9 °C) and the fully I-substituted duplex (46.9 °C). The T_m values for the six mixed cases vary by 3.2 K.

Table 2. Permutational analysis of the thermal stability of pairs of **I** and **Y** in a two-base system.

		Duplex	T_m[a,c] [°C]
	1	5′-GATGAC**YY**GCTAG 3′-CTACTG**YY**CGATC	42.9
	2	5′-GATGAC**II**GCTAG 3′-CTACTG**II**CGATC	46.9[b]
	3	5′-GATGAC**II**GCTAG 3′-CTACTG**YY**CGATC	45.7
	4	5′-GATGAC**YY**GCTAG 3′-CTACTG**II**CGATC	45.2
	5	5′-GATGAC**YI**GCTAG 3′-CTACTG**YI**CGATC	44.8
	6	5′-GATGAC**YI**GCTAG 3′-CTACTG**IY**CGATC	43.4
	7	5′-GATGAC**YY**GCTAG 3′-CTACTG**YI**CGATC	46.6
	8	5′-GATGAC**YY**GCTAG 3′-CTACTG**IY**CGATC	43.7

[a] $c = 1.2$ μM, 10 mM NaH_2PO_4, 0.15 M NaCl, 1 mM EDTA, pH 7.0. [b] Same buffer without EDTA. [c] Estimated error in $T_m = \pm 0.5$ °C.

The relative differences in duplex stability as a function of the nature of the aromatic units is most expressed in the homoaromatic systems and to a lesser extend in the mixed series. Nevertheless there are distinct sequence effects which eventually may establish a recognition code that is entirely decoupled from hydrogen-bonding interactions and relies only on differential stacking interactions, or possibly differential solvation.

Pairing properties of duplexes containing I- or Y-abasic site pairs in an alternating fashion ("relaxed backbone" series): In contrast to I–I or Y–Y base pairs that upon intercalation inevitably lead to a sugar phosphate backbone that is extended along the helical axis, we also investigated alternating arrangements of I–H and Y–H pairs (H = abasic site). In these systems interstrand stacking of the aromatic units is still possible without necessarily invoking a change of the backbone conformation relative to that of an unconstrained B-DNA duplex ("relaxed backbone" series, Table 3).

Within this structural series we first investigated duplexes containing two H–H, T–H and G–H pairs (Table 3, entries 1–3). The T_m data clearly show an increase in the order of the H–H duplex (19 °C) to the T–H duplex (27.5 °C) to the G–H duplex (32.6 °C). This remarkable increase in T_m goes parallel with the increasing potential of the bases to overlap, and thus is in agreement with the interstrand stacking model also in the "relaxed backbone" series. Within the same structural motif the nonpolar aromatic I–H (35.7 °C), Y–H (39.5 °C) and P–H (43.3 °C) pairs lead to duplexes with considerably higher T_m values. The T_m values for the mixed, I–H/Y–H, duplexes (entries 7 and 8) are lying in between the T_m values for the I–H (entry 4) or Y–H (entry 5) substituted duplex. Thus, the order of stability of the mixed I- and Y duplexes follows exactly that of the "nonrelaxed" backbone series (Table 1).

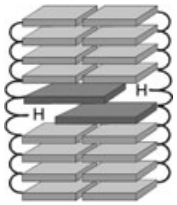
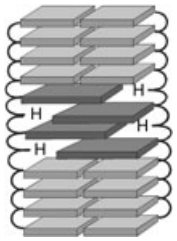
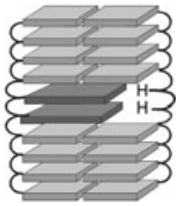
Melting experiments of duplexes containing four such base pairs (entries 9–11) show a dramatic decrease of duplex stabilities at low salt conc. (150 mM NaCl) up to an extent where no distinct T_m value can be determined. Therefore the T_m values were also measured at high salt (1 M NaCl). Interestingly the Y–H and I–H duplexes (entries 9 and 10) led to higher T_m values than the P–H duplex (entry 11). This is in agreement with earlier data on duplexes carrying up to four P–H pairs that resulted in strongly reduced or no helix formation, respectively.[6] The failure of duplex formation with multiple P units may be due to competing hydrophobic aggregation of these residues in the corresponding single strands.

We also investigated duplexes where the Y–H and I–H base pairs are in a nonalternating sequence context (Table 3, entries 12 and 13). A comparison of the T_m values of the corresponding alternating duplexes (entries 5 and 6) shows that in the Y–H case the alternating arrangement is favored by 3.2 K. In the I–H case there is no difference in T_m regardless whether the pairs are arranged in an alternating or nonalternating manner. Generally, higher T_m values were expected in the alternating relative to the nonalternating cases due to energetic differences arising from interstrand versus intrastrand stacking. However, this is only encountered in the Y–H case. Surprisingly it does not happen in the I–H case. The reason for it is yet unknown, although we note that differences in the preferred orientation of the biarylic axis in the units I relative to Y could account for it. In any case more structural data are necessary for a more detailed interpretation of this fact.

CD experiments of selected duplexes (Table 3, entries 9–11) are overlaid in Figure 6. No distinct deviations from a B-DNA conformation are observed.

Asymmetric distribution of I- and Y base pairs in the duplex: Unpaired bases as occurring in hairpin–loop struc-

Table 3. Thermal stabilities of hydrophobic pairs in the "relaxed backbone" structural alignment, determined by UV-melting curves; **I** (**1a**), **Y** (**1b**), **P** (**1c**), **H** (**1d**), **G** and **T** = deoxyguanosine and thymidine, respectively

		duplex	T_m[a,c] [°C]	T_m[b,c] [°C]
	1	5′-GATGAC**H**HGCTAG 3′-CTACTGH**H**CGATC	19.0	
	2	5′-GATGAC**T**HGCTAG 3′-CTACTGH**T**CGATC	27.5	
	3	5′-GATGAC**G**HGCTAG 3′-CTACTGH**G**CGATC	32.6	
	4	5′-GATGAC**I**HGCTAG 3′-CTACTGH**I**CGATC	35.7	
	5	5′-GATGAC**Y**HGCTAG 3′-CTACTGH**Y**CGATC	39.5	
	6	5′-GATGAC**P**HGCTAG 3′-CTACTGH**P**CGATC	43.3	
	7	5′-GATGAC**I**HGCTAG 3′-CTACTGH**Y**CGATC	36.7	
	8	5′-GATGAC**Y**HGCTAG 3′-CTACTGH**I**CGATC	36.8	
	9	5′-GATGA**HIHI**CTAG 3′-CTACT**IHIH**GATC	~5	20.0
	10	5′-GATGA**HYHY**CTAG 3′-CTACT**YHYH**GATC	~10	23.0
	11	5′-GATGA**HPHP**CTAG 3′-CTACT**PHPH**GATC	~5	18.0
	12	5′-GATGAC**YY**GCTAG 3′-CTACTG**HH**CGATC	36.3	
	13	5′-GATGAC**II**GCTAG 3′-CTACTG**HH**CGATC	35.6	

[a] $c = 1.2$ μM, 10 mM NaH_2PO_4, 0.15 M NaCl, pH 7.0. [b] $c = 1.2$ μM, 10 mM NaH_2PO_4, 1.0 M NaCl, pH 7.0. [c] Estimated error in $T_m = \pm 0.5$ °C.

tures or in bulged structures are common natural structural motifs. We became interested in the question of how the biarylic units I and Y behave in bulge positions and therefore investigated duplexes with an asymmetric distribution of I units in both strands. In system A (Table 4) we investigated single bulge situations (n/$n+1$, $n = 1$–6) with both I- and Y residues. In system B (Table 4) we looked into duplexes with variable bulge size (2/n+2, $n = 1$–5) in the case of the biphenylic unit I. The T_m values for system A and B are summarized in Table 4 and graphically represented in Figure 7.

In the I series, single bulges within an increasing number of I pairs (system A) go along with an increase in the T_m of the duplexes much in the same way as their symmetric equivalents (Table 1). The analogous negative trend in T_m is observed in the Y series. Interestingly, asymmetric variation of the bulge size (system B) in the case of I is accompanied with an almost invariant T_m which is in between that of the symmetric duplexes with $n = 2$ and 3 (Table 1). Only with bulge sizes larger than $n = 4$ (Table 4, system B, entries 2–6 and 2–7), the duplexes become thermally less stable. The CD spectra of the duplexes of system A and B are not significantly different from that of the corresponding symmetric sequences (data not shown).

Sequences with terminal I stretches or without natural base pairs: To explore whether natural base pairs flanking the I pairs are necessary in order to maintain discrete duplex

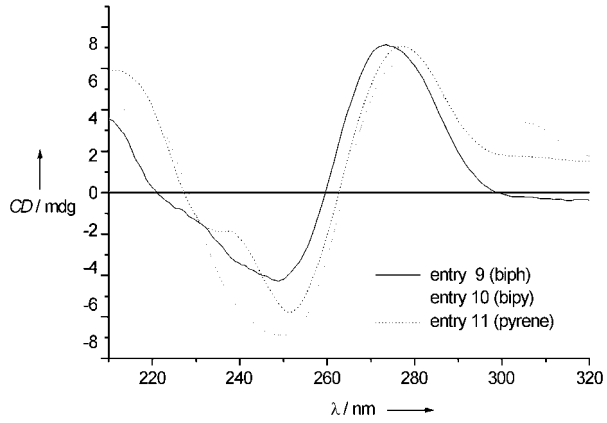


Figure 6. CD spectra ($T = 20$ °C) of selected duplexes from Table 3.

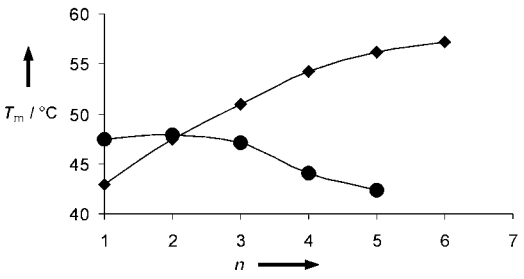


Figure 7. Graphical representation of the thermal stabilities (T_m) vs the number of I residues (n) of duplexes in the asymmetric systems A (◆) and B (●) from Table 4.

Table 4. Thermal stabilities of duplexes containing an asymmetric distribution of hydrophobic pairs. System A: $n/n+1$; System B: $2/2+n$; **I** (**1a**), **Y** (**1b**).

System A $n/n+1$	X=I T_m[a,d] [°C]	Duplex	X=Y T_m[b,d] [°C]
1/2	42.9	5′-GATGAC**X**GCTAG 3′-CTACTG**X**CGATC	43.7
2/3	47.5	5′-GATGAC**XX**GCTAG 3′-CTACTG**XX**CGATC	41.4
3/4	51.0	5′-GATGAC**XXX**GCTAG 3′-CTACTG**XXX**CGATC	39.6
4/5	54.2	5′-GATGAC**XXXX**GCTAG 3′-CTACTG**XXXX**CGATC	38.4
5/6	56.2	5′-GATGAC**XXXXX**GCTAG 3′-CTACTG**XXXXX**CGATC	n.d.[c]
6/7	57.2	5′-GATGAC**XXXXXX**GCTAG 3′-CTACTG**XXXXXX**CGATC	n.d.[c]

System B $2/n+2$	X=I T_m[a]	Duplex
2/3	47.5	5′-GATGAC**XX**GCTAG 3′-CTACTG**XXX**CGATC
2/4	47.9	5′-GATGAC**XX**GCTAG 3′-CTACTG**XXXX**CGATC
2/5	47.1	5′-GATGAC**XX**GCTAG 3′-CTACTG**XXXXX**CGATC
2/6	44.1	5′-GATGAC**XX**GCTAG 3′-CTACTG**XXXXXX**CGATC
2/7	42.4	5′-GATGAC**XX**GCTAG 3′-CTACTG**XXXXXXX**CGATC

[a] $c=1.2$ μM, 10 mM NaH_2PO_4, 0.15 M NaCl, pH 7.0. [b] $c=1.2$ μM, 10 mM NaH_2PO_4, 0.15 M NaCl, 1 mM EDTA, pH 7.0. [c] Not determined. [d] Estimated error in $T_m=\pm0.5$ °C.

Table 5. I-modified oligonucleotide single strands and duplexes used in the gel retardation experiments, the corresponding T_m data from UV-melting curves and discernible secondary structure formation.

Entry	Sequence	T_m[a,d] [°C]	Secondary structure
1	5′-GATGAC**IIIIII**	mmp[b]	aggregate
2	5′-GATGAC**IIIIII** 3′-CTACTG**IIIIII**	mmp[b]	aggregate
3	5′-GATGAC**IIIIII**GCTAG 3′-CTACTG**IIIIII**	mmp[b]	aggregate
4	5′-GATGAC**IIIIII** 3′-CTACTG**IIIIII**CGATC	51.3	duplex
5	5′-GATGAC**IIIIII**GCTAG 3′-CTACTG**IIIIII**CGATC	57.0	duplex
6	5′-GAT**IIIIII**TAG	–	single strand
7	5′-GAT**IIIIII**TAG 3′-CTA**IIIIII**ATC	n.d.	duplex
8	3′-CTA**IIIIII**ATC	–	single strand
9	5′-T**IIIIIIIIIII**T	n.d.	aggregate
10	5′-TTTTTTTT**IIIIIIIIIII**T	n.d.	single strand
11	5′-TTTTTTTT**IIIIIIIIIII**T 3′-T**IIIIIIIIIII**T	n.d.	mixed aggregate

[a] $c=1.2$ μM, 10 mM NaH_2PO_4, 0.15 M NaCl, pH 7.0. [b] Multiple overlaying melting processes were observed [c] n.d. = no T_m detectable due to low hyperchromicity. [d] Estimated error in $T_m=\pm0.5$ °C.

structures we prepared and analyzed the duplexes listed in Table 5. A 13-mer duplex containing seven I pairs and six natural base pairs at one end (Table 5, entry 2) failed to show a sigmoidal transition in the UV-melting curve. A gel mobility experiment (Figure 8, lane 2) showed multiple bands reminiscent of less-defined aggregate formation of the two strands with themselves or with each other. Aggregate formation occurs also with a strand containing a 3′-natural overhang (entry 3) or with one of the single strands (entry 1) alone. By contrast, the duplex containing a 5′-natural overhang (entry 4) shows a clear and highly sigmoidal transition in the UV-melting curve (T_m 51.3 °C) and a single duplex band in the gel (Figure 8, entry 4). This T_m is only 5.7 K lower than that of the duplex containing natural base pairs at both ends (Table 5, entry 5) and is 42.4 K higher than that of the natural hexamer duplex consisting of the first six base pairs ($T_m=8.9$ °C).[34] From these experiments we conclude that stable I pairs can exist also in duplexes with natural base pairs on one end and a 5′-overhang of natural bases on the other end, while duplexes with blunt-ended I pairs or a 3′-overhang tend to aggregate. The reason of the differential behavior of duplexes with 5′- or 3′-overhang is unknown at present, but is expected to be due to stacking of the overhang on the neighboring I pair. In another gel mobility experiment we investigated duplexes with shortened ends of natural base pairs Figure 8, entries 6–8). In this case it becomes clear that three natural base pairs at both ends suffice to produce a discrete duplex structure. Its T_m could not be determined due to the observed low hyperchromicity.

In a next step we wanted to explore whether a pure I duplex can exist without the help of flanking natural base pairs. Due to the expected low hyperchromicity upon melting of a fully biphenylic system, we adopted an experimental approach developed by Switzer et al.[35] We prepared two single strands containing 12 I residues each, that were tagged with one and eight thymidine residues at their 5′-end, respectively, resulting in different oligonucleotide lengths (Table 5, entries 9 and 10). Assuming that the flank-

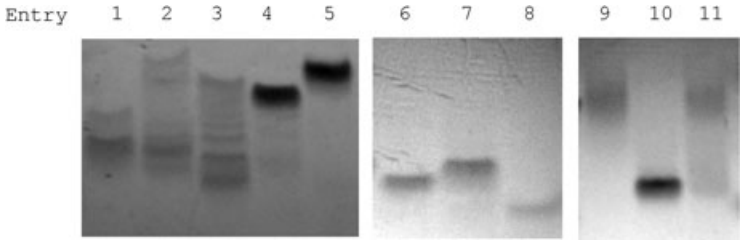


Figure 8. 20% Polyacrylamide gel electrophoresis of single strands and mixtures corresponding to entries 1–11 from Table 5). Bands were visualized by UV light (260 nm). Entries correspond to Table 5.

ing T residues do not alter strand affinities, in a 1:1 mixture of both strands one would expect three bands in the case of statistic duplex formation. However, no discrete band of a single strand was found for the shorter oligonucleotide (Figure 8, lane 9), while a sharp band with the expected mobility of a single strand was observed for the longer oligonucleotide (entry 10). The mixture of both showed the disappearance of the band of the longer oligonucleotide without the appearance of discrete new bands (entry 11) eventually indicating cross-pairing.

Discussion

I- versus Y pairs: The increasing thermal stability of duplexes with an increasing number of I pairs contrasts findings in the Y series where decreasing thermal stability with increasing number of bipyridyl residues is observed. Obvious factors responsible for this behavior are differences in stacking or solvation energies, or differences in the preferred orientation of the biarylic axis. In general, Kool et al.[4] found that hydrophobic effects are more important in stabilizing stacking than other effects (electrostatic effects, dispersion forces). However, the natural DNA bases are found to be less dependent on hydrophobic effects than the more nonpolar compounds. We exclude differences in the polarizability (α_m) of the two aromatic systems as the major factor, as calculations revealed very similar α_m values (18.64 Å^3 for Y, both *syn* and *anti* isomer, 20.15 Å^3 for I). More likely, differences in solvation of the edges of the aromatic units in the major or minor groove in the duplex state, or structural differences around the biphenylic axis (intrinsically planar in the case of Y and intrinsically nonplanar in the case of I) are responsible for the stability differences. The latter could also explain why a single I base pair is less stable compared with a single Y base pair, if its stacking interaction with the nearest neighbor natural base pair is taken into account.

Relaxed versus stretched backbone: It is long known that ethidium bromide can intercalate up to an extent of one ethidium per 2.5 base pairs in DNA–polymer duplexes.[1] This illustrates the flexibility of the DNA backbone which allows for considerable breathing along the helical axis. It is thus not surprising that a zipper-motif as described here in the "stretched backbone" series is tolerated by the phosphodiester backbone. It is notable that this zipper corresponds to an intercalator/base-pair ratio of 1:1; this indicates that even further extension along the helical axis is tolerated by the backbone. It is unclear at present if and where the limitations in length of this recognition motif lie. The fact that the "relaxed backbone" series with Y–H and I–H pairs are in all cases less stable than the "stretched backbone" series with I–I and Y–Y pairs may well have its origin in differences in the entropy of duplex formation, the former series having more degrees of conformational freedom than the latter series.

An interesting question is also, whether the RNA backbone, which is intrinsically more compact and stiffer, will tolerate such a zipper-like intercalation motif or not. This would certainly have implications on research in the area of novel base pairs for biotechnological applications.[14–22]

Fully modified duplexes: An obvious question is, whether fully modified duplexes built only from I- or Y base pairs exist and are stable under standard conditions. We found that defined DNA duplex formation from two single strands containing six I residues and three natural base pairs at either end readily takes place (Figure 8, lane 7). Furthermore, we know that natural base pairs on one end of the duplex can (but must not) be sufficient for defined duplex formation (Figure 8, lane 4). On the other hand, oligonucleotides that have no matching natural base pairs, or that show blunt ended I pairs (Figure 8, lane 1, 2, 9–11) either tend to aggregation or do not pair at all. Therefore a few natural base pairs at one end of the duplex seem to be required to align two strands in a defined structural register and to prevent undefined aggregation.

Conclusion

With the structural and biophysical data contained in this communication we lend further evidence for the existence of the interstrand-stacking structural model of non-hydrogen bonding aromatics as base substitutes that we first proposed a couple of years ago.[23,24] These findings underline the more and more recognized importance of interstrand stacking interactions for nucleic acid double-helix stability. In addition it demonstrates the structural dynamics of the DNA backbone towards an elongation along the helical axis.

Besides representing an ideal scaffold for the study of the energetics of stacking interactions it could also be a structural motif of interest in nanotechnology or molecular electronics. In this context we are currently interested in the following questions: i) how do stacking energies depend on the electronic nature of the aromatic units; ii) does there exist a recognition code based entirely on differential stacking interactions; iii) can donor and acceptor substituted biphenyls form extended charge transfer complexes when alternatingly stacked within the center of a helix; and iv) can electrons or charges be efficiently transported across the zipper much in the same way as in natural DNA, and can this be kinetically

and thermodynamically tuned via electronic variation (chemical substitution) of the biphenyl units.

Experimental Section

General: Reactions were performed under argon in distilled, anhydrous solvents. All chemicals were reagent grade from Fluka or Aldrich. ^{1}H NMR (300 MHz, 500 MHz) spectra were recorded on a Bruker AC-300 or Bruker DRX-500 spectrometer; the chemical shifts δ/ppm were referenced to residual undeuterated solvent ([D]chloroform) = 7.27, [D]methanol = 3.35); coupling constants J in Hz. ^{13}C NMR (75 MHz) were recorded on a Bruker AC300; the chemical shifts δ/ppm were referenced to residual undeuterated solvent ([D]chloroform = 77.00, [D]methanol = 49.3). Carbon multiplicity (s,d,t,q) from DEPT spectra.^{31}P NMR spectra (162 MHz) were recorded on a Bruker DRX-400 spectrometer; the chemical shift δ/ppm was referenced to 85% H_3PO_4 as external standard. LSIMS and EI mass spectra were recorded on a Auto Speq Q VG at 70 eV, ESI-MS mass spectra on a Fisons Instrument VG Platform. For TLC, pre-coated plates SIL-G UV254 (Macherey Nagel) have been used and visualized by UV and/or dipping into a solution of $Ce(SO_4)_2$ (10.5 g), phosphormolybdic acid (21 g), H_2SO_4 (60 mL), and H_2O (900 mL). Flash chromatography (FC) was performed with silica gel 60 (230–400 mesh). Abbreviations: EtOAc: ethyl acetate; TEA: triethylamine; TIPDS: 1,3-dichloro-1,1,3,3-tetraisopropyl disiloxane; DMT: 4,4′-dimethoxytrityl.

Compound 5: A solution of *n*BuLi (1.6 M in hexane, 62.7 mL, 104 mmol) was diluted with dry Et_2O (45 mL). A solution of **4** (10.2 mL, 103.7 mmol) in dry Et_2O (100 mL) was added at −78 °C. The mixture was stirred for 1 h at −78 °C. A solution of Me_3SnCl (20.0 g, 104 mmol) in dry THF (100 mL) was then added dropwise at −78 °C. The cooling bath was removed and the reaction mixture was allowed to warm to RT. The color changed from brown to green and finally to yellow. The reaction was quenched with sat. NH_4Cl and the organic layer was washed with H_2O, brine, dried ($MgSO_4$) and evaporated. Compound **5** (21.2 g, 93%) was obtained by distillation under reduced pressure as a colorless liquid. ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ = 8.73 (d, ^3J(H,H) = 4.4 Hz, 1H), 7.54–7.42 (m, 2H), 7.15–7.11 (m, 1H), 0.34 (s, 9H; CH_3); ^{13}C NMR (75 MHz, $CDCl_3$, 25 °C): δ = 173.51 (s), 150.51 (d), 133.51 (d), 131.58 (d), 122.24 (d), −9.50 (q); MS (70 eV, EI): *m*/*z* (%): 243/241 (25/20) $[M]^+$, 228/226 (100/75), 198/196 (77/42), 135/133 (53/41).

Compound 6: 2,5-Dibromopyridine (36.9 g, 0.156 mol) was added to a solution of 2-trimethylstannylpyridine (**5**, 34.5 g, 0.141 mol) in *m*-xylene (300 mL), and the reaction mixture was degassed by gently bubbling argon through the solution for 1 h. Then, $[Pd(PPh_3)_4]$ (1.64 g, 1 mol%) was added and heated to 120 °C under stirring. After 12 h the mixture was cooled and poured into 2 M NaOH. The phases were separated, and the aqueous layer was extracted with toluene (2 × 1000 mL). The combined organic phases were dried ($MgSO_4$) and evaporated under reduced pressure. Compound **6** (28.53 g, 86%) was obtained after FC (hexane/EtOAc/TEA 8:2:2%) as a white solid. R_f = 0.30; ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ = 8.73 (d, ^3J(H,H) = 2.5 Hz, 1H), 8.67 (dt, ^3J(H,H) = 0.8, 3.2 Hz; 1H), 8.38 (d, ^3J(H,H) = 7.9 Hz, 1H), 8.32 (d, ^3J(H,H) = 8.5 Hz, 1H), 7.94 (dd, ^3J(H,H) = 2.3, 8.5 Hz, 1H) 7.82 (td, ^3J(H,H) = 1.7, 7.5 Hz, 1H), 7.32 (ddd, ^3J(H,H) = 1.5, 4.7, 7.4 Hz, 1H); ^{13}C NMR (75 MHz, $CDCl_3$, 25 °C): δ = 155.04 (s), 154.42 (s), 150.18 (d), 149.08 (d), 139.49 (d), 137.14 (d), 124.00 (d), 122.36 (d), 121.17 (s), 121.02 (d); MS (70 eV, EI): *m*/*z* (%): 234/236 (100/100) $[M]^+$, 155 (76), 128 (82), 78 (39).

Compound 3b: A Schlenk tube was charged with compound **6** (8 g, 0.034 mol), CuI (0.323 g, 0.0017 mol) and NaI (10.2 g, 0.068 mol), briefly evacuated and flushed with argon. Racemic *trans-N,N′*-dimethyl-1,2-cyclohexanediamine[28] (0.544 mL, 0.0034 mol) and dioxane (34 mL) were added. The Schlenk tube was sealed and the reaction mixture stirred at 110 °C for 70 h. The brown suspension was allowed to reach RT, was diluted with 30% aqueous ammonia (180 mL), poured into water (500 mL) and extracted with CH_2Cl_2 (3 × 200 mL). The combined organic layers were dried ($MgSO_4$) and concentrated in vacuo. Compound **3b** (9.02 g, 94%) was obtained after FC (hexane/EtOAc/TEA 8:2:2%) as a white solid. R_f = 0.30 (hexane/EtOAc/TEA 8:2:2%); m.p. 104 °C; ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ = 8.86 (d, ^3J(H,H) = 1.5 Hz, 1H; C6H), 8.65 (dd, ^3J(H,H) = 0.8, 4.8 Hz, 1H; C6H), 8.36 (d, ^3J(H,H) = 8.1 Hz, 1H; C3H), 8.20 (d, ^3J(H,H) = 7.8 Hz, 1H; C3H), 8.10 (dd, ^3J(H,H) = 2.2, 8.5 Hz, 1H; C4H), 7.80 (td, ^3J(H,H) = 1.8, 7.7 Hz, 1H; C4H), 7.31 (ddd, ^3J(H,H) = 1.1, 4.8, 5.9 Hz, 1H; C5H); ^{13}C NMR (75 MHz, $CDCl_3$, 25 °C): δ = 155.18 (s), 155.10 (d), 154.83 (s), 149.12 (d), 145.11 (d), 136.98 (d), 124.02 (d), 122.74 (d), 120.88 (d), 93.83 (s); MS (70 eV, EI): *m*/*z* (%): 282 (100) $[M]^+$, 155 (57), 128 (25), 78 (30); IR (KBr): $\tilde{\nu}$ = 1586, 1573, 1558, 1540, 1455, 1435, 1358, 998, 791, 633 cm^{-1}.

Compound 7a: *n*BuLi (1.6 M in hexane, 9.3 mL, 14.9 mmol) was added dropwise at −78 °C to a solution of 4-bromo-biphenyl (**3a**, 3.5 g, 15.0 mmol) in dry THF (130 mL). After 1 h at −78 °C, **2** (6.16 g, 14.71 mmol) in dry THF (20 mL) was added. The mixture was stirred for 3 h at −78 °C and allowed to warm up to RT over night. Sat. $NaHCO_3$ (300 mL) was added and the mixture was extracted with Et_2O (3 × 250 mL). The organic layer was washed with H_2O (100 mL), dried ($MgSO_4$) and concentrated in vacuo. The yellow residue was dissolved in dry CH_2Cl_2 (40 mL) and cooled to −78 °C. Et_3SiH (11.7 mL, 73.6 mmol) and $BF_3{\cdot}OEt_2$ (9.3 mL, 73.6 mmol) were added dropwise. The reaction mixture was allowed to warm to RT over night, quenched with HCl (1 M, 30 mL) and stirred for 30 min at RT. The mixture was neutralized with 2% NaOH and extracted with EtOAc (4 × 150 mL). The organic layer was washed with brine (50 mL), dried ($MgSO_4$) and concentrated in vacuo. Compound **7a** (2.93 g, 35%) was obtained after FC (hexane/EtOAc 9:1) as a slightly yellow solid. R_f = 0.15 (hexane/EtOAc 9:1); m.p. 80–87 °C; ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ = 7.65–7.63 (d, ^3J(H,H) = 2.7 Hz, 2H; ArH), 7.60–7.54 (m, 4H; ArH), 7.51–7.46 (m, 3H; ArH), 7.44–7.23 (m, 15H; ArH), 5.15 (d, ^3J(H,H) = 6.6 Hz, 1H; C1′H), 4.69–4.56 (m, 6H; CH_2), 4.44 (m, 1H; C4′H), 4.11 (t, ^3J(H,H) = 4.8 Hz, 1H; C3′H), 3.93 (t, ^3J(H,H) = 5.9 Hz, 1H; C2′H), 3.73 (ddd, $^{2,3}J$(H,H) = 4.0, 10.7, 14.7 Hz, 2H; C5′H); ^{1}H NMR NOE (400 MHz, $CDCl_3$, 25 °C): δ = 5.15 (C1′H) → 7.61 (ArH; 8.8%), 4.44 (C4′H; 3.0%), 3.93 (C2′H; 1.7%); 4.44 (C4′H) → 5.15 (C1′H; 3.4%), 4.11 (C3′H; 2.3%), 3.73 (C5′H; 2.7%); 4.11 (C3′H) → 3.93 (C2′H, 7.6%); ^{13}C NMR (75 MHz, $CDCl_3$, 25 °C): δ = 140.98 (s, ArC), 140.57 (s, ArC), 139.47 (s, ArC), 138.17, 137.96, 137.80 (3s, ArC), 128.73, 128.37, 128.28, 128.07, 127.75, 127.64, 127.60, 127.19, 127.06, 127.02, 126.72 (11d, ArC), 83.71 (d, C2′), 82.38 (d, C1′), 81.78 (d, C4′), 77.54 (d, C3′), 73.47, 72.25, 71.95 (3t, CH_2), 70.47 (t, C5′); MS (70 eV, EI): *m*/*z* (%): 556 (0.05) $[M]^+$, 465 (5), 448 (3), 357 (10), 153 (12), 107 (8), 91 (100), 77 (13).

Compound 7b: This compound was prepared as described for **7a**, from **3b** (6.0 g, 0.021 mol), *n*BuLi (13.3 mL) in THF (280 mL). Lactone **2** (8.35 g, 0.0199 mol) in THF (70 mL) was added. After extraction the yellow residue was dissolved in CH_2Cl_2 (54.5 mL), Et_3SiH (15.81 mL, 0.0995 mol) and $BF_3{\cdot}OEt_2$ (12.5 mL, 0.0995 mol) were added. Compound **7b** (3.51 g, 32%) was obtained after FC (hexane/EtOAc 1:1 followed by a second column hexane/EtOAc/TEA 8:2:2%) as a slightly yellow oil. R_f = 0.21 (hexane/EtOAc/TEA 9:1:2%); ^{1}H NMR (500 MHz, $CDCl_3$, 25 °C): δ = 8.73 (d, ^3J(H,H) = 2.7 Hz, 1H; ArH), 8.71 (ddd, ^3J(H,H) = 1.2, 2.4, 6.1 Hz, 1H; ArH), 8.44 (d, ^3J(H,H) = 9.9 Hz, 1H; ArH), 8.35 (d, ^3J(H,H) = 10.2 Hz, 1H; ArH), 7.87–7.83 (m, 2H; ArH), 7.40–7.30 (m, 10H; ArH), 7.26–7.24 (m, 3H; ArH), 7.20–7.18 (m, 2H; ArH), 7.18 (d, ^3J(H,H) = 3.0 Hz, 1H; ArH), 5.10 (d, ^3J(H,H) = 9.3 Hz, 1H; C1′H), 4.63–4.51 (m, 5H; CH_2), 4.44 (d, ^3J(H,H) = 15.0 Hz, 1H; CH_2), 4.40 (dd, ^3J(H,H) = 4.7, 8.9 Hz, 1H; C4′H), 4.05 (dd, ^3J(H,H) = 4.0, 6.4 Hz, 1H; C3′H), 3.85 (dd, ^3J(H,H) = 6.6, 9.4 Hz, 1H; C2′H), 3.68 (dd, $^{2,3}J$(H,H) = 5.0, 12.9 Hz, 1H; C5′H), 3.62 (dd, $^{2,3}J$(H,H) = 5.0, 12.9 Hz, 1H; C5′H); ^{1}H NMR NOE (500 MHz, $CDCl_3$, 25 °C): δ = 5.10 (C1′H) → 4.44 (C4′H; 3.0%); 4.05 (C3′H) → 3.85 (C2′H; 11%); 3.85 (C2′H) → 4.05 (C3′H, 10%); ^{13}C NMR (75 MHz, $CDCl_3$, 25 °C): δ = 155.94 (s, ArC), 155.54 (s, ArC), 149.08 (d, ArC), 147.47 (d, ArC), 137.89 (s, ArC), 137.72 (s, ArC), 137.36 (s, ArC), 136.87 (d, ArC), 136.04 (s, ArC), 134.75 (d, ArC), 128.38–127.58 (15d, ArC), 123.59 (d, ArC), 121.06 (d, ArC), 120.68 (d, ArC), 83.80 (d, C2′), 82.30 (d, C1′), 80.07 (d, C4′), 77.42 (d, C3′), 73.49, 72.50, 71.92 (3t, CH_2), 70.38 (t, C5′); LSIMS: *m*/*z* (%): 559 (100) $[M+H]^+$, 185 (37), 147 (32), 136 (45).

Compound 8a: A solution of **7a** (2.23 g, 4.0 mmol) in dry CH_2Cl_2 (40 mL) was cooled to −78 °C and treated with BBr_3 (1 M in CH_2Cl_2, 14.0 mL, 14.0 mmol). After 4 h at −78 °C the reaction mixture was quenched with MeOH (45 mL) and allowed to warm up to RT over night. The solvent was evaporated; the residue dissolved in MeOH (200 mL) and washed with hexane (3×40 mL). The aqueous layer was evaporated and residual water was removed by coevaporation from pyridine. The crude product (1.14 g) was dissolved in dry pyridine (44 mL) and 1,3-dichloro-1,1,3,3-tetraiso-propyldisiloxane (1.25 mL, 4.0 mmol) was added dropwise at 0 °C. After stirring for 5 h at RT the solvent was evaporated. The residue was dissolved in saturated aqueous $NaHCO_3$ (80 mL) and extracted with EtOAc (4×40 mL). The organic layer was dried ($MgSO_4$) and concentrated in vacuo.

Compound 8a (761 mg, 36% over 2 steps) was obtained after FC (hexane/EtOAc 9:1) as a slightly yellow oil. R_f=0.24 (hexane/EtOAc 8:2); ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ =7.62–7.58 (m, 4H; ArH), 7.53–7.43 (m, 4H; ArH), 7.38–7.33 (m, 1H; ArH), 4.91 (d, ^3J(H,H)= 3.7 Hz, 1H; C1′H), 4.43 (t, ^3J(H,H)=6.6 Hz, 1H; C3′H), 4.15–4.13 (m, 2H; C5′H), 4.09–4.05 (m, 1H; C4′H), 4.02 (dd, ^3J(H,H)=3.7, 5.9 Hz, 1H; C2′H), 3.20–2.80 (br, 1H; HO), 1.14–1.04 (m, 28H; *i*Pr); ^{13}C NMR (75 MHz, $CDCl_3$, 25 °C): δ =140.91 (s, ArC), 140.60 (s, ArC), 139.11 (s, ArC), 128.75 (d, ArC), 127.26 (s, ArC), 127.17 (d, ArC), 127.11 (d, ArC), 126.27 (d, ArC), 85.26 (d, C1′), 82.46 (d, C2′), 77.30 (d, C4′), 71.62 (d, C3′), 62.45 (t, C5′), 17.51, 17.40, 17.38, 17.33, 17.15, 17.11, 17.00 (7q, *i*Pr), 13.42, 13.20, 12.89, 12.65 (4d, CH-*i*Pr); MS (70 eV, EI): *m*/*z* (%): 528 (0.8) $[M]^+$, 485 (29), 467 (15), 455 (9), 395 (7), 235 (100), 205 (29), 152 (17), 77 (9).

Compound 8b: This compound was prepared as described for **8a**, from **7b** (552 mg, 0.99 mmol), BBr_3 (1 M in CH_2Cl_2, 3.45 mL, 3.45 mmol) in CH_2Cl_2 (10 mL). The reaction was quenched with MeOH (12 mL), the mixture evaporated and the residue dissolved in H_2O (50 mL) and washed with CH_2Cl_2 (1×20 mL). The aqueous layer was evaporated and dried by coevaporation from pyridine. To the crude product (285 mg) 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane (0.37 mL, 0.99 mmol) in dry pyridine (11 mL) was added. Compound **8b** (239 mg, 46% over two steps) was obtained after FC (EtOAc/hexane 3:8) as a slightly yellow oil. R_f= 0.26 (EtOAc/hexane 3:8); ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ =8.71 (d, ^3J(H,H)=2.2 Hz, 1H; ArH), 8.70–8.68 (m, 1H; ArH), 8.43–8.38 (m, 2H; ArH), 7.90 (dd, ^3J(H,H)=1.9, 8.1 Hz, 1H; ArH), 7.83 (dt, ^3J(H,H)= 1.8, 7.7 Hz, 1H; ArH), 7.32 (ddd, ^3J(H,H)=1.1, 4.8, 7.4 Hz, 1H; ArH), 4.92 (d, ^3J(H,H)=3.7 Hz, 1H; C1′H), 4.40 (t, ^3J(H,H)=7.8 Hz, 1H; C3′H), 4.16–4.06 (m, 3H; C4′H, C5′H), 3.99 (m, 1H; C2′H), 3.06 (d, ^3J(H,H)=4.1 Hz, 1H; (HO), 1.12–1.02 (m, 28H; *i*Pr); ^{1}H NMR NOE (400 MHz, $CDCl_3$, 25 °C): δ =4.92 (C1′H) → 8.70 (ArH; 6.4%), 7.90 (ArH; 2.7%), 4.06 (C4′H; 4.2%), 3.99 (C2′H; 2.6%); 4.40 (C3′H) → 8.71 (ArH; 1.1%), 7.90 (ArH; 2.2%), 3.99 (C2′H; 10.7%); 3.99 (C2′H) → 8.70 (ArH; 2.2%), 7.91 (ArH, 1.5%), 4.91 (C1′H; 3.2%), 4.41 (C3′H; 10.0%); 4.06 (C4′H) → 4.92 (C1′H; 3.6%), 4.40 (C3′H; 2.5%); ^{13}C NMR (101 MHz, $CDCl_3$, 25 °C): δ =155.73 (s, ArC), 155.42 (s, ArC), 148.95 (d, ArC), 146.91 (d, ArC), 136.81 (d, ArC), 135.48 (s, ArC), 134.27 (d, ArC), 123.53 (d, ArC), 121.01 (d, ArC), 120.63 (d, ArC), 83.09 (d, C1′), 82.70 (d, C2′), 76.53 (d, C4′), 71.57 (d, C3′), 62.29 (t, C5′), 17.31, 17.20, 17.18, 17.14, 16.94, 16.91, 16.89, 16.78 (8q, *i*Pr), 13.20, 13.01, 12.74, 12.46 (4d, CH-*i*Pr); LSIMS: *m*/*z* (%): 531 (100) $[M+H]^+$, 199 (19), 185 (28), 133 (14).

Compound 9a: 1,1′-Thiocarbonyl diimidazole (300 mg, 1.7 mmol) was added to a solution of **8a** (740 mg, 1.40 mmol) in dry acetonitrile (6.1 mL). After stirring for 9 h at RT a second portion of 1,1′-thiocarbonyl diimidazole (150 mg, 0.85 mmol) was added and stirred for another 9 h. The suspension was concentrated in vacuo, the residue dissolved in CH_2Cl_2 (100 mL) and washed with saturated aqueous $NaHCO_3$ solution (3×40 mL). The organic layer was dried ($MgSO_4$) and concentrated in vacuo. Compound **9a** (751 mg, 84%) was obtained after FC (hexane/EtOAc 7:3) as a white solid. R_f=0.41 (hexane/EtOAc 7:3); m.p. 112–116 °C; ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ =8.54 (s, 1H; ImH), 7.76 (s, 1H; ImH) 7.64–7.59 (m, 6H; ArH), 7.46 (t, ^3J(H,H)=7.35 Hz, 2H; ArH), 7.39–7.35 (m, 1H; ArH), 7.16 (s, 1H; ImH), 5.84 (d, ^3J(H,H)= 4.8 Hz, 1H; C2′H), 5.32 (s, 1H; C1′H), 4.71 (dd, ^3J(H,H)=4.8, 8.8 Hz, 1H; C3′), 4.33–4.28 (m, 1H; C5′H), 4.17–4.10 (m, 2H; C4′H, C5′H), 1.17–0.94 (m, 28H; *i*Pr); ^{13}C NMR (75 MHz, $CDCl_3$, 25 °C): δ =141.2, 140.68, 137.69, 136.81, 136.80 (5s, ArC), 130.81, 128.80, 127.44, 127.30, 127.12, 126.89, 126.43, 124.08, 118.10 (9d, ArC, ImC), 87.64 (d, C2′), 82.89 (d, C1′), 82.12 (d, C4′), 69.21 (d, C3′), 60.72 (t, C5′), 17.48, 17.36, 17.34, 17.27, 17.08, 16.98, 16.90 (7q, *i*Pr), 13.37, 13.05, 12.89, 12.67 (4d, CH-*i*Pr); LSIMS: *m*/*z* (%): 639 (38) $[M+H]^+$, 511 (10), 467 (8), 261 (75), 233 (100).

Compound 9b: This compound was prepared as described for **9a**, from **8b** (139 mg, 0.262 mmol), 1,1′-thiocarbonyl diimidazole (56 mg, 0.31 mmol) in acetonitrile (1.2 mL). The residue was dissolved in EtOAc (50 mL) and washed with saturated aqueous $NaHCO_3$ (3×40 mL). Compound **9b** (120 mg, 71%) was obtained after FC (EtOAc/hexane 4:1) as a white foam. R_f=0.25 (EtOAc/hexane 3:1); ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ =8.89 (d, ^3J(H,H)=1.8 Hz, 1H; ArH), 8.69 (d, ^3J(H,H)= 4.0 Hz, 1H; ArH), 8.45 (m, 1H; ImH), 8.43–8.42 (m, 2H; ArH), 8.07 (dd, ^3J(H,H)=1.8, 8.0 Hz, 1H; ArH), 7.84 (dd, ^3J(H,H)=1.8, 7.7 Hz, 1H; ArH), 7.71 (m, 1H; ImH), 7.35–7.31 (m, 1H; ArH), 7.09 (s, 1H; ImH), 5.78 (d, ^3J(H,H)=4.0 Hz, 1H; C2′H), 5.36 (s, 1H; C1′H), 4.67 (dd, ^3J(H,H)=5.2, 8.8 Hz, 1H; C3′H), 4.29 (m, 1H; C5′H), 4.18–4.09 (m, 2H; C4′H, C5′H), 1.11–0.92 (m, 28H; *i*Pr); ^{13}C NMR (75 MHz, $CDCl_3$, 25 °C): δ=183.35 (s, C=S), 156.17, 155.72 (2s, ArC), 149.16 (d, ArC), 147.27 (d, ArC), 136.87 (d, ImC), 136.77, 134.53, 134.29 (3d, ArC), 131.08 (d, ImC), 130.67, 123.79, 121.13 (3d, ArC), 118.04 (d, ImC), 87.08 (d, C2′), 82.17 (d, C1′), 81.09 (d, C4′), 69.16 (d, C3′), 60.53 (t, C5′), 17.41, 17.30, 17.27, 17.22, 17.01, 16.92, 16.88, 16.84 (8q, *i*Pr), 13.26, 12.97, 12.83, 12.61 (4d, CH-*i*Pr); LSIMS: *m*/*z* (%): 641 (10) $[M+H]^+$, 531 (15), 261 (18), 235 (40), 278 (64), 252 (66), 215 (40), 181 (28), 160 (38), 105 (18).

Compound 10a: A solution of **9a** (700 mg, 1.1 mmol) in toluene (7.2 mL) was heated at 80 °C, and a solution of Bu_3SnH (0.58 mL, 2.19 mmol) and α,α′-azobisisobutyronitrile (34 mg, 0.21 mmol) in toluene (18 mL) was added dropwise over a period of 2 h. The reaction mixture was stirred for 2 h at 80 °C and then allowed to cool to RT. The mixture was concentrated in vacuo. Compound **10a** (438 mg, 78%) was obtained after FC (hexane/EtOAc 15:0.3) as a colorless oil. R_f=0.61 (hexane/EtOAc 8:2); ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ =7.62–7.58 (m, 4H; ArH), 7.48–7.34 (m, 5H; ArH), 5.17 (t, ^3J(H,H)=7.35 Hz, 1H; C1′H), 4.60 (m, 1H; C3′H), 4.19 (m, 1H; C5′H), 3.96–3.90 (m, 2H; C4′H, C5′H), 2.48–2.40 (m, 1H; C2′H), 2.20–2.11 (m, 1H; C2′H), 1.14–1.07 (m, 28H; *i*Pr); ^{13}C NMR (75 MHz, $CDCl_3$, 25 °C): δ =141.09, 140.93, 140.44 (3s, ArC), 128.71, 127.19, 127.11, 127.06, 126.29 (5d, ArC), 86.46 (d, C4′), 78.84 (d, C1′), 73.43 (d, C3′), 63.76 (t, C5′), 43.11 (t, C2′), 17.59, 17.46, 17.44, 17.39, 17.26, 17.11, 17.08, 16.99 (8q, *i*Pr), 13.52, 13.40, 13.03, 12.57 (4d, CH-*i*Pr); MS (70 eV, EI): *m*/*z* (%): 512 (0.02) $[M]^+$, 469 (86), 451 (19), 439 (5), 235 (100).

Compound 10b: This compound was prepared as described for **10a**. Bu_3SnH (80 μL, 0.30 mmol) and α,α′-azobisisobutyronitrile (4.6 mg, 28 mol) in toluene (0.4 mL) were added to **9b** (119 mg, 0.186 mmol) in toluene (2 mL). Compound **10b** (69 mg, 72%) was obtained after FC (EtOAc/hexane/TEA 1:3:0.05) as a slightly yellow oil. R_f=0.36 (EtOAc/hexane 1:3); ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ =8.67 (d, ^3J(H,H)= 4.0 Hz, 1H; ArH), 8.62 (d, ^3J(H,H)=1.5 Hz, 1H; ArH), 8.39 (d, ^3J(H,H)=4.8 Hz, 1H; ArH), 8.36 (d, ^3J(H,H)=5.2 Hz, 1H; ArH), 7.84–7.77 (m, 2H, ArH), 7.31–7.27 (m, 1H; ArH), 5.18 (t, ^3J(H,H)=7.4 Hz, 1H; C1′H), 4.56 (m, 1H; C3′H), 4.15 (m, 1H; C5′H), 3.97–3.87 (m, 2H; C4′H, C5′H) 2.45 (ddd, $^{2,3}J$(H,H)=4.8, 7.0, 12.5 Hz, 1H; C2′H), 2.15–2.03 (m, 1H; C2′H), 1.09–1.03 (m, 28H; *i*Pr); ^{13}C NMR (75 MHz, $CDCl_3$, 25 °C): δ =156.23, 155.72 (2s, ArC), 149.35, 147.33, 137.85, 137.05, 134.55, 123.81, 121.24, 120.96 (8d, ArC), 86.81 (d, C4′), 77.62 (d, C1′), 73.47 (d, C3′), 63.79 (t, C5′), 43.21 (t, C2′), 17.76, 17.64, 17.61, 17.56, 17.43, 17.27, 17.16, 16.06 (8q, *i*Pr), 13.80, 13.69, 13.58, 13.24 (4d, CH-*i*Pr); LSIMS: *m*/*z* (%): 515 (100) $[M+H]^+$, 183 (20).

Compound 1a: $NEt_3{\cdot}3HF$ (1.0 mL, 6.3 mmol) was added dropwise to a solution of **10a** (323 mg, 0.63 mmol) in dry THF (32 mL). After stirring for 15 h at RT the reaction mixture was concentrated in vacuo. Compound **1a** (167 mg, 98%) was obtained after FC (EtOAc/toluene/MeOH 5:5:0.3) as a white solid. R_f=0.37 (EtOAc/toluene/MeOH 5:5:1); ^{1}H NMR (300 MHz, [D]methanol, 25 °C): δ=7.65–7.60 (m, 4H; ArH),

7.51–7.33 (m, 5H; ArH), 5.21 (dd, ^3J(H,H) = 5.2, 10.7 Hz, 1H; C1′H), 4.40–4.38 (m, 1H; C3′H), 4.05–4.00 (m, 1H; C4′H), 3.75 (m, 2H; C5′H), 2.30–2.23 (ddd, ^3J(H,H) = 1.5, 5.5, 13.3 Hz, 1H; C2′H), 2.07–1.96 (m, 1H; C2′H); ^{13}C NMR (75 MHz, [D]methanol, 25 °C) δ = 142.57, 142.42, 142.06 (3s, ArC), 130.13, 128.59, 128.19, 127.97 (4d, ArC), 89.49 (d, C4′), 81.65 (d, C1′), 74.74 (d, C3′), 64.36 (t, C5′), 45.22 (t, C2′); MS (70 eV, EI): *m*/*z* (%): 270 (53) $[M]^+$, 252 (4), 234 (10), 180 (83), 167 (100), 152 (44), 115 (18), 77 (19); IR (KBr): $\tilde{\nu}$ = 3394, 2927, 1486, 1084, 1071, 1045, 1008, 830, 764, 701 cm^{-1}.

Compound 1b: This compound was prepared as described for **1a**, from **10b** (61 mg, 0.12 mmol) and NEt_3·3HF (0.20 mL, 1.2 mmol) in THF (7 mL). Compound **1b** (31 mg, 95%) was obtained after FC (EtOAc/toluene/MeOH/TEA 5:5:3:2%) as a slightly pink oil. R_f = 0.48 (EtOAc/toluene/MeOH/TEA 5:5:3:2%); ^{1}H NMR (300 MHz, [D]methanol, 25 °C): δ = 8.72–8.69 (m, 2H; ArH), 8.33–8.31 (m, 2H; ArH), 8.06–7.95 (m, 2H; ArH), 7.48 (m, 1H; ArH), 5.29 (dd, ^3J(H,H) = 5.5, 10.7 Hz, 1H; C1′H), 4.43–4.41 (m, 1H; C3′H), 4.00 (m, 1H; C4′H), 3.75 (d, ^3J(H,H) = 5.2 Hz, 2H; C5′H), 2.34 (dd, ^3J(H,H) = 5.5, 13.23 Hz, 1H; C2′H), 2.11–1.96 (m, 1H; C2′H); ^{13}C NMR (75 MHz, [D]methanol, 25 °C): δ = 157.23, 156.67 (2s, ArC), 150.52, 148.67 (2d, ArC), 140.04 (s, ArC), 139.04, 136.77, 125.58, 122.93, 122.59 (5d, ArC), 89.77 (d, C4′), 79.41 (d, C1′), 74.69 (d, C3′), 64.22 (t, C5′), 45.10 (t, C2′); LSIMS: *m*/*z* (%): 273 (15) $[M+H]^+$, 155 (100), 135 (18), 119 (89).

Compound 11a: 4,4′-Dimethoxytrityl chloride (220 mg, 0.65 mmol) was added in three portions to a solution of **1a** (146 mg, 0.54 mmol) in dry pyridine (2.2 mL). After 6 h, toluene (2 mL) was added and the reaction mixture was evaporated. The residue was dissolved in CH_2Cl_2 (30 mL) and washed with saturated aqueous $NaHCO_3$ (3×15 mL), dried ($MgSO_4$) and concentrated in vacuo. Compound **11a** (240 mg, 85%) was obtained after FC (EtOAc/toluene/TEA 3:6:0.3) as a white foam. R_f = 0.37 (EtOAc/toluene/TEA 3:6:0.2); ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ = 7.62–7.19 (m, 18H; ArH), 6.87 (s, 2H; ArH), 6.84 (s, 2H; ArH), 5.24 (dd, ^3J(H,H) = 5.5, 9.9 Hz, 1H; C1′H), 4.47 (m, 1H; C3′H), 4.10 (m, 1H; C4′H), 3.80 (s, 6H; OCH_3), 3.40 (dd, ^3J(H,H) = 4.8, 9.9 Hz, 1H; C5′H), 3.20 (dd, ^3J(H,H) = 5.5, 9.9 Hz, 1H; C5′H), 2.38 (s, 1H; OH), 2.30 (ddd, $^{2,3}J$(H,H) = 2.2, 5.9, 7.7 Hz, 1H; C2′H), 2.16–2.09 (m, 1H; C2′H); ^{1}H NMR NOE (400 MHz, $CDCl_3$, 25 °C): δ = 5.24 (C1′H) → 7.4 (ArH; 7.6%), 4.10 (C4′H; 3.5%), 2.30 (C2′H; 5.5%); 4.47 (C3′H) → 4.10 (C4′H; 2.7%), 3.40 (C5′H; 1.2%), 3.20 (C5′H; 1.4%); 4.10 (C4′H) → 5.24 (C1′H; 4.2%), 4.47 (C3′H; 2.4%), 3.40 (C5′H; 3.0%), 3.20 (C5′H; 1.8%); 2.30 (C2′H) → 5.24 (C1′H); 10.3%), 4.47 (C3′H; 2.5%); ^{13}C NMR (75 MHz, $CDCl_3$, 25 °C): δ = 158.47, 144.86, 140.93, 140.85, 140.50 (5s, ArC), 136.04, 130.13, 129.05, 128.77, 128.24, 128.21, 127.87, 127.25, 127.13, 127.10, 126.82, 126.51, 113.14 (13d, ArC), 86.34 (s, ArC), 86.24 (d, C4′), 79.80 (d, C1′), 74.76 (d, C3′), 64.51 (t, C5′), 55.22 (q, OCH_3), 43.84 (t, C2′); LSIMS: *m*/*z* (%): 572 (4) $[M+H]^+$, 303 (100).

Compound 11b: This compound was prepared as described for **11a**, from **1b** (30 mg, 0.11 mmol), 4,4′-dimethoxytrityl chloride (49 mg, 0.14 mmol) in dry pyridine (0.4 mL). Compound **11b** (48 mg, 75%) was obtained after FC (EtOAc/toluene/MeOH/TEA 1:1:0.03:0.03) as a slightly yellow foam. R_f = 0.4 (EtOAc/toluene/MeOH/TEA 1:1:0.03:0.03); ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ = 8.59 (m, 2H; ArH), 8.30 (m, 2H; ArH), 7.82–7.70 (m, 2H; ArH), 7.41–7.39 (m, 2H; ArH), 7.30–7.09 (m, 12H; ArH), 6.77 (s, 2H; ArH), 6.74 (s, 2H; ArH), 5.19 (dd, ^3J(H,H) = 5.5, 10.3 Hz, C1′H), 4.39 (m, 1H, C3′H), 4.07 (m, 1H; C4′H), 3.78 (s, 6H; OCH_3), 3.32–3.20 (m, 2H; C5′H), 2.27–2.21 (m, 1H; C2′H) 2.11–2.01 (m, 1H; C2′H); ^{13}C NMR (75 MHz, $CDCl_3$, 25 °C): δ = 158.46 (s, ArC), 155.93, 155.39 (2s, ArC), 149.08, 147.20 (2d, ArC), 144.72 (2s, ArC), 137.50 (s, ArC), 136.88 (d, ArC), 135.92 (s, ArC), 134.56 (d, ArC), 130.02, 128.13, 127.80, 126.78 (4d, ArC) 123.60, 121.06, 120.79 (3d, ArC), 113.11 (d, ArC), 86.56 (s, ArC), 86.23 (d, C4′H), 78.97 (d, C1′H), 74.36 (d, C3′), 64.36 (t, C5′), 55.14 (q, CH_3), 43.68 (t, C2′); HRMS, LSI-MS: *m*/*z*: calcd for $C_{36}H_{35}N_2O_5$: 575.25537; found: 575.25460 $[M+H]^+$.

Compound 12a: *N*,*N*-Diisopropylethylamine (213 μL, 1.24 mmol) was added followed by 2-cyanoethyl diisopropylchlorophosphor amidite (138 μL, 0.62 mmol) at RT to a solution of **11a** (237 mg, 0.41 mmol) in dry THF (10.6 mL). After 1.5 h CH_2Cl_2 (40 mL) was added, and the mixture was extracted with saturated aqueous $NaHCO_3$ (3×20 mL), dried ($NaSO_4$) and concentrated in vacuo. Compound **12a** (284 mg, 89%) was obtained after FC (toluene/EtOAc/TEA 8:2:0.2) as a white foam. R_f = 0.72 (toluene/EtOAc/TEA 8:2:0.2); ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ = 7.61–7.18 (m, 18H; ArH), 6.84 (d, ^3J(H,H) = 3.7 Hz, 2H; ArH), 6.81 (d, ^3J(H,H) = 3.67 Hz, 2H; ArH), 5.22 (m, 1H; C1′H), 4.55 (dd, ^3J(H,H) = 5.9, 10.7 Hz, 1H; C4′H), 4.27 (m, 1H; C3′H), 3.84–3.71 (m, 2H; CH_2CO), 3.79 (s, 6H; OCH_3), 3.65–3.52 (m, 2H; CH-*i*Pr), 3.41–3.23 (m, 2H; C5′H), 2.65 (t, ^3J(H,H) = 6.30 Hz, 1H; CH_2CN), 2.50 (t, ^3J(H,H) = 6.30, 1H; CH_2CN), 2.45–2.39 (m, 1H; C2′H), 2.13 (m, 1H; C2′H), 1.22–1.09 (m, 12H; CH_3-*i*Pr); ^{13}C NMR (75 MHz, $CDCl_3$, 25 °C): δ = 158.39, 144.87, 140.90, 140.87, 140.75, 140.72, 136.09, 136.03 (8s, ArC), 130.12, 130.09, 128.68, 128.26, 128.22, 127.72, 127.15, 127.02, 126.70, 126.66, 126.49 (11d, ArC), 177.48, 117.42 (2s, CN), 113.03 (d, ArC), 86.04 (s, ArC), 85.98, 85.72, 85.64 (3d, C4′H), 80.08, 80.01 (2d, C1′), 76.34, 76.11, 75.92, 75.70 (4d, C3′), 64.18, 64.13 (2t, C5′), 58.42, 58.39, 58.17, 58.14 (4t, CH_2), 55.14, 55.13 (2q, OCH_3), 43.25, 43.21 (2t, C2′), 43.08, 43.04 (2d, CH-*i*Pr), 24.64, 24.52, 24.46, 24.36 (4q, CH_3-*i*Pr), 20.34, 20.25, 20.19, 20.09 (4t, CH_2CN); ^{31}P NMR (202 MHz, $CDCl_3$, 25 °C): δ = 149.30, 149.09; HRMS, LSI-MS: *m*/*z*: calcd for $C_{47}H_{54}N_2O_6P_1$: 773.3719; found: 773.3687 $[M+H]^+$.

Compound 12b: This compound was prepared as described for **12a**, from **11b** (30 mg, 0.06 mmol), *N*,*N*-diisopropylethylamine (30 μL, 0.17 mmol) and 2-cyanoethyl diisopropylchlorophosphoramidite (20 μL, 0.09 mmol) in dry THF (1.6 mL). Compound **12b** (28 mg, 69%) was obtained after FC (EtOAc/hexane/TEA 1:1:2%) as a slightly yellow foam. R_f = 0.45, 0.53 (EtOAc/hexane/TEA 1:1:2%); ^{1}H NMR (300 MHz, $CDCl_3$, 25 °C): δ = 8.69–8.67 (m, 2H; ArH), 8.42–8.36 (m, 2H; ArH), 7.89 (dt, ^3J(H,H) = 2.9, 8.1 Hz, 1H; ArH), 7.82 (td, ^3J(H,H) = 1.5, 7.7 Hz, 1H; ArH), 7.48 (m, 2H; ArH), 7.38–7.20 (m, 10H; ArH), 6.85–6.81 (m, 4H; ArH), 5.27 (dd, ^3J(H,H) = 4.8, 10.3 Hz, 1H; C1′H), 4.60–4.55 (m, 1H; C4′), 4.29 (m, 1H; C(3′H), 3.81–3.64 (m, 8H; CH_2CO, OCH_3), 3.62–3.59 (m, 2H; CH-*i*Pr), 3.35–3.27 (m, 2H; C5′H), 2.63 (t, ^3J(H,H) = 6.3, 1H; CH_2CN), 2.51–2.45 (m, 1H; CH_2CN) 2.11–2.05 (m, 1H; C2′H), 1.75–1.65 (m, 1H; C2′H), 1.22–1.10 (m, 12H; CH_3-*i*Pr); ^{13}C NMR (75 MHz, $CDCl_3$, 25 °C): δ = 158.48 (s, ArC), 156.04, 155.57 (2s, ArC), 149.15, 147.32 (2d, ArC), 144.75 (2s, ArC), 137.28 (s, ArC), 136.85 (d, ArC), 136.00, 135.97 (2s, ArC), 134.58 (d, ArC), 130.1, 128.13, 128.25, 128.2, 127.80, 126.78 (5d, ArC), 123.61, 121.04, 120.76 (3d, ArC), 117.47, 117.41 (2s, CN), 113.10 (d, ArC), 86.23, 86.18 (2d, C4′H), 85.95, 85.90 (2s, 2H; ArC), 78.08, 78.04 (2d, C1′), 76.25, 76.11, 75.80, 75.66 (4d, C3′), 64.08, 64.03 (t, C5′), 58.38, 58.23 (2t, CH_2OP), 55.18 (q, OCH_3), 43.29, 43.24 (2t, C2′), 43.19, 43.07 (2d, CH-*i*Pr), 24.62, 24.56, 24.49, 24.44 (4q, CH_3-*i*Pr), 20.39, 20.34, 20.22, 20.16 (4t, CH_2-CN); ^{31}P NMR (202 MHz, $CDCl_3$, 25 °C): δ = 148.57, 148.44; HRMS, LSI-MS: *m*/*z*: calcd for $C_{45}H_{52}N_4O_6P_1$: 775.3625; found: 775.3612 $[M+H]^+$.

Synthesis and purification of the oligonucleotides: All oligonucleotides were synthesized on a 1 μmol scale on an Applied Biosystems Expedite Nucleic Acid Synthesizer (8909) using standard phosphoramidite chemistry. The phosphoramidites of the natural nucleosides as well as the nucleoside derived CPG solid supports were purchased from Glen Research. The universal support was purchased from CT-Gen (San Jose). The solvents and reagents used for the synthesis were prepared according to the manufacturer's indications in the trityl-off mode. The coupling time for the modified phosphoramidites was extended to 6 min and 2-ethylthio-1*H*-tetrazole was used as activator. After synthesis, the oligonucleotides were detached and deprotected in concentrated aqueous ammonia (12–18 h at 55 °C) and filtered through Titan filters (Teflon, 0.45 μm, Infochroma AB). HPLC was performed on an Äkta Basic 10/100 system (Amersham Pharmacia Biotech). All modified oligonucleotides were characterized by ESI$^-$-mass spectrometry (Table 6). Concentrations of oligonucleotides were determined by UV absorption. For the modified bases, extinction coefficients $\varepsilon_{260\,nm}$ = 7100 (bipy) and $\varepsilon_{260\,nm}$ = 19000 (biph) were used.

UV-melting experiments and CD spectra: All UV melting curves were recorded on a Cary 3E UV/VIS spectrometer (Varian) equipped with a Peltier block and Varian WinUV software at 260 nm. Unless otherwise indicated, the oligonucleotide concentration was kept at 1.2 μM in a buffer solution (10 mM NaH_2PO_4, 0.15 mM NaCl, pH 7.0) throughout all

Table 6. MS data and HPLC or FPLC purification of modified oligonucleotides.

Entry	Sequence	Calcd $[M-H]^-$	Found $[M-H]^-$	HPLC Purification AE[a] [%]	t_R [min]	RP[c] [%]	t_R [min]	$OD_{260\,nm}$
1	5′-CTAGC**HH**GTCATC-3′	3652.42	3652.76	40–55 B[1*]	17.3	10–25 B	18.7	31.4
2	3′-GATCG**HH**CAGTAG-5′	3741.48	3741.86	40–70 B[1*]	18.0	10–25 B	19.5	13.7
3	5′-CTAGC**TH**GTCATC-3′	3776.52	3776.25	40–55 B[1*]	17.9	10–25 B	16.4	30.7
4	3′-GATCG**HT**CAGTAG-5′	3865.58	3865.13	40–55 B[1*]	17.5	10–25 B	17.3	31.7
5	5′-CTAGC**GH**GTCATC-3′	3801.53	3800.75	40–55 B[1*]	19.2	10–25 B	16.1	22.1
6	3′-GATCG**HG**CAGTAG-5′	3890.59	3889.88	40–60 B[1*]	19.9	10–25 B	17.0	30.8
7	5′-CTAGC**IH**GTCATC-3′	3804.61	3804.89	20–45 B[1[b]]	23.0	10–25 B	24.2	16.1
8	3′-GATCG**HI**CAGTAG-5′	3893.67	3893.00	20–45 B[1[b]]	24.2	10–25 B	25.3	25.4
9	5′-CTAGC**YH**GTCATC-3′	3806.47	3806.74	40–70 B[1*]	23.2	5–30 B	21.0	16.9
10	3′-GATCG**HY**CAGTAG-5′	3895.52	3895.38	40–70 B[1*]	24.6	5–30 B	21.2	32.0
11	5′-CTAGC**PH**GTCATC-3′	3852.65	3852.90	20–45 B[1[b]]	25.0	10–25 B	24.8	25.8
12	3′-GATCG**HP**CAGTAG-5′	3941.71	3941.65	20–45 B[1[b]]	25.8	10–25 B	25.4	24.2
13	5′-CTAG**HIHI**TCATC-3′	3698.60	3698.70	60–82 B[1]	19.5	10–35 B	26.9	35.7
14	3′-GATC**IHIH**AGTAG-5′	3787.66	3787.00	60–82 B[1]	23.5	10–35 B	27.8	11.1
15	5′-CTAG**HYHY**TCATC-3′	3702.32	3701.63	30–65 B[1]	26.0	10–25 B	23.4	20.7
16	3′-GATC**YHYH**AGTAG-5′	3791.38	3790.75	50–80 B[1]	29.3	10–25 B	23.7	19.6
17	5′-CTAG**HPHP**TCATC-3′	3794.68	3793.75	40–80 B[1]	16.6	10–35 B	27.4	24.2
18	3′-GATC**PHPH**AGTAG-5′	3883.74	3882.88	60–95 B[1]	24.5	10–35 B	26.7	29.9
19	5′-CTAGC**I**GTCATC-3′	3624.51	3624.00	25–65 B[2*]	23.2	10–25 B	23.8	26.0
20	3′-GATCG**I**CAGTAG-5′	3713.57	3713.25	20–60 B[2*]	28.1	10–25 B	24.4	19.1
21	5′-CTAGC**II**GTCATC-3′	3956.80	3955.88	60–75 B[1]	16.5	10–35 B	25.6	38.6
22	3′-GATCG**II**CAGTAG-5′	4045.86	4045.25	40–95 B[1*]	19.2	10–40 B	18.4	25.9
23	5′-CTAGC**III**GTCATC-3′	4289.09	4288.25	50–90 B[1*]	27.5	10–40 B	27.4	28.6
24	3′-GATCG**III**CAGTAG-5′	4378.15	4377.25	50–90 B[1*]	27.5	10–45 B	24.8	18.6
25	5′-CTAGC**IIII**GTCATC-3′	4621.38	4621.63	60–95 B[1*]	23.0	20–55 B	23.8	37.1
26	3′-GATCG**IIII**CAGTAG-5′	4710.44	4710.63	60–95 B[1*]	25.9	20–55 B	23.8	13.5
27	5′-CTAGC**IIIII**GTCATC-3′	4953.67	4953.25	–	–	20–55 B[*]	24.0	30.0
28	3′-GATCG**IIIII**CAGTAG-5′	5042.73	5042.75	–	–	20–55 B[*]	24.8	36.3
29	5′-CTAGC**IIIIII**GTCATC-3′	5285.96	5285.00	–	–	20–55 B[*]	28.9	22.1
30	3′-GATCG**IIIIII**CAGTAG-5′	5375.02	5374.13	–	–	30–65 B[*]	20.0	18.8
31	5′-CTAGC**IIIIIII**GTCATC-3′	5618.25	5617.25	–	–	30–65 B[*]	24.0	11.2
32	3′-GATCG**IIIIIII**CAGTAG-5′	5707.31	5706.13	–	–	30–65 B[*]	24.0	24.1
33	5′-TTTTTTTT**IIIIIIIIII**T-3′	6663.30	6661.75	–	–	30–65 B[*]	25.0	29.9
34	3′-T**IIIIIIIIII**T-5′	4533.92	4532.13	–	–	30–65 B[*]	25.3	16.5
35	5′-**IIIIII**GTCATC-3′	4093.26	4092.75	–	–	20–80 B[*]	23.9	30.4
36	3′-**IIIIII**CAGTAG-5′	4142.29	4341.5	–	–	40–70 B	19.2	36.1
37	5′-CTA**IIIIII**ATC-3′	3744.97	3744.75	–	–	20–100 B[*]	19.0	43.0
38	3′-GAT**IIIIII**TAG-5′	3825.02	3824.13	–	–	20–80 B[*]	23.0	32.9
39	5′-CTAGC**Y**GTCATC-3′	3626.37	3626.25	40–70 B[1*]	21.5	5–30 B	21.0	n.d.
40	3′-GATCG**Y**CAGTAG-5′	3715.43	3715.25	40–70 B[1*]	21.0	5–30 B	21.5	n.d.
41	5′-CTAGC**YY**GTCATC-3′	3969.52	3959.88	50–70 B[1]	16.0	10–35 B	16.0	28.8
42	3′-GATCG**YY**CAGTAG-5′	4049.58	4049.00	40–60 B[1*]	20.7	10–20 B	13.7	13.8
43	5′-CTAGC**YYY**GTCATC-3′	4294.67	4294.38	30–70 B[2*]	22.7	10–25 B	14.2	12.8
44	3′-GATCG**YYY**CAGTAG-5′	4383.74	4383.50	40–70 B[2*]	22.4	10–25 B	14.3	15.2
45	5′-CTAGC**YYYY**GTCATC-3′	4628.82	4628.38	40–95 B[2*]	16.2	10–25 B	15.3	21.0
46	3′-GATCG**YYYY**CAGTAG-5′	4717.89	4718.00	40–80 B[2*]	17.0	10–25 B	15.6	10.7
47	5′-CTAGC**YYYYY**GTCATC-3′	4962.98	4962.63	–	–	5–30 B[*]	24.5	12.0
48	3′-GATCG**YYYYY**CAGTAG-5′	5052.04	5051.63	–	–	5–30 B[*]	24.6	14.0
49	5′-CTAGC**YYYYYY**GTCATC-3′	5297.13	5296.38	–	–	5–40 B[*]	22.5	20.7
50	3′-GATCG**YYYYYY**CAGTAG-5′	5386.19	5385.13	–	–	5–40 B[*]	23.0	14.1
51	5′-CTAGC**IY**GTCATC-3′	3958.66	3957.88	40–80 B[2*]	20.2	10–25 B	20.1	12.7
52	3′-GATCG**IY**CAGTAG-5′	4047.72	4046.88	40–80 B[2*]	22.0	10–25 B	20.6	10.1
53	5′-CTAGC**YI**GTCATC-3′	3958.66	3958.00	40–80 B[2*]	20.6	10–25 B	20.2	19.8

[a] AE = Anion-exchange chromatography: Nucleogen DEAE 60–7, 125 × 4 mm, with Nucleogen-Guard column 30 × 4 mm (Machery–Nagel)for HPLC; solvent A^1 = 20 mM KH_2PO_4, pH 6, in H_2O/MeCN 8:2, and solvent B^1 = 20 mM KH_2PO_4, 1 M KCl in H_2O/MeCN 8:2; solvent B^2 = 20 mM KH_2PO_4, 1 M KCl in H_2O/MeCN 7:3. [b] Mono Q HR 5/5: Pharmacia Biotech for FPLC; solvent A^1 and B^1. [c] RP = reversed-phase chromatography: Aquapore RP-300, 220 × 6 mm (7 μm) with RP-C18 Newguard (7 μm) (Brownlee Labs); solvent A = 0.10 M $(Et_3NH)OAc$ in H_2O and solvent B = 0.10 M $(Et_3NH)OAc$ in H_2O/MeCN 1:4. All chromatographic purification were done at a flow rate of 1 mL min^{-1} and at rt or heated to 60 °C when indicated with *, and compounds were detected by UV at 260 nm.

measurements. Consecutive heating-cooling-heating cycles in the temperature interval of 0 or 90 °C or 10 to 90 °C were applied with a linear gradient of 0.5 °C min^{-1}. Heating and cooling ramps were in most cases superimposable, exceptions are indicated. T_m values were defined as the maximum of the first derivative of the melting curve. CD spectra were measured on a JASCO J-715 spectropolarimeter at the temperature indicated.

NMR experiments: Temperature dependent ^{1}H NMR spectra of the duplex 5′d(GATGACIGCTAG)-d(CTAGCIGTCATC) were recorded on a Bruker DRX 500 spectrometer at 500.13 MHz. The residual H_2O was suppressed by CW presaturation. Proton chemical shifts were referenced to sodium 3-(trimethylsilyl) 2,2,3,3-[D_4]propionate (TSP). NMR data were processed and analyzed using Bruker software (WinNMR, Version 6.0). The duplex was dissolved in 10 mM NaH_2PO_4, 150 mM NaCl at pH 7 and repeatedly freeze-dried from D_2O prior to spectra acquisition. Final addition of 100% D_2O resulted in a sample-concentration of about 1 mM.

Gel electrophoresis: 20% Nondenaturing polyacrylamide gels (19:1 mono/bis) were prepared according to standard procedures.[36] DNA samples (900 pmol) were dried on a Speedvac at low temperature for 2 h, taken up in the loading buffer (90 mM Tris-borat, 8% saccharose) and prior to loading on the gel heated up to 90 °C and then slowly cooled to 4 °C. The gels were run at 4 °C and 100 V for 16 h in TBE buffer (90 mM Tris-borat, pH 7.2). Bands were visualized by UV light.

[1] C. R. Cantor, P. R. Schimmel, in *Biophysical Chemistry Part III: The behavior of biological macromolecules,Vol. 2* (Ed.: P. C. Vapnek), W. H. Freeman, New York, **1980**, pp. 1109–1181.
[2] W. Saenger, *Principles of Nucleic Acid Structure*, Springer, New York, **1984**.
[3] K. M. Guckian, B. A. Schweitzer, R. X.-F. Ren, C. J. Sheils, P. L. Paris, D. C. Tahmassebi, E. T. Kool, *J. Am. Chem. Soc.* **1996**, *118*, 8182–8183.
[4] K. M. Guckian, B. A. Schweitzer, R. X.-F. Ren, C. J. Sheils, P. L. Paris, D. C. Tahmassebi, E. T. Kool, *J. Am. Chem. Soc.* **2000**, *112*, 2213–2222.
[5] B. A. Schweitzer, E. T. Kool, *J. Am. Chem. Soc.* **1995**, *117*, 1863–1872.
[6] T. J. Matray, E. T. Kool, *J. Am. Chem. Soc.* **1998**, *120*, 6191–6192.
[7] S. Moran, R. X.-F. Ren, E. T. Kool, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 10506–10511.
[8] K. M. Guckian, J. C. Morales, E. T. Kool, *J. Org. Chem.* **1998**, *63*, 9652–9656.
[9] J. C. Morales, E. T. Kool, *Nat. Struct. Biol.* **1998**, *5*, 950–954.
[10] J. C. Delaney, P. T. Henderson, S. A. Helquist, J. C. Morales, J. M. Essigmann, E. T. Kool, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 4469–4473.
[11] J. C. Morales, K. M. Guckian, E. T. Kool, *Angew. Chem.* **2000**, *112*, 1046–1068; *Angew. Chem. Int. Ed.* **2000**, *39*, 990–1009.
[12] T. J. Matray, E. T. Kool, *Nature* **1999**, *399*, 704–708.
[13] K. M. Guckian, T. R. Krugh, E. T. Kool, *Nat. Struct. Biol.* **1998**, *5*, 954–959.
[14] M. Berger, A. K. Ogawa, D. L. McMinn, Y. Q. Wu, P. G. Schultz, F. E. Romesberg, *Angew. Chem.* **2000**, *112*, 3069–3071; *Angew. Chem. Int. Ed.* **2000**, *39*, 2940–2942.
[15] M. Berger, S. D. Luzzi, A. A. Henry, F. E. Romesberg, *J. Am. Chem. Soc.* **2002**, *124*, 1222–1226.
[16] D. L. McMinn, A. K. Ogawa, Y. Wu, J. Liu, P. G. Schultz, F. E. Romesberg, *J. Am. Chem. Soc.* **1999**, *121*, 11585–11586.
[17] A. K. Ogawa, Y. Q. Wu, D. L. McMinn, J. Liu, P. G. Schultz, F. E. Romesberg, *J. Am. Chem. Soc.* **2000**, *122*, 3274–3287.
[18] E. L. Tae, Y. Wu, G. Xia, P. G. Schultz, F. E. Romesberg, *J. Am. Chem. Soc.* **2001**, *123*, 7439–7440.
[19] C. Yu, A. A. Henry, F. E. Romesberg, P. G. Schultz, *Angew. Chem.* **2002**, *114*, 3997–4000; *Angew. Chem. Int. Ed.* **2002**, *41*, 3841–3844.
[20] S. Atwell, E. Meggers, G. Spraggon, P. G. Schultz, *J. Am. Chem. Soc.* **2001**, *123*, 12364–12367.
[21] Y. Wu, A. K. Ogawa, M. Berger, D. L. McMinn, P. G. Schultz, F. E. Romesberg, *J. Am. Chem. Soc.* **2000**, *122*, 7621–7632.
[22] S. Matsuda, A. A. Henry, P. G. Schultz, F. E. Romesberg, *J. Am. Chem. Soc.* **2003**, *125*, 6134–6139.
[23] C. Brotschi, A. Häberli, C. J. Leumann, *Angew. Chem.* **2001**, *113*, 3101–3103; *Angew. Chem. Int. Ed.* **2001**, *40*, 3012–3014.
[24] C. Brotschi and C. J. Leumann, *Angew. Chem.* **2003**, *115*, 1694–1697; *Angew. Chem. Int. Ed.* **2003**, *42*, 1655–1658.
[25] I. Singh, W. Hecker, A. K. Prasad, V. S. Parmar, O. Seitz, *Chem. Commun.* **2002**, 500–501.
[26] C. Beuck, I. Singh, A. Bhattacharya, W. Hecker, V. S. Parmar, O. Seitz, E. Weinhold, *Angew. Chem.* **2003**, *115*, 4088–4091; *Angew. Chem. Int. Ed.* **2003**, *42*, 3958–3960.
[27] G. A. Kraus, M. T. Molina, *J. Org. Chem.* **1988**, *53*, 752–753.
[28] K. Krohn, H. Heins, K. Wielckens, *J. Med. Chem.* **1992**, *35*, 511–517.
[29] W. Timpe, K. Dax, N. Wolf, H. Weidmann, *Carbohydr. Res.* **1975**, *39*, 53–60.
[30] A. Klapars, X. H. Huang, S. L. Buchwald, *J. Am. Chem. Soc.* **2002**, *124*, 7421–7428.
[31] A. Klapars, S. L. Buchwald, *J. Am. Chem. Soc.* **2002**, *124*, 14844–14845.
[32] P. F. H. Schwab, F. Fleischer, J. Michl, *J. Org. Chem.* **2002**, *67*, 443–449.
[33] C. B. Nielsen, M. Petersen, E. B. Pedersen, P. E. Hansen, U. B. Christensen, *Bioconjugate Chem.* **2004**, *15*, 260–269.
[34] J. Becaud, I. Pompizi, C. J. Leumann, *J. Am. Chem. Soc.* **2003**, *125*, 15338–15342.
[35] J. C. Chaput, C. Switzer, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10614–10619.
[36] J. Sambrook, E. F. Fritsch, T. Maniatis in *Molecular Cloning, A Laboratory Manual*, *Vol. 1*, Cold Spring Harbor Laboratory Press, **1989**, pp. 6.36.

Received: August 19, 2004
Published online: January 31, 2005

