Nucleic acid secondary structures containing the double-headed nucleoside 5′(*S*)-*C*-(2-(thymin-1-yl)ethyl)thymidine

Charlotte Andersen,[a] Pawan K. Sharma,[b] Mikkel S. Christensen,[a] Signe I. Steffansen,[a] Charlotte M. Madsen[a] and Poul Nielsen*[a]

Received 27th June 2008, Accepted 5th August 2008
First published as an Advance Article on the web 18th September 2008
DOI: 10.1039/b810930c

Oligodeoxynucleotides containing the double-headed nucleoside 5′(*S*)-*C*-(2-(thymin-1-yl)ethyl)thymidine were prepared by standard solid phase synthesis. The synthetic building block for incorporating the double-headed moiety was prepared from thymidine, which was stereoselectively converted to a protected 5′(*S*)-*C*-hydroxyethyl derivative and used to alkylate the additional thymine by a Mitsunobu reaction. The oligodeoxynucleotides were studied in different nucleic acid secondary structures: duplexes, bulged duplexes, three-way junctions and artificial DNA zipper motifs. The thermal stability of these complexes was studied, demonstrating an almost uniform thermal penalty of incorporating one double-headed nucleoside moiety into a duplex or a bulged duplex, comparable to the effects of the previously reported double-headed nucleoside 5′(*S*)-*C*-(thymin-1-yl)methylthymidine. The additional base showed only very small effects when incorporated into DNA or RNA three-way junctions. The various DNA zipper arrangements indicated that extending the linker from methylene to ethylene almost completely removed the selective minor groove base–base stacking interactions observed for the methylene linker in a (–3)-zipper, whereas interactions, although somewhat smaller, were observed for the ethylene linker in a (–4)-zipper motif.

Introduction

Chemically modified oligonucleotides can serve in nucleic acid nanotechnology,[1,2] oligonucleotide-based diagnostics[3] and as antisense drugs.[4] Altered or improved selective nucleic acid recognition is a key aim, as is varying the properties of DNA as a skeleton for nanoscale engineering. The idea of using double-headed 2′-deoxynucleosides, in which additional nucleobases have been attached to 2′-deoxynucleotides, for these purposes has recently been introduced by us and others.[5–9] The additional nucleobase constitutes a potential recognition site on the outside of a duplex or another nucleic acid secondary structure. The nucleobase can participate in specific hydrogen bonding as well as stacking and hereby pave the way for a second coding function, for interstrand communication and for thermal stabilisation of a secondary structure.

As our first example of a double-headed nucleoside, we introduced **1** (see Fig. 1) containing an additional nucleobase in the 2′-position.[6] The hybridisation properties of oligodeoxynucleotides (ODNs) containing this building block showed a slight thermal destabilisation of both DNA : RNA and DNA : DNA duplexes as compared with unmodified duplexes. By using the same ODNs and targets containing a hairpin structure, **1** was also incorporated in the branching point of a three-way junction, and this arrangement demonstrated a stabilising effect of the additional nucleobase as judged from an increase in melting temperature (T_m) of +4.2 °C in a DNA context and +2.5 °C with an RNA complement.[6]

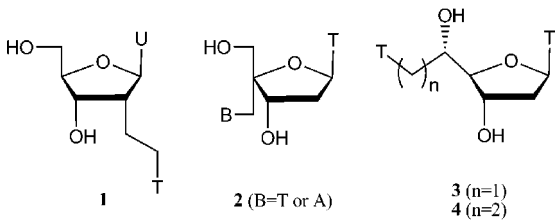


Fig. 1 Double-headed nucleosides. T = thymin-1-yl, U = uracil-1-yl, A = adenin-9-yl.

Herdewijn and co-workers introduced **2** as another example of a double-headed nucleoside with either thymine or adenine as the additional nucleobase pointing into the minor groove when incorporated into duplexes.[7] With the purpose of finding base–base interactions, preferably a classic A–T base-pairing in the minor grove, the most significant results were found when using solely the adenine moiety, and these interactions were rather weak and indicated to be a result of stacking and not of Watson–Crick type base-paring.[7]

Another recent example from our laboratory was the double-headed nucleoside **3**, where the additional thymine has been attached to the 5′-position *via* a methylene linker.[8] By using the 5′(*S*)-configuration, the substituent points towards the minor groove in a B-type duplex. By incorporating **3** into ODNs, the effect of the additional base was studied in different secondary nucleic acid structures. In standard duplexes the modification caused a constant drop in T_m of 4–6 °C in a DNA : DNA and 3.3 °C in a DNA : RNA duplex. However, interesting base–base

[a]*Nucleic Acid Center†, Department of Chemistry and Physics, University of Southern Denmark, 5230 Odense M, Denmark. E-mail: pon@ifk.sdu.dk; Fax: +45 66158780; Tel: +45 65502565*
[b]*Department of Chemistry, Kurukshetra University, Kurukshetra-136 119, India*

† The Nucleic Acid Center is funded by the Danish National Research Foundation for studies on nucleic acid chemical biology.

interactions in the minor groove were detected in an artificial DNA zipper arrangement of **3**. Hence, hybridisation studies and molecular modelling demonstrated a specific stacking interaction between the additional nucleobases when arranged in the (–3)-zipper arrangement, that is positioned on each complementary strand in the duplex with two interspacing base-pairs. Thus, the destabilisation of the duplex, which in all other cases was additive with two or more modifications, was counteracted by 6–7 °C with the (–3)-zipper arrangement.[8]

On the other hand, hybridisation studies of ODNs with **3** incorporated in the branching point of a three-way junction in the same way as for **1** did not show any thermal stabilisation in either a DNA : DNA or a DNA : RNA context. Molecular modelling of a DNA three-way junction with a CC bulge adjacent to the strand exchange site indicated that the additional base is positioned near the region of the bulge pyrimidines but with no specific contact.[8] Although the thermal studies of this type of structure did not indicate a clear interaction, *i.e.* neither a significant destabilisation or stabilisation, it is known that this type of structure exhibits folding isomerism with regard to the site of bulging out, *i.e.* between the junction and either the 3′ end or the 5′ end.[10] It is therefore entirely possible that the modelling structure and the experimental system display different isomeric conformations. To investigate this end further, we reasoned that a longer linker such as homologue **4** (Fig. 1) would increase the possibility of interaction between the extra base and the bulge.

Also the influence of linker length on the base–base interactions observed in a DNA zipper should be investigated. Hence, a modelling study of **4** similar to the homologue **3**[8] with a (–3)-zipper arrangement in a B-DNA duplex suggested that the extra nucleobases stack in the minor groove (Fig. 2). However, whereas the shorter methylene linker places the two additional thymine bases in an arrangement almost co-planar to the bases in the core strands (Fig. 2a),[8] the longer ethylene linker cannot accommodate such a conformation and places the extra nucleobases in a perpendicular arrangement (Fig. 2b).

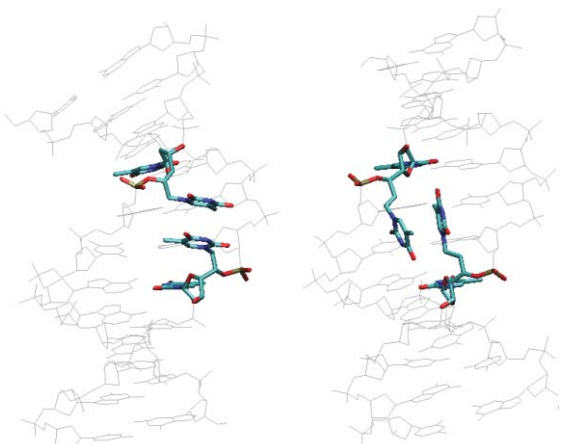


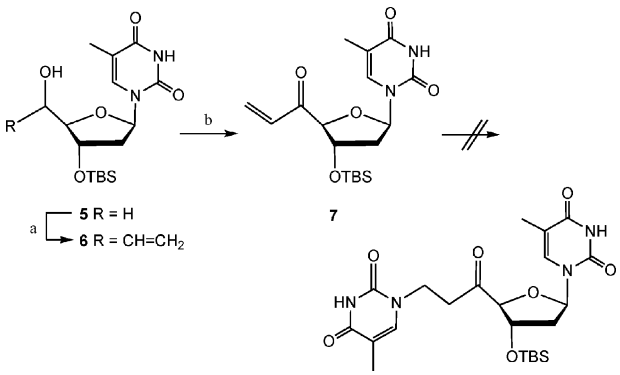
Fig. 2 Average structures from MD-simulation at 298 K. (a) Duplex with **3** incorporated in each strand in a (–3)-zipper arrangement.[8] (b) The corresponding duplex with **4**.

Herein we report the synthesis and evaluation of **4** as a homologue of **3** with the linker one carbon extended, in order to investigate the effect of the longer linker on DNA secondary structures including modified DNA zippers.

Results and discussion

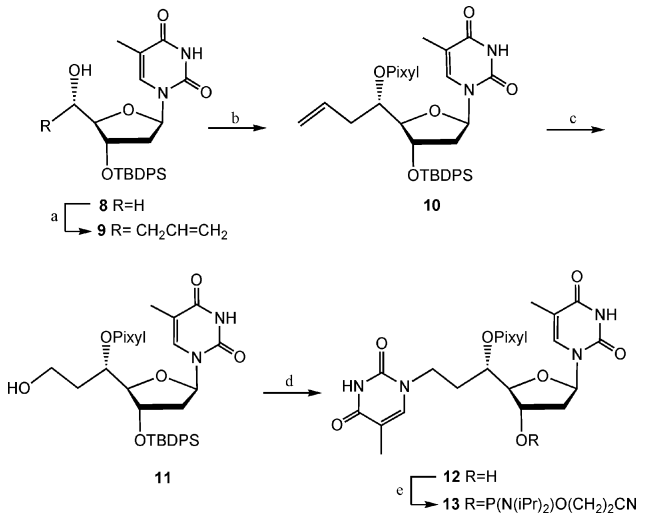
Chemical synthesis

The double-headed nucleoside **4** was synthetically approached by a linear strategy with thymidine as the starting material. The first attempted strategy was designed using a Michael-type addition as the key step for introducing the additional thymine (Scheme 1). Previously, several acyclic nucleosides have been synthesised using a Michael-type strategy.[11–13] Using a known method, the 3′-protected derivative **5** was converted to the 5′-*C*-vinylthymidine derivative **6** as an epimeric mixture by oxidation, Grignard reaction affording the 5′-*C*-ethynyl derivative and Lindlar hydrogenation.[14] The Michael substrate **7** was obtained by oxidation of the secondary alcohol in **6** using Dess–Martin periodinane. The Michael-type addition was attempted with either thymine or 3-*N*-benzoylthymine using numerous different conditions including heating in DMSO with or without the addition of different bases or treatment in ionic liquid.[13] Unfortunately, no double-headed nucleoside was isolated, probably due to the reversibility of the Michael-type reaction.



Scheme 1 *Reagents and conditions*: a, ref. 14: (i) DMP, DCM, 76%, (ii) HCCMgBr, THF, 70%, (iii) Lindlar cat., H_2, quinoline, EtOAc, 92%; b, DMP, DCM, 76%; TBS = *tert*-butyldimethylsilyl.

The synthetic strategy was then changed, and the introduction of the additional thymine was approached *via* a Mitsunobu reaction as applied in the preparation of **1**. Firstly, attempts were made to transform compound **6** to the 1,3-dihydroxyl derivative by hydroboration of the double bond, but only with limited success. Therefore, we decided to approach the target by introducing a 5′-*C*-allyl group, which could be subjected to an oxidative cleavage (Scheme 2). We have previously reported[15] that the 5′-*C*-allylation can be performed with complete stereoselectivity by the use of the TBDPS protecting group instead of the TBS group.[16] Hence, the 5′-hydroxy group of the thymidine derivative **8** was oxidised and the allyl linker was introduced *via* a Lewis acid catalyzed reaction to give exclusively the 5′(*S*)-configured product **9**.[15] The secondary alcohol of compound **9** was protected with the pixyl group to give **10** in 96% yield. The pixyl group has proven superior to the more traditional DMT (4,4′-dimethoxytrityl) group for protection of the secondary 5′-hydroxy groups and equally efficient in the subsequent automatic DNA-synthesis.[8,15,17] The double bond was



Scheme 2 *Reagents and conditions*: a, ref. 15: (i) DMP, DCM, 81%, (ii) allyltrimethylsilane, $BF_3{\cdot}OEt_2$, DCM, 75%; b, pixylchloride, pyridine, 96%; c, (i) OsO_4, $NaIO_4$ THF, (ii) $NaBH_4$, THF, H_2O, 53% over 2 steps; d, (i) T^{Bz}, PPh_3, DEAD, (ii) TBAF, THF, (iii) NH_3(aq), CH_3OH, 80% over 3 steps; e, $NC(CH_2)_2OP(N(iPr)_2)_2$, diisopropylammonium tetrazolide, DMF, CH_3CN, 45%; TBDPS = *tert*-butyldiphenylsilyl, pixyl = 9-phenylxanthen-9-yl.

cleaved to give the primary alcohol **11** in 53% yield *via* a standard procedure consisting of dihydroxylation by OsO_4, oxidative cleavage by $NaIO_4$ and reduction using $NaBH_4$. Introduction of the additional nucleobase was accomplished using the Mitsunobu reaction with 3-*N*-benzoylthymine as the nucleophile, giving the desired double-headed nucleoside. Subsequent treatment with TBAF deprotected the 3′-hydroxy group, and cleavage of the benzoyl group was obtained by treatment with 25% aqueous ammonia and methanol. Compound **12** was obtained in 80% yield over the 3 steps and subsequently converted to the desired phosphoramidite **13** using a standard protocol.

Preparation and evaluation of oligodeoxynucleotides

The double-headed nucleoside phosphoramidite **13** was incorporated into DNA sequences using standard automated solid phase DNA-synthesis with 1*H*-tetrazole as the activator and a prolonged coupling time for **13** of 20 minutes. The coupling yield for **13** was >90% and standard acidic treatment after the coupling also removed the pixyl group. The constitution and purity of the ODNs were verified by MALDI-MS and ion exchange/RP-HPLC, respectively. The ODN sequences applied in this study (**ON1–ON9**) are shown in Table 1 and 2. The hybridisation properties of the ODNs with different complementary DNA and RNA sequences were evaluated by thermal stability experiments, and the results from these experiments are further depicted in Tables 1 and 2.

Firstly, the modified sequence **ON3**, containing one central double-headed nucleoside **4**, was studied and compared with the unmodified sequences **ON1** and **ON2**. The effect of the modification was evaluated in duplexes, bulged duplexes and three-way junctions formed with complementary DNA and RNA (Table 1). The duplexes formed between **ON3** and complementary DNA and RNA were destabilised in both cases with ΔT_m of –5 °C and –2.5 °C, respectively, compared with unmodified **ON1**. This indicates that incorporation of the additional nucleobase in the minor groove is not favourable, but on the other hand is reasonably well accepted in the duplex. This result is similar to the previously reported results with the shorter linker in the double-headed nucleoside **3**.[8]

Subsequently, the hybridisation properties of **ON3** with complementary bulged DNA and RNA were studied. The penalty for the incorporation of **4** is seen in all the formed duplexes as a drop in T_m of around 4 °C, indicating that the bulges are willing to accommodate the additional thymine in the same way as duplexes. On the other hand, no positive influence is indicated by the additional base in the bulged duplexes. It is clear when comparing the T_m values of **ON2** against the complementary strand forming a regular duplex and **ON3** with the same complement forming an A-bulge that the large drop in T_m of 12.5 °C strongly indicates that the double-headed nucleoside does not behave as a TT-dinucleotide. This also indicates that the additional nucleobase is not likely intercalated in the duplex core. Similar destabilisation is seen in duplexes with GA bulges opposing the double-headed nucleoside.

For the three-way junction studies (Table 1), two different sequences containing a hairpin-loop were used. Thus, the DNA complement contains an additional CC-bulge as compared with

Table 1 Hybridisation data for duplexes, bulged duplexes and three-way junctions[a]

		5′-GCT CAC **X**CT CCC A				
		X = T	**X** = TT	**X** = **4**		
		(**ON1**)	(**ON2**)	(**ON3**)	ΔT_m 1[b]	ΔT_m 2[c]
DNA	3′- d(CGA GTG AGA GGG T)	51.0	—	46.0	–5.0	—
RNA	3′- r(CGA GUG AGA GGG U)	58.5	—	56.0	–2.5	—
DNA, A-bulge	3′- d(CGA GTG AAG AGG GT)	43.0	52.0	39.0	–4.0	–13.0
RNA, A-bulge	3′- r(CGA GUG AAG AGG GU)	49.5	58.5	46.0	–3.5	–12.5
DNA, GA-bulge	3′- d(CGA GTG AGA GAG GGT)	42.5	42.0	38.0	–4.5	–4.0
RNA, GA-bulge	3′- r(CGA GUG AGA GAG GGU)	49.5	48.0	46.0	–3.5	–2.0
DNA hairpin	3′- d(CGA GTG ACC CGC GTT TTC GCG AGA GGG T)	23.5	25.0	22.5	–1.0	–2.5
RNA hairpin	3′- r(CGA GUG ACG CGU UUU CGC GAG AGG GU)	36.0	36.5	35.0	–1.0	–1.5

[a] Melting temperatures (T_m values/°C) obtained from the maxima of the first derivatives of the melting curves (A_{260} *vs.* temperature) recorded in a medium salt buffer (Na_2HPO_4 (7.5 mM), NaCl (100 mM), EDTA (0.1 mM), pH 7.0) using 1.0 μM concentrations of each strand. [b] $\Delta T_m 1 = T_m$ (**ON3**) – T_m (**ON1**), [c] $\Delta T_m 2 = T_m$ (**ON3**) – T_m (**ON2**).

Table 2 Hybridisation data for the DNA zipper motifs[a]

	Duplex sequence	T_m	ΔT_m	ΔT_m3		Duplex sequence	T_m	ΔT_m	$\Delta\Delta T_m$	ΔT_m3	$\Delta\Delta T_m3$
	5′-CGC ATA TTC GC 3′-GCG TAT AAG CG	45.5			**ON7** **ON4**	5′-CGC ATA T**X**C GC 3′-GCG **X**AT AAG CG	36.0	–9.5	**+1.2**	–9.8	0.0
ON4	5′-CGC ATA TTC GC 3′-GCG **X**AT AAG CG	39.0	–6.5	–5.4	**ON8** **ON4**	5′-CGC ATA **X**TC GC 3′-GCG **X**AT AAG CG	37.5	–8.0	**+2.7**	–3.8	**+6.5**
ON5	5′-CGC ATA TTC GC 3′-GCG TA**X** AAG CG	40.0	–5.5	–4.7	**ON7** **ON5**	5′-CGC ATA T**X**C GC 3′-GCG TA**X** AAG CG	35.8	–9.7	0.0	–10.0	–0.9
ON6	5′-CGC ATA TTC GC 3′-GCG **X**A**X** AAG CG	33.5	–12.0	–10.9	**ON8** **ON5**	5′-CGC ATA **X**TC GC 3′-GCG TA**X** AAG CG	35.5	–10.0	–0.3	–10.4	–0.8
					ON9 **ON4**	5′-CGC A**X**A **X**TC GC 3′-GCG **X**AT AAG CG	31.0	–14.5	**+1.0**	–10.4	**+7.0**
ON7	5′-CGC ATA T**X**C GC 3′-GCG TAT AAG CG	41.3	–4.2	–4.4	**ON9** **ON5**	5′-CGC A**X**A **X**TC GC 3′-GCG TA**X** AAG CG	31.3	–14.2	+0.3	–18.2	–1.5
ON8	5′-CGC ATA **X**TC GC 3′-GCG TAT AAG CG	41.3	–4.2	–4.9	**ON7** **ON6**	5′-CGC ATA T**X**C GC 3′-GCG **X**A**X** AAG CG	31.0	–14.5	**+1.7**	–17.0	–1.7
ON9	5′-CGC A**X**A **X**TC GC 3′-GCG TAT AAG CG	36.5	–9.0	–12.0	**ON8** **ON6**	5′-CGC ATA **X**TC GC 3′-GCG **X**A**X** AAG CG	29.8	–15.7	+0.5	–9.0	**+6.8**
					ON9 **ON6**	5′-CGC A**X**A **X**TC GC 3′-GCG **X**A**X** AAG CG	24.5	–21.0	0.0	–16.8	**+6.1**

[a] Melting temperatures (T_m values/°C) obtained from the maxima of the first derivatives of the meltng curves (A_{260} *vs.* temperature) recorded in a medium salt buffer (Na_2HPO_4 (7.5 mM), NaCl (100 mM), EDTA (0.1 mM), pH 7.0) using 1.0 μM concentrations of each strand. **X** corresponds to **4** by the incorporation of **13**. $\Delta\Delta T_m$ values for duplex x : y are defined as $\Delta T_{m(x:y)} - (\Delta T_{m(x:ref)} + \Delta T_{m(y:ref)})$. ΔT_m3 and $\Delta\Delta T_m3$ correspond to the similar measurements for sequences containing **3** taken from ref. 8.

the RNA-complement in order to form a more well-defined secondary structure,[10] as explained in our former study.[8] The hybridisation studies of **ON3** show, however, that the double-headed nucleoside **4** is not able to stabilise either of the two three-way junctions when compared to the dinucleotide in **ON2**, or the single thymidine in **ON1**, as small destabilisations of 1–2 °C are seen in all cases. Comparing this destabilisation with the penalty observed with incorporation into duplexes, it is clear that the additional base is better accommodated in the three-way junctions, possibly due to the more flexible system. On the other hand, no stabilisation is observed, and this indicates that the linker extension on **4** as compared with **3** is not leading the additional base in the DNA : DNA-complex into a contact with the CC-bulge. The results with **4** are not different from the results with **3**,[8] and this may also suggest that the isomerism is different for the modelling system and the experimental system.

In order to examine the possibility of minor groove base–base interactions in an artificial DNA zipper motif, the incorporation of **4** into the same series of sequences as **3** was successfully performed. These results are illustrated in Table 2 in combination with the results from the previous study of **3** for direct comparison. Firstly, single incorporations were studied with the sequences **ON4–ON9** hybridised to their unmodified complements, and in all cases decreases in melting temperatures of 4.2–6.5 °C were detected with the expected additive effect of the two modifications in **ON6** and **ON9**. When comparing the two modifications **4** and **3** (ΔT_m and ΔT_m3, respectively), the thermal penalty paid in one sequence tends to be higher for **4** than for **3**, whereas it tends to be opposite when the complementary sequence is modified. Nevertheless, the decrease in T_m is relatively uniform in all cases. Hereafter, the zipper motifs in which the additional nucleobases are positioned in the opposite sequences with distances of 0–3 interspacing base-pairs (corresponding to (–1) to (–4)-zippers) were studied. The earlier results with **3**, shown in the right hand column of Table 2, demonstrate the very selective thermally stable formation of (–3)-zippers. Only in that specific positioning was a relative increase in T_m (defined as $\Delta\Delta T_m$) of 6–7 °C observed, most likely as a result of a strong stacking interaction in the minor groove as supported by modelling experiments (Fig. 2).[8] In the case of **4**, the results are completely different, demonstrating only a much smaller effect when positioned in a (–3)-zipper, as seen by a $\Delta\Delta T_m$ of +2.7 °C in the case of **ON8** : **ON4**, decreasing to +1.0 °C, +0.5 °C and even 0 °C in **ON9** : **ON4**, **ON8** : **ON6**, and **ON9** : **ON6**, respectively. Therefore, although the initial modelling study indicates the two extra bases to be stacked (Fig. 2), this particular arrangement leads to significantly poorer interactions than seen with the simple methylene linker arrangement. On the other hand, a small increase is also seen with the (–4)-zipper in **ON7** : **ON4** and **ON7** : **ON6**, with a $\Delta\Delta T_m$ of +1.2 °C and +1.7 °C, respectively. In all other cases, T_m indicated an additive destabilisation when **4** was incorporated multiple times. These results demonstrate the unique ability of **3** to position the additional nucleobases in the minor groove in a position allowing efficient stacking in a (–3)-zipper motif, and suggests that the extension of the linker in **4** induces a flexibility that is not accepted. Due to the same flexibility, the

indicated base–base contact in a (−4)-zipper is much smaller and less conclusive.

Conclusion

The synthesis of the double-headed nucleoside 5′(*S*)-*C*-(2-(thymin-1-yl)ethyl)thymidine **4** has been accomplished and its incorporation into oligodeoxynucleotides was straightforward. The relatively small thermal penalty paid by incorporating this moiety and positioning the additional nucleobase on the surface of a nucleic acid duplex showed that this is well accommodated and available for recognition. On the other hand, no stabilisation by the additional base in combination with the ethylene linker was detected in the secondary structures studied. The data proved, first of all, that the specific base–base contact in a DNA (−3)-zipper positioning of the two opposite moieties of **3** is unique and lost by increasing the linker length. Therefore, we are concentrating our future studies on this motif by varying the additional nucleobase moiety.

Experimental section

All commercial reagents were used as supplied. All reactions were conducted under a nitrogen atmosphere. Column chromatography was carried out using silica gel 60 (0.040–0.063 mm). HiRes MALDI mass spectra were recorded on an Ionspec Ultima Fourier Transform mass spectrometer using 2,5-dihydroxybenzoic acid as a matrix. NMR spectra were recorded on a 300 MHz Varian spectrometer. The values for δ are in ppm relative to tetramethylsilane as an internal standard for the ^{1}H NMR assignments, relative to the solvent signal (DMSO-d_6 39.52; $CDCl_3$ 77.16) for the ^{13}C-NMR assignments, and relative to 85% H_3PO_4 as an external standard for the ^{31}P NMR spectra. Assignments of NMR spectra are based on 2D and/or DEPT spectra and follow nucleoside naming style, *i.e.* the anomeric carbon is C-1′. No distinction has been made between the two thymine nucleobases.

Preparation of 5′-*C*-vinyl-5′-oxo-3′-*O*-(*tert*-butyldimethylsilyl)thymidine (7)

A solution of compound **6** (100 mg, 0.26 mmol) in anhydrous CH_2Cl_2 was stirred at room temperature. Dess–Martin periodinane (128 mg, 0.30 mmol) was added and the mixture was stirred for 20 h. The reaction mixture was diluted with CH_2Cl_2 (20 mL) and washed with a saturated aqueous solution of $Na_2S_2O_3$ (10 mL) and a saturated aqueous solution of $NaHCO_3$ (10 mL), dried (Na_2SO_4) and concentrated under reduced pressure. The residue was purified by column chromatography (0–80% ethyl acetate in petroleum ether) to give **7** as a white foam (76 mg, 76%). R_f 0.71 (75% ethyl acetate in petroleum ether). ^{1}H NMR ($CDCl_3$): δ 9.03 (1H, br s, NH), 8.09 (1H, d, J = 1.2, H6), 6.59–6.42 (3H, m, H1′, H7′), 6.02 (1H, dd, J = 1.5, 9.6, H6′), 4.82 (1H, d, J = 1.8, H4′), 4.48 (1H, m, H3′), 2.17–1.95 (2H, m, H2′), 1.98 (3H, s, CH_3), 0.94–0.85 (9H, m, $C(CH_3)_3$), 0.14–0.07 (6H, m, $SiCH_3$). ^{13}C NMR ($CDCl_3$): δ 197.0 (C5′), 164.1 (C4), 150.5 (C2), 136.4 (C6), 133.0, 131.3 (C6′, C7′), 111.2 (C5), 88.9 (C1′), 86.2 (C4′), 74.1 (C3′), 39.8 (C2′), 25.6 ($C(CH_3)_3$), 17.9 ($CCH_3)_3$), 12.6 (CH_3), −4.6, −4.8 ($SiCH_3$). HiRes MALDI MS m/z (M + Na^+) found/calc. 403.1667/403.1660.

Preparation of 5′(*S*)-*C*-(2-hydroxyethyl)-5′-*O*-(9-phenylxanthen-9-yl)-3′-*O*-(*tert*-butyldiphenylsilyl)thymidine (11)

To a solution of compound **10**[15] (2.00 g, 2.58 mmol) in THF (20 mL) stirred at room temperature was added a solution of $NaIO_4$ (2.10 g, 9.79 mmol) in H_2O (20 mL) and a 2.5% solution of OsO_4 in *tert*-butyl alcohol (0.26 mL, 26 μmol). The reaction mixture was stirred at room temperature for 26 h. Ethylene glycol (2 mL, 144 mmol) was added, and the resulting mixture was stirred for 15 min whereupon H_2O (40 mL) was added. The mixture was extracted by CH_2Cl_2 (2 × 50 mL), and the organic phase was dried (Na_2SO_4) and concentrated under reduced pressure. The residue was dissolved in a mixture of THF (20 mL) and H_2O (20 mL), and $NaBH_4$ (195 mg, 5.16 mmol) was added. After being stirred at room temperature for 1 h, the reaction mixture was extracted with CH_2Cl_2 (2 × 50 mL). The organic phase was washed with a saturated aqueous solution of $NaHCO_3$ (2 × 50 mL), dried (Na_2SO_4) and concentrated under reduced pressure. The residue was purified by column chromatography (0–2.5% CH_3OH and 0.1% pyridine in CH_2Cl_2) to give the product **11** as a white foam (1.06 g, 53%): R_f 0.45 (50% ethyl acetate in petroleum ether). ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.36 (br s, 1H, NH) 7.81 (d, 1H, J = 1.2 Hz, H-6), 7.60–7.10 (m, 17H, Ph), 7.05–6.80 (m, 4H, Ph), 6.70 (m, 1H, Ph), 6.45 (m, 1H, Ph), 6.35 (dd, 1H, J = 4.8, 9.3 Hz, H-1′), 3.55–3.50 (m, 2H, H-3′, H-5′), 3.10–2.90 (m, 3H, H-4′, H-7′), 2.28 (dd, 1H, J = 4.8, 12.3 Hz, H-2′), 2.01 (d, 3H, J = 1.2 Hz, CH_3), 1.92 (m, 1H, H-2′), 1.25 (m, 1H, H-6′), 1.18 (m, 1H, OH), 0.98 (s, 9H, $SiC(CH_3)_3$) 0.75 (m, 1H, H-6′); ^{13}C NMR (75 MHz, $CDCl_3$) δ 163.7 (C-4), 151.7, 151.6, 150.1, 147.3 (C-2, Ph), 135.9, 135.7, 133.5, 133.1, 131.7, 130.4, 130.0, 129.8, 128.1, 127.9, 127.6, 127.2, 123.5, 123.3, 123.2, 123.1, 116.7, 116.5 (C-6, Ph), 110.7 (C-5), 89.8, 85.7, 75.2, 71.3, 58.7 (C-1′, C-3′, C-4′, C-5′, C-7′), 41.3 (C-2′), 35.4 (C-6′), 26.9 ($C(CH_3)_3$), 18.7 ($CCH_3)_3$), 12.8 (CH_3). HiRes MALDI MS m/z (M + Na^+) found/calc. 803.3104/873.3128.

Preparation of 5′(*S*)-*C*-(2-(thymin-1-yl)ethyl)-5′-*O*-(9-phenylxanthen-9-yl)thymidine (12)

A solution of **11** (1.00 g, 1.28 mmol), 3-*N*-benzoylthymine (649 mg, 2.82 mmol) and PPh_3 (1.01 g, 3.85 mmol) in anhydrous THF (30 mL) was stirred at 0 °C, and DEAD (1.56 mL, 3.59 mmol) was added dropwise. The solution was stirred for 6 h at room temperature and concentrated under reduced pressure. The residue was diluted with ethyl acetate (50 mL) and washed with a saturated aqueous solution of $NaHCO_3$ (50 mL). The aqueous layer was extracted with ethyl acetate (2 × 20 mL), and the combined organic phase was dried ($MgSO_4$) and concentrated under reduced pressure. The residue was dissolved in anhydrous THF (30 mL) and a 1.0 M solution of tetrabutylammonium fluoride in THF (2.20 mL) was added. The solution was stirred at room temperature for 2 h and then partitioned between water (50 mL) and ethyl acetate (50 mL). The aqueous phase was extracted with ethyl acetate (2 × 25 mL), and the combined organic phase was washed with brine (50 mL), dried (Na_2SO_4) and concentrated under reduced pressure. The residue was dissolved in methanol (30 mL) and a 25% aqueous solution of ammonia (3.1 mL) was added dropwise. The reaction mixture was stirred for 16 h at room temperature, concentrated under reduced pressure and purified by silica gel column chromatography (1–9% CH_3OH

and 0.1% pyridine in CH_2Cl_2) to give the product **12** as a white foam (663 mg, 80%): R_f 0.15 (7.5% CH_3OH in ethyl acetate). ^1H NMR (300 MHz, DMSO-d_6) δ 11.32 (br s, 1H, NH), 11.12 (br s, 1H, NH), 7.81 (d, 1H, $J = 1.2$ Hz, H-6), 7.40–6.94 (m, 14H, Ph, H-6), 6.00 (dd, 1H, $J = 6.6$, 7.2 Hz, H-1′), 5.08 (d, 1H, $J = 4.5$ Hz, OH), 3.76 (m, 1H, H-3′), 3.63 (t, 1H, $J = 3.6$ Hz, H-4′), 3.50–3.15 (m, 3H, H-5′, H-7′), 2.05–1.90 (m, 2H, H-2′), 1.78 (s, 3H, CH_3), 1.72 (d, 3H, $J = 1.2$ Hz, CH_3), 1.40–1.25 (m, 2H, H-6′); ^{13}C NMR (75 MHz, DMSO-d_6) δ 164.2 (C-4), 163.73 (C-4), 150.9, 150.7, 150.5, 150.3, 148.0 (C-2, Ph), 137.7, 135.6, 131.2, 131.1, 129.8, 129.7, 127.7, 127.0, 126.8, 123.4, 123.2, 123.1, 122.5, 116.3, 115.9 (C-6, Ph), 109.7, 108.5 (C-5), 86.5, 83.1, 76.2, 71.1, 70.3 (C-1′, C-3′, C-4′, C-5′, CPh_3), 54.9 (C-7′), 43.5 (C-2′), 30.8 (C-6′), 12.3 (CH_3), 11.9 (CH_3). HiRes ESI MS m/z (M + Na^+) found/calc. 673.2269/673.2274.

Preparation of 5′(*S*)-*C*-(2-(thymin-1-yl)ethyl)-5′-*O*-(9-phenylxanthen-9-yl)-3′-*O*-(*P*-β-cyanoethoxy-*N,N*-diisopropylaminophosphinyl)thymidine (13)

A solution of compound **12** (500 mg, 0.923 mmol) and diisopropylammonium tetrazolide (236 mg, 1.39 mmol) in anhydrous DMF (10 mL) and anhydrous CH_3CN (10 mL) was stirred at room temperature and *N,N,N′,N′*-tetraisopropylphosphoramidite (0.44 mL, 1.38 mmol) was added. The reaction mixture was stirred for 10 h and diluted with ethyl acetate (20 mL). The solution was washed with brine (2 × 20 mL). The combined aqueous phase was extracted with ethyl acetate (20 mL), and the combined organic phase was dried (Na_2SO_4) and concentrated under reduced pressure. The residue was purified by column chromatography (1–6% CH_3OH and 0.1% pyridine in CH_2Cl_2) to give the product **13** as a white foam (291 mg, 45%): R_f 0.25 (7.5% CH_3OH in ethyl acetate). ^{31}P NMR (121.5 MHz, $CDCl_3$) δ 150.26, 149.88. HiRes ESI MS m/z (M + Na^+) found/calc. 873.3347/873.3353.

Oligonucleotide synthesis and hybridisation experiments

The oligodeoxynucleotides were synthesised using an automated Expedite 8909 nucleic acid synthesis system following the phosphoramidite approach. Synthesis of oligonucleotides **ON3–ON9** was performed on a 0.2 μmol scale by using 2-cyanoethyl phosphoramidites of standard 2′-deoxynucleosides in combination with the modified phosphoramidite **13**. The synthesis followed the regular protocol employing standard CPG supports and 4,5-dicyanoimidazole as the activator except in the case of the modified amidites, which were manually coupled using 0.05 M amidite and 0.5 M tetrazole as the activator in CH_3CN for 20 min. The coupling yields for **13** in combination with the following unmodified amidites were in the range of 91–100%. The 5′-*O*-DMT oligonucleotides were removed from the solid support by treatment with concentrated aqueous ammonia at 55 °C for 16–24 h, which also removed the protecting groups. The oligonucleotides were purified by reversed-phase HPLC on a Waters 600 system using an X_{terra} prep MS C_{18}; 10 μm; 7.8 × 150 mm column; gradient of buffer (0.05M triethylammonium acetate) in 75% CH_3CN(aq): 0–70% buffer, 38 min; 70–100% buffer, 7 min; 100% buffer, 10 min. All fractions containing 5′-*O*-DMT protected oligonucleotide (retention time 20–30 minutes) were collected and concentrated. The products were detritylated by treatment with an 80% aqueous solution of acetic acid for 20 min, and finally isolated by precipitation with ethanol at –18 °C overnight. MALDI-MS m/z (found/calcd); **ON3** (3979.1/3982.6), **ON4** (3532.3/3533.4), **ON5** (3530.2/3533.4), **ON6** (3682.4/3685.4), **ON7** (3442.8/3444.3), **ON8** (3444.3/3444.3), **ON9** (3594.4/3596.3). After dissolution in double distilled water, the concentrations were determined spectrometrically at 260 nm in the pH 7.0 buffer, assuming an extinction coefficient for the modified double-headed nucleoside of equal to twice that of thymidine. The UV-melting curves were acquired at 260 nm with a scan rate of 0.2 °C min^{-1} cycled between 15 and 75 °C after denaturation at 80 °C. All melting curves were found to be reversible. The melting temperatures were taken as the first derivatives of the absorbance *versus* temperature up-curve and reported as the average of two measurements.

Modelling experiment

The MD-simulations in Fig. 2b were obtained by the same protocol as used in ref. 8.

Acknowledgements

The Danish National Research Foundation, The Danish Research Agency and Møllerens Fond are thanked for financial support. Mrs Birthe Haack is thanked for technical assistance.

References and Notes

1 H. Yan, *Science*, 2004, **306**, 2048.
2 J. Wengel, *Org. Biomol. Chem.*, 2004, **2**, 277.
3 P. Mouritzen, A. Toftgaard Nielsen, H. M. Pfundheller, Y. Choleva, L. Kongsbak and S. Møller, *Expert Rev. Mol. Diagn.*, 2003, **3**, 89.
4 J. Kurreck, *Eur. J. Biochem.*, 2003, **270**, 1628.
5 T. Wu, M. Froeyen, G. Schepers, K. Mullens, J. Rozenski, R. Busson, A. Van Aerscot and P. Herdewijn, *Org. Lett.*, 2004, **6**, 51.
6 S. L. Pedersen and P. Nielsen, *Org. Biomol. Chem.*, 2005, **3**, 3570.
7 T. Wu, K. Nauwelaerts, A. Van Aerschot, M. Froeyen, E. Lescrinier and P. Herdewijn, *J. Org. Chem.*, 2006, **71**, 5423.
8 M. S. Christensen, C. M. Madsen and P. Nielsen, *Org. Biomol. Chem.*, 2007, **5**, 1586.
9 M. S. Christensen, A. D. Bond and P. Nielsen, *Org. Biomol. Chem.*, 2008, **6**, 81.
10 N. B. Leontis, W. Kwok and J. S. Newman, *Nucleic Acids Res.*, 1991, **19**, 759.
11 H. B. Lazrek, H. Khaïder, A. Rachdi, J.-L. Barascut and J.-L. Imbach, *Tetrahedron Lett.*, 1996, **37**, 4701.
12 S. Guaillarme, S. Legoupy, A.-M. Aubertin, C. Olicard, N. Bourgougnon and F. Huet, *Tetrahedron*, 2003, **59**, 2177.
13 J.-M. Xu, C. Qian, B.-K. Liu, Q. Wu and X.-F. Lin, *Tetrahedron*, 2007, **63**, 986.
14 A. M. Sørensen, K. E. Nielsen, B. Vogg, J. P. Jacobsen and P. Nielsen, *Tetrahedron*, 2001, **57**, 10191.
15 P. K. Sharma, B. H. Mikkelsen, M. S. Christensen, K. E. Nielsen, C. Kirchhoff, S. L. Pedersen, A. M. Sørensen, K. Østergaard and P. Nielsen, *Org. Biomol. Chem.*, 2006, **4**, 2433.
16 V. Banuls and J.-M. Escudier, *Tetrahedron*, 1999, **55**, 5831.
17 H. Trafelet, E. Stulz and C. Leumann, *Helv. Chim. Acta*, 2001, **84**, 87.

