# Synthesis and modelling of DNA junction and minor groove zipper motifs incorporating the double-headed nucleoside 5′(*S*)–*C*-(thymine-1-ylmethyl)thymidine†

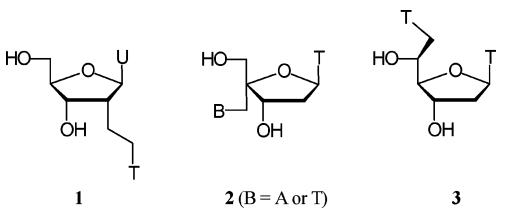
**Mikkel S. Christensen, Charlotte M. Madsen and Poul Nielsen***

***Received 18th January 2007, Accepted 23rd March 2007***
***First published as an Advance Article on the web 17th April 2007***
**DOI: 10.1039/b700852j**

A nucleoside with two nucleobases, a so-called double-headed nucleoside, 5′(*S*)–*C*-(thymine-1-ylmethyl)thymidine **3**, is synthesised and incorporated into oligonucleotides. The additional nucleobase is hereby positioned in the minor groove of the duplexes, which are formed with complementary DNA and RNA-sequences. Slight thermal destabilisation of these duplexes as compared to unmodified duplexes is observed. With other target sequences forming bulged duplexes or three-way junctions, no additional influence of the additional base on the thermal stability is observed. On the other hand, a base–base stacking interaction and subsequent stabilisation is observed when two double-headed nucleotide moieties are positioned in two complementary DNA-sequences forming a DNA-zipper motif.

## Introduction

In nucleic acid chemistry, the search for functional nucleic acid fragments has been intense and motivated by potentially therapeutic nucleic acid targeting and by nucleic acid chemical biology.[1–3] In the search for simple building blocks for functional nucleic acids, we have recently introduced the idea of positioning an additional nucleobase on an otherwise natural nucleotide moiety.[4,5] This use of "double-headed" nucleosides is versatile, and both the stabilisation of secondary structures like bulges or three-way junctions as well as the design of double-coding nucleic acids were envisioned.[4,5] By using an additional nucleobase in a nucleic acid sequence, both stacking and hydrogen-bonding interactions with the target are possible. Our first example of a double-headed nucleoside was 2′-deoxy-2′-*C*-(2-(thymin-1-yl)ethyl)uridine **1** (Fig. 1) which was incorporated into oligodeoxynucleotides (ODN's) and studied as a building block in different secondary nucleic acid structures.[4] Incorporated in the centre of a standard 13-mer DNA duplex or in the corresponding DNA : RNA hybrid duplex, the thermal stability decreased slightly compared to unmodified duplexes demonstrating that the additional base is reasonably well-accommodated in the duplexes but with no favourable new binding interactions. On the other hand, when **1** was incorporated in the branching point of a three-way junction, a slight stabilisation of the secondary structure was observed.[4] Herdewijn and co-workers has recently introduced another example, **2**, prepared with both thymine and adenine as the second base.[6] The focus of that study was to find an additional base–base interaction in the minor groove by the incorporation of two moieties of **2** at different positions in two complementary ODN's. The best interaction was found between two adenine moieties and not between a thymine and an adenine indicating that the interaction was not a Watson–Crick type base-pairing. An NMR-experiment indicated a weak hydrophobic interaction between the two bases in a perpendicular position.[6] Other double-headed nucleosides have been shown in the literature without any incorporation into ODN's. The second base has been positioned either directly on the ribofuranose on the 2′-position[7] or as a replacement of either the 3′-hydroxy group[7] or the 5′-hydroxy group.[8–10] In the present study, the additional nucleobase is positioned at the 5′-carbon through a methylene linker. Hence, the (*S*)-configured epimer of 5′-*C*-(thymine-1-ylmethyl)thymidine **3** was found by simple molecular modelling to give the most interesting opportunities, as the additional base of this double-headed nucleoside moiety will be positioned in the minor groove of nucleic acid duplexes and additional base–base interactions between complementary strands can be predicted. Furthermore, the ability to form stabilised secondary structures like three-way junctions will be studied.




**1** **2** (B = A or T) **3**

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

## Results and discussion

### Chemical synthesis

The synthesis of the double headed nucleoside **3** was approached with a linear strategy and thymidine as a convenient starting

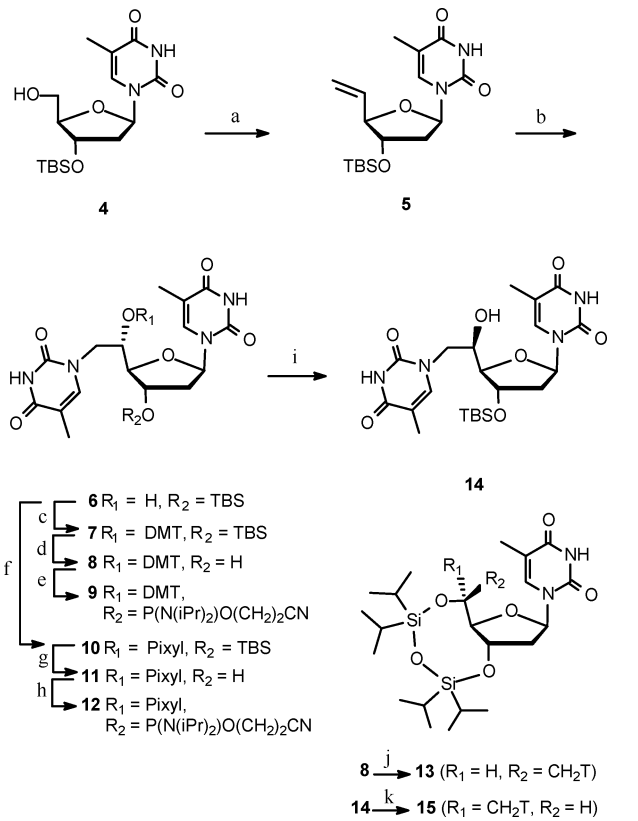
*Nucleic Acid Center‡, Department of Chemistry and Physics, University of Southern Denmark, 5230, Odense M, Denmark. E-mail: pon@ifk.sdu.dk; Fax: +45 66158780; Tel: +45 65502565*

† Electronic supplementary information (ESI) available: UV-melting profiles for self-complemenatary sequences **27**–**30**, Van't Hoff plots for obtaining the data in Table 6, and modelling data. See DOI: 10.1039/b700852j

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

material. Thus, several examples of introducing additional carbons and functionalisation on the 5′-*C* position have been shown in the literature.[11–17] From standard protocols, the 3′-protected derivative **4** was obtained and converted by oxidation and a Wittig olefination to the known 5′-deoxy-5′-methylene derivative **5** (Scheme 1).[11,12] The stereoselective epoxidation of this alkene has been described by Wang and Middleton,[12] and we used this preparation of the epimeric mixture of epoxides ($S : R \sim 9 : 1$), and a subsequent treatment with thymine and potassium carbonate in hot DMF to give the (*S*)-configured protected double-headed nucleoside **6** in 39% yield. The *N*1-alkylation of the extra thymine was assessed by the absence of coupling between any NH and any of the two H-6 in the ($^1$H,$^1$H)-COSY spectrum.




**Scheme 1** *Reagents and conditions*: a, ref. 11, 12; b, (i) mCPBA, $CH_2Cl_2$, 42%, (ii) thymine, $K_2CO_3$, DMF, 39%; c, DMTCl, pyridine, $CH_3CN$; d, TBAF, THF, 41% (2 steps); e, $NC(CH_2)_2OP(Cl)N(iPr)_2$, DIPEA, $CH_2Cl_2$, 58%; f, Pixyl-Cl, pyridine, 79%; g, TBAF, THF, 54%; h, $NC(CH_2)_2OP(Cl)N(iPr)_2$, DIPEA, $CH_2Cl_2$, $CHCl_3$, 64%; i, (i) $Tf_2O$, DMAP, $CH_2Cl_2$, (ii), NaOH (aq.), EtOH, 40%; j, (i) $Cl_2CHCOOH$, $CDCl_3$, (ii) $TIPDSCl_2$, pyridine, 39%; k, (i) TBAF, THF, (ii) $TIPDSCl_2$, pyridine, 53%. TBS = *tert*-butyldimethylsilyl. DMT = 4,4′-dimethoxytrityl. Pixyl = 9-phenylxanthen-9-yl.

In order to prepare the nucleoside for incorporation into ODN's, the 5′-hydroxy group was protected with the 4,4′-dimethoxytrityl (DMT) group. However, even prolonged exposure to a large excess of DMTCl in pyridine did not give a complete conversion, and after immediate desilylation of the crude product **7** by treatment with TBAF in THF, **8** was obtained in a mere yield of 41% over the two steps. This compound was converted to the phosphoramidite **9** by a standard protocol in 58% yield. Subsequent DNA synthesis using this amidite was successful (*vide infra*). Nevertheless, the difficulties in protecting the sterically hindered secondary alcohol in **6** motivated a replacement of the DMT-group with the Pixyl-group. Hence, the Pixyl group has been reported to give a more efficient protection of similar secondary alcohols and to be equally efficient in standard automated solid phase DNA-synthesis.[17,18] Treatment of **6** with Pixyl-Cl in pyridine afforded **10** in a good yield. Subsequent desilylation gave **11** and phosphitylation the phosphoramidite **12**, which was used in DNA synthesis with similar efficiency as **9** (*vide infra*).

In order to verify the stereochemistry of the double-headed nucleoside **6**, we employed a method used before by us[16,18] and others[11,12,17] for other 5′-*C*-alkylated thymidine derivatives. We converted an analytical amount of compound **8** into the restrained disiloxane compound **13** by *in situ* deprotection and treatment with an excess of the disiloxyl dichloride reagent in 39% overall yield. An NMR-analysis of this compound using coupling constants as well as NOE-difference data should give the relevant structural information. In the $^1$H NMR spectrum, the signal from H-4′ exhibited both a fairly weak coupling constant (2.7 Hz) and a strong coupling constant (8.4 Hz) indicating a *gauche* and a *trans* conformation in accordance with a moderate rigid bicyclic structure of **13**. However, a signal overlap of the H-5′ and H-3′ signals in the spectrum rendered the NOE method ambiguous. Therefore, the 5′-epimer of **6** was made on the analytical scale by treatment of **6** with triflic anhydride followed by hydrolysis of the triflate with aqueous base to give **14**. This reaction probably proceeded through an intramolecular ring-closure by the 2-oxygen of the 5′-positioned thymine giving a five-membered intermediate. Compound **14** was converted to **15** by *in situ* desilylation followed by treatment with an excess of the disiloxyl dichloride reagent. In contrast to **13**, compound **15** exhibited very well resolved signals in the $^1$H NMR spectrum showing a strong 7% NOE correlations between H-5′ and H-3′ as expected for this configuration, as well as insignificant enhancement of either of the latter two protons when H-4′ was irradiated. This proved an 5′(*R*)-configuration for compound **14** obtained by the inversion of the 5′(*S*)-configured **6**, and in conclusion, this analysis confirmed the 5′(*S*)-configuration of **6** that was also expected from the reported 5′(*S*)-configuration of the intermediate epoxide.[12]

## Preparation and evaluation of oligonucleotides

The oligonucleotides prepared for this study by standard automated solid phase DNA-synthesis are shown in Table 1. The reference sequences **16** and **17** are the same as used in former studies[18,19] including the study of our first double-headed nucleoside **1**.[4] The phosphoramidite **9** was incorporated in the central position of this sequence giving **18**. A prolonged coupling time for **9** was used with pyridinium chloride as the activator followed by prolonged capping time. The coupling efficiency for **9** was >90%. A prolonged coupling time was also used for the proceeding unmodified phosphoramidite using 4,5-dicyanoimidazole as the activator. For the extension of the study by the investigation of a DNA zipper motif with predicted base–base interaction in the minor groove, and due to the demand for an easier access to a phosphoramidite (*vide supra*), we prepared the other ODN's (Table 1) from the pixyl-protected phosphoramidite **12**. Thus, **12** was incorporated in different positions of the two complementary standard sequences **19** and **23** giving the series of ODN's

**Table 1** Prepared oligodeoxynucleotide sequences and their MS-data

| | ODN sequences[a] | MW (found/calc.)[b] |
|---|---|---|
| **16** | 5′-GCTCACTCTCCCA | — |
| **17** | 5′-GCTCACTTCTCCCA | — |
| **18** | 5′-GCTCAC**X**$^{T}$CTCCCA | 3968.9/3967.8 |
| **19** | 5′-CGCATATTCGC-3′ | — |
| **20** | 5′-CGCATAT**X**$^{T}$CGC-3′ | 3428.8/3429.7 |
| **21** | 5′-CGCATA**X**$^{T}$TCGC-3′ | 3427.8/3429.7 |
| **22** | 5′-CGCA**X**$^{T}$A**X**$^{T}$TCGC-3′ | 3567.7/3567.7 |
| **23** | 3′-GCGTATAAGCG-5′ | — |
| **24** | 3′-GCG**X**$^{T}$ATAAGCG-5′ | 3516.8/3518.7 |
| **25** | 3′-GCGTA**X**$^{T}$AAGCG-5′ | 3516.8/3518.7 |
| **26** | 3′-GCG**X**$^{T}$A**X**$^{T}$AAGCG-5′ | 3655.9/3656.7 |
| **27** | 5′-ATATATATATAT-3′ | — |
| **28** | 5′-ATATATA**X**$^{T}$ATAT-3′ | 3781.4/3779.7 |
| **29** | 5′-ATATA**X**$^{T}$ATA**X**$^{T}$AT-3′ | 3920.0/3917.8 |
| **30** | 5′-ATATA**X**$^{T}$A**X**$^{T}$A**X**$^{T}$AT-3′ | 4056.8/4055.8 |

[a] **X**$^{T}$ refers to the incorporation of **9** or **12**. [b] MALDI-MS negative mode.

**20–22** and **24–26**. For the study of a more elaborate zipper, also the self-complementary sequence **27** was prepared and modified with one, two or three incorporations of the double-headed modification in the sequences **28–30**. The coupling efficiency for **12** was >80% using prolonged coupling and capping times and pyridinium chloride as the activator. The constitution and purity of the oligonucleotides was verified by MALDI-MS (Table 1) and HPLC-profiles, respectively.

In order to evaluate the effect of one additional double-headed nucleoside moiety **3** in duplexes as well as in bulged duplexes and three-way junctions, the hybridisation properties of the oligonucleotides **16–18** with different complementary DNA and RNA sequences were studied by thermal stability examinations (Table 2). First, the standard duplexes formed with the fully matched DNA and RNA complements were examined. Comparing the duplexes formed by the sequences **16** and **18**, respectively, the double-headed nucleotide induced a drop in $T_m$ of 4.0 °C in a DNA : DNA duplex and of 3.3 °C in a DNA : RNA duplex. This indicates that the additional nucleobase is relatively well accommodated in standard duplexes but also that a small penalty of placing the additional base in the minor groove is paid. Next, a duplex extended with one adenosine bulge in the complementary sequence was studied (Table 2). Comparing **18** with **16**, small decreases in $T_m$ of ~3 °C were observed. Thus, the additional thymine is well accommodated into the bulged duplexes but no stabilisation due to base-pairing or stacking from the additional thymine is seen. In comparison with **17**, large drops in $T_m$ of the complexes of ~13 °C with both DNA and RNA complements were observed. In other words, the ability of **3** to behave as a TT-dinucleotide opposite an AA-complement is very poor. In another bulged duplex in which an AGA-sequence is opposing the double-headed nucleotide **3**, similar results were observed (Table 2). Thus, the comparison of **18** with either **16** or **17** revealed decreases in $T_m$ of ~3 °C, and the additional thymine does not seem to influence the structure significantly.

Next, we studied two three-way junctions (TWJ's) with a standard stable stem–loop sequence with two single stranded regions being complementary to the oligonucleotides **16–18** (Table 2). In **18**, the double-headed nucleotide **3** is positioned in the branching point. In our former study,[18] we found the DNA TWJ without any additional bulge in the branching point to be relatively flexible. This has also been observed by Leontis *et al.*[20] and Seeman *et al.*[21] and allows for structural deviation in the backbone in the vicinity of the junction. Therefore, a DNA TWJ with an additional CC bulge[18] was applied for this study in order to render the DNA TWJ less flexible.[20–22] This TWJ has a melting temperature similar to the TWJ missing the CC bulge,[18] but displays a much sharper melting transition, which indicates a more well-defined structure and a significantly larger negative enthalpy of formation.[23] Incorporating the double-headed nucleoside in this structure, a slight destabilisation ($\Delta T_m = -2.0$ °C) with **18** compared to **17** was observed. Considering that a backbone length corresponding to one nucleotide has been removed, this finding seems somewhat encouraging. In order to investigate the structure of the core of the junction, a model was built based on the coordinates from an NMR study[22] of a TWJ containing a CT bulge. Maintaining the original flanking duplexes of this structure but replacing the bulge and junction basepairs, gave a perturbed model, which was minimised and warmed up in a 100 ps MD-simulation to give the average structure in Fig. 2. The modeling result suggests that the extra thymine base is bulged out and may interact with the CC-bulge, while its opposite adenine base forms a wedged bulge. Thermal stability comparison of **18** with **16**, which would be anticipated to give the wedge adenine

**Table 2** Hybridisation data for the prepared ODN's with different DNA- and RNA-complements[a]

| | Regular or bulged complements | | | | | | Complements with an intrastrand stem-loop[f] | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Fully matched | | Additional A | | Additional GA | | DNA[g] | | RNA[h] | |
| | DNA[b] | RNA[b] | DNA[c] | RNA[c] | DNA[d] | RNA[d] | 0 mM | 10 mM | 0 mM | 10 mM |
| **16** | 51.0 | 58.2 | 41.0[e] | 48.0[e] | 41.1[e] | 47.8[e] | 25.1 | — | — | — |
| **17** | — | — | 51.3 | 57.6 | 40.5 | 47.8 | 24.9 | 35.3 | 36.5 | 44.7 |
| **18** | 47.0 (−4.0) | 54.9 (−3.3) | 37.9 (−3.1) (−13.4) | 45.1 (−2.9) (−12.5) | 37.9 (−3.2) (−2.6) | 44.8 (−3.0) (−3.0) | 22.9 (−2.2) (−2.0) | 31.9 (−3.4) | 35.1 (−1.4) | 43.1 (−1.6) |

[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. All $T_m$ values are given as averages of double determinations. $\Delta T_m$ values are given in brackets. [b] DNA 3′-CGAGTGAGAGGGT, RNA 3′-CGAGUGAGAGGGU. [c] DNA 3′-CGAGTGAAGAGGGT, RNA 3′-CGAGUGAAGAGGGU. [d] DNA 3′-CGAGTGA**G**AGAGGGT, RNA 3′-CGAGUGA**G**AGAGGGU. [e] Data taken from ref. 4. [f] 0 or 10 mM $Mg^{2+}$ obtained by the addition of $MgCl_2$. [g] DNA 3′-CGAGTGA**CCCGCGTTTTCGCG**AGAGGGT, [h] RNA 3′-CGAGUGA**CGCGUUUUCGCG**AG-AGGGU. Bold sequences are bulges and stem-loops.

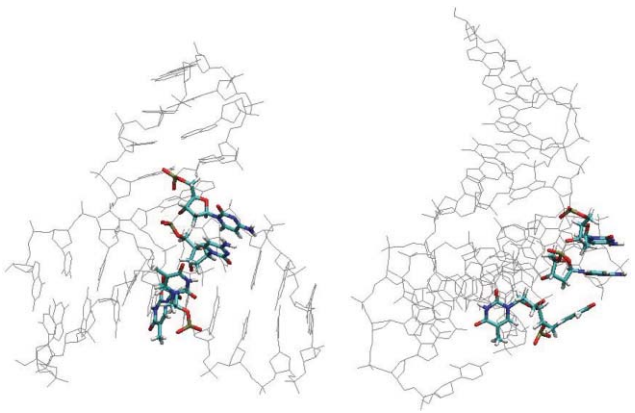


**Fig. 2** Molecular modelling of the core of the junction in the TWJ **18**:DNA. Average structure of the final 10 ps of a 100 ps MD-simulation at 303 K. The central double-headed nucleotide and the bulge are highlighted.

bulge without any possibility for a tertiary interaction with the CC bulge, does however not suggest any additional stabilisation from the extra thymine. In fact, the structures with **16** and **17** display very similar stability. It is well known that tertiary interactions are rather weak[24] and so the TWJ structure with **18** may, nevertheless, exhibit tertiary interactions as indicated by the modeling.

On the other hand, the DNA : RNA TWJ without a bulge in the branching point (*i.e.* **17** with RNA, Table 2) was found in our former study to be stable and well-defined.[18] A slight destabilisation by the incorporation of the double-headed moiety was observed with a $\Delta T_m$ −1.4 °C comparing **18** to **17**. Both TWJ's were stabilised significantly by the addition of $Mg^{2+}$ to a 10 mM concentration (Table 2), *i.e.* increases in $T_m$ of 8–10 °C as expected.[20] However, the unmodified TWJ's were in both cases slightly more stabilised than the modified TWJ's. Nevertheless, no significant differences in any $Mg^{2+}$-induced conformational changes in the TWJ's are indicated.

Inspired by the relatively modest but also remarkably constant decrease in $T_m$ observed with the incorporation of the double-headed nucleotide **3**, we decided to explore the possibility of a zipper motif with a stabilisation of the unpaired nucleobase moiety *via* intrastrand interaction with another unpaired nucleobase in the minor groove. A range of possible motifs was exploited by the preparation of a series of complementary sequences **19–26** (Table 1). A standard 11-mer duplex **19** : **23** was studied, and a single incorporation of **3** into this duplex (*i.e.* replacing **19** with **20** or **21**, *or* replacing **23** with **24** or **25**), in all cases demonstrated a decrease in $T_m$ of 4.4 to 5.4 °C (Table 3). This means that the penalty of accommodating one double-headed nucleoside moiety **3** in a duplex is remarkably constant taking also the duplex in Table 2 into account ($\Delta T_m = -4.0$ °C). Introducing two of these moieties on the same strand (as with **22** and **26**) gives the double or an even slightly higher penalty ($\Delta T_m = -12.0$ and −10.9 °C, respectively). On the other hand, an interesting contact between the thymine moieties in the minor groove is indicated with the duplex **21** : **24**, which is more thermally stable than either of the duplexes with just one modification, *i.e.* **21** : **23** and **19** : **24**. This suggests a stabilising correlation between the two thymine nucleobases positioned in opposite strands with two interspacing base pairs. Furthermore, it can be seen that moving the nucleobases further apart, *i.e.* **20** : **24** with three

**Table 3** Hybridisation data for the prepared ODN's in a DNA zipper motif[a]

| | **19** | **20** | **21** | **22** |
|---|---|---|---|---|
| **23** | 44.6 | 40.2 | 39.7 | 32.6 |
| | | (−4.4) | (−4.9) | (−12.0) |
| **24** | 39.2 | 34.8 | 40.8 | 34.2 |
| | (−5.4) | (−9.8) | (−3.8) | (−10.4) |
| | | **(0.0)** | **(+6.5)** | **(+7.0)** |
| **25** | 39.9 | 34.6 | 34.2 | 26.6 |
| | (−4.7) | (−10.0) | (−10.4) | (−18.2) |
| | | **(−0.9)** | **(−0.8)** | **(−1.5)** |
| **26** | 33.7 | 27.6 | 35.6 | 27.8 |
| | (−10.9) | (−17.0) | (−9.0) | (−16.8) |
| | | **(−1.7)** | **(+6.8)** | **(+6.1)** |

[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. All $T_m$ values are given as averages of double determinations. $\Delta T_m$ values are given in brackets. $\Delta\Delta T_m$ values for duplex **x : y** (defined as $\Delta T_{m(x:y)}-(\Delta T_{m(x:23)} + \Delta T_{m(19:y)})$) are given in bold brackets.

interspacing base pairs, has virtually the opposite effect giving a thermal destabilisation of −9.8 °C, which corresponds roughly to the sum of the two monosubstituted duplexes. Moving the opposite nucleobases closer with either one interspacing base pair, *i.e.* **20** : **25**, or no interspacing base pair, *i.e.* **21** : **25** shows similar destabilisation of about −10 °C. Thus, in the latter three motifs the opposite nucleobases seem to operate independent of each other. This is emphasised by substituting with more nucleobases that are not positioned with two interspacing base pairs. Each nucleobase introduces roughly a −5.5 to −6 °C thermal destabilisation (*i.e.* **22** : **25** and **20** : **26**). When two interspacing base pairs are again present, as in **22** : **24**, **21** : **26** and in **22** : **26**, the decrease in $T_m$ is again smaller than expected just by the sum of double-headed moieties. In summary, the disubstituted duplex with opposite nucleobases and with two interspacing canonical base pairs appears to show a significant stabilisation that should be calibrated against the additive effects of introducing the independent double-headed nucleotides, which is roughly −10.0 °C. By this measure (as printed in bold in Table 3) the **21** : **24** motif appears stabilised by +6.5 °C by interactions between the opposite bases. The same is evident with **22** : **24**, **21** : **26** and **22** : **26** with similar stabilisation of +6–7 °C demonstrating the generality of this interaction.

These encouraging results were explored in another sequence context, in which incorporation of several modifications into the same strand using the configuration with two interspacing nucleotides in opposite strands was achieved, *i.e.* in the self-complementary sequence **27** (Table 1). Although this opens up the possibility of hairpin formation no biphasic transitions were observed on the UV-melting curves (ESI-Fig. 1†). Table 4 shows the melting temperatures obtained in high salt buffer. In this palindromic duplex, a single monomer modification (**28**) is seen to lower the thermal stability with 4 °C, which was about the same as for the corresponding modified duplex (**21** : **24**) with the non-self-complementary sequence ($\Delta T_m = -3.8$ °C) and less than expected by two independent moieties of the double-headed nucleoside. Although not directly comparable due to the lower total strand concentration (1.0 μM compared to 2.0 μM total concentration), the data does indicate consistency in the observed

**Table 4** Hybridisation data for the self-complementary DNA zipper[a]

| | **27** | **28** | **29** | **30** |
|---|---|---|---|---|
| | 25 | 21<br>(−4) | 18<br>(−7) | 18<br>(−7) |

[a] Melting temperatures ($T_m$ values/°C) recorded in a high salt buffer ($Na_2HPO_4$ (7.5 mM), NaCl (1.0 M), EDTA (0.1 mM), pH 7.0) using 1.0 μM concentrations of the strands. $\Delta T_m$ values are given in brackets.

correlation. Incorporation of two modifications three nucleotides apart in the same strand (**29**) results in a $\Delta T_m$ of −7 °C, which is almost twice the value for the singly modified strand as expected for an additive effect. Interestingly, however, incorporation of a third modification (**30**) between the latter two introduces no further destabilisation as $\Delta T_m$ again is −7 °C. The reason for this observation is unclear, but may indicate a change in conformation to better accommodate the extra thymines in the minor groove. Sequences composed of AT have been shown to display polymorphism with an anti-parallel Hoogsteen conformation in a crystalline state and a conformation with the classical Watson–Crick base pairs in solution.[25,26] Although there is no evidence of a Hoogsteen conformation in solution, a theoretical study indicate that the energy difference between the two conformations is small.[27] Thus, a change in conformation for this sequence is not unlikely and a Hoogsteen conformation with its narrower minor groove may indeed be better for accommodating the extra thymines. Regardless of the conformation the flattening out in the melting temperature drop, when installing more modifications does indicate that the sequence is capable of forming a quite stable duplex presumably with the thymines stacked in a zipper motif.

The duplex motif **21** : **24** was subjected to further UV-melting experiments to investigate the origins of this seemingly stabilised duplex. Firstly, the salt environment was varied from only 10 mM $Na^+$ to a high concentration at 1.0 M $Na^+$. As shown in Table 5, this did not effect the relative stability as compared with the reference duplex **19** : **23**. Likewise, adding 20 vol% isopropanol as cosolvent to lower the water activity[28] although indicating a positive effect of water only gave minor perturbation of the stability as compared to the reference, *i.e.* $\Delta\Delta T_m = -0.6$ °C. Although the effect is small, the binding of water seems more important in the modified duplex.

Next, the Van't Hoff thermodynamic parameters were determined using the $1/T_m$ *vs.* $\ln(C)$ method[28] as shown in Table 5 (see plots in ESI-Fig. 2†). From these two-state model dependent parameters, the duplex **21** : **24** shows the least favourable enthalpy of formation and any gain in stability seems to be entropy driven. In contrast with the thermal stability, the thermodynamic stability of the zipper motif duplex **21** : **24** appears to be slightly smaller than of the duplex **19** : **24**. In fact, the latter shows an even larger enthalpy of formation than the unmodified reference duplex **19** : **23** and **21** : **23** indicating that a sequence specific interaction with the extra nucleobase and the core duplex may also possibly be operating.

To investigate the molecular origin of the interactions in the minor groove, the motif **21** : **24** was subjected to a 1 ns molecular dynamics simulation, where the initial conformation adopts an A-form. Theoretically, this allows for a wide hydrated minor groove to close up and let the extra nucleobases displace water molecules in the minor groove as a transformation progresses towards a B-form.[29] Indeed, this is actually observed after approx. 0.5 ns of simulation. The final conformation stabilises in a B-type as shown from the structure averaged over the last 50 ps in Fig. 3. Hereby, an interaction between the extra nucleobases that appears to have a co-stacking nature is revealed. The distance between the two thymines is ~3.6 Å.

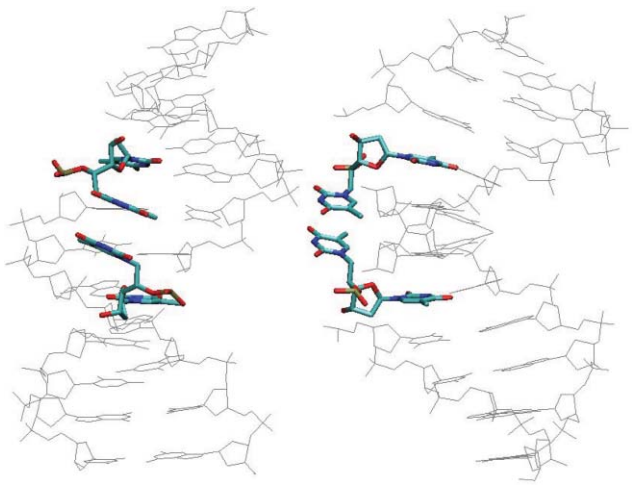


**Fig. 3** Average structure of duplex **21** : **24** from the final 50 ps in a 1 ns MD-simulation at 298 K. The central double-headed nucleotides are highlighted.

The present results demonstrate a strong interstrand interaction between two additional bases positioned in the minor groove. Thus, the orientation of the additional bases towards the minor groove by using the simple double-headed moiety **3** is strong and seems precisely suitable for a minor groove stacking interaction. Compared to **2**,[6] a much stronger base–base interaction in the minor groove is observed, and the interaction seems from the modelling experiment to be of a true π-stacking nature. With two moieties of **2**, the interstrand interaction was thermally weaker and of a perpendicular nature.[6] On the other hand, the thermal penalty of positioning one moiety of **2** with thymine as the additional base in a duplex is smaller (around −2 °C)[6] than the corresponding

**Table 5** Hybridisation data for the DNA zipper motif in different buffer systems[a]

| | 10 mM $Na^+$ | 110 mM $Na^+$ | 1.0 M $Na^+$ | 110 mM $Na^+$ in 20% *i*PrOH |
|---|---|---|---|---|
| **19** : **23** | 26.4 | 44.6 | 51.9 | 38.3 |
| **21** : **24** | 23.1<br>(−3.3) | 40.8<br>(−3.8) | 48.4<br>(−3.5) | 33.9<br>(−4.4) |

[a] Melting temperatures ($T_m$ values/°C), see Table 3.

penalty of **3** (around −5 °C). Nevertheless, adenine as the additional base in **2** is not accommodated very well with a penalty of −6–10 °C.[6] For the purpose of designing more complex structures with pair-wise interstrand communication in the minor groove, the 5′(*S*)-configured skeleton of **3** seems optimal. Similar moieties with other additional nucleobases or aromatic systems will therefore be studied. Recently, a somewhat similar interaction between pyrene moieties oriented towards the minor groove by the 2′-*N*-position of 2′-*N*-amino-LNA monomers has been studied.[30] Large increases in stability due to these interactions were observed in duplexes that are further stabilised by the conformationally restricted LNA-nucleotides.[30,31] The combination of 5′(*S*)-positioned additional bases and the duplex stabilising properties of LNA will be a matter of a future study.

For the influence on the secondary structure of TWJ′s, the influence of **3** is not optimal when the purpose is base–base interaction with core bases on the complementary strand. On the other hand, base–base interaction with a bulged region is indicated by modeling. Obviously, the length of the linker between the additional base and the 5′-carbon in **3** can be very important for this kind of interaction. The short methylene linker between the additional thymine and the nucleotide limits the flexibility and while being ideal for the zipper motif above, the modeling indicates it is too short for an ideal interaction in the studied junction motifs.

A hydrophobic zipper structure with 5′(*S*)-*C*-butyl or isopentyl 2′-deoxynucleotide moieties have been studied by Leumann and co-workers.[17,32] No stabilisation of duplexes was observed with this zipper structure. However, any favourable alkyl-zipper formation might be outweighed by a distortion in the hydration pattern of the duplex.[32] Although the zipper motif introduced in our case by the introduction of two opposite bases in the duplex **21** : **24** is hydrophilic, the penalty of positioning a base on the outside of a duplex might be due to a costly entropy of hydration as compared to a stacked situation.[33] In the zipper motif, the reduction of the water activity (Table 5) does seem to influence the zipper more than the unmodified duplex as the driving force for stacking could be diminished. The thermodynamic data (Table 6) also indicates that the driving force is entropic as the mono-modified duplexes **21** : **23** and **19** : **24** display similar or larger entropic penalties of duplex formation as compared to the unmodified **19** : **23**, while the zipper **21** : **24** displays a lower entropic penalty. The relative gain in thermal stability of approx. 6 °C by positioning two thymines in the minor groove, which is a remarkable result, strongly supports this. It arises from only one stacking interaction and only two pyrimidine moieties are involved. The effect seems additive when several of these base–base interactions are adjacently positioned in the minor groove.

**Table 6** Van't Hoff data for the DNA zipper duplexes[a]

| | $\Delta H°$ /kJ mol$^{-1}$ | $\Delta S°$ /J mol$^{-1}$ K$^{-1}$ | $\Delta G°$ (25 °C)/kJ mol$^{-1}$ |
|---|---|---|---|
| **19** : **23** | −351 | −986 | −57.0 |
| **21** : **23** | −336 | −960 | −49.8 |
| **19** : **24** | −392 | −1140 | −52.1 |
| **21** : **24** | −325 | −920 | −50.7 |

[a] Data obtained from $1/T_m$ *vs.* ln(*C*) plots at 110 mM $Na^+$, pH 7.0, see ESI-Fig. 2†.

## Conclusion

In conclusion, the incorporation of 5′(*S*)-*C*-(thymine-1-ylmethyl)thymidine **3** into duplexes has revealed the 5′(*S*)-*C*-position as ideal for the orientation of additional nucleobases towards the minor groove. Hence, efficient base–base stacking interactions are obtained in the minor groove with the double-headed nucleotide moieties positioned in opposite strand with two interspacing base pairs. This inter-strand communication can be an important feature in the design of functional nucleic acid structures and for the development of self-assembling nucleic acid nanostructures. Moreover, the right combination of double-headed moieties can reveal an even more delicate pattern of minor groove interactions, and a double-coding function of DNA using both an "inner" and an "outer" code can be envisioned.

## Experimental

All commercial reagents were used as supplied. All reactions were conducted under an argon atmosphere. Column chromatography was carried out using silica gel 60 (0.040–0.063 mm). HRMALDI mass spectra were recorded on an Ionspec Ultima Fourier Transform mass spectrometer with a DHB-matrix. NMR spectra were recorded on a 300 MHz Varian spectrometer. The values for $\delta$ are in ppm relative to tetramethylsilane as 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 external standard for the $^{31}$P-NMR spectra. Assignments of NMR spectra are based on ($^1$H,$^1$H)-COSY, ($^1$H,$^{13}$C)-hetcor and/or DEPT spectra following nucleoside naming style, *i.e.* the anomeric carbon is C-1′. No distinction has been made between the two thymine nucleobases.

### Preparation of 5′(*S*)-*C*-(thymin-1-ylmethyl)-3′-*O*-(*tert*-butyldimethylsilyl)thymidine (6)

To a solution of 5′-methylidenethymidine **5** (3.86 g, 10.9 mmol) in $CH_2Cl_2$ (100 mL) was added a solution of mCPBA (3.71 g, 16.5 mmol) in $CH_2Cl_2$ (20 mL). The mixture was stirred for 24 h and another portion of mCPBA (1.11 g, 4.9 mmol) in $CH_2Cl_2$ (20 mL) was added. After stirring for 24 h, the reaction mixture was extracted with a saturated aqueous solution of $NaHCO_3$ (200 mL), dried ($Na_2SO_4$) and concentrated under reduced pressure. The residue was purified by column chromatography (67–0% petrol ether in diethyl ether) to give the crude epoxide as a mixture of epimers (1.62 g) (*S* : *R* ~9 : 1). Thymine (2.71 g, 21.5 mmol) and $K_2CO_3$ (1.00 g, 7.4 mmol) was dissolved in DMF (20 mL) and the mixture was stirred at 85 °C for 30 min. A solution of the crude epoxide (1.16 g, 3.15 mmol) in DMF (12 mL) was added in portions of 1 mL over 3 h. The mixture was stirred at 85 °C for 1 h and then at room temperature for 12 h and neutralised with acetic acid. The mixture was filtered and concentrated under reduced pressure. The residue was purified by column chromatography (2–5% $CH_3OH$ in $CH_2Cl_2$) to give the product (0.60 g, 39%) as a white foam. $R_f$ 0.28 (10% EtOH in $CHCl_3$). $\delta_H$ (300 MHz, DMSO-$d_6$, $Me_4Si$) 11.30 (1H, s, NH), 11.22 (1H, s, NH), 7.88 (1H, d, *J* 1.2 Hz, H-6), 7.39 (1H, d, *J* 0.9 Hz, H-6), 6.21 (1H, dd, *J* 6.3 Hz, *J* 7.5 Hz, H-1′), 5.61 (1H, d, *J* 5.4 Hz, OH), 4.45 (1H, dd, *J* 2.4 Hz, *J* 2.7 Hz, H-3′), 3.95 (1H, m, H-5′), 3.78–3.64 (3H, m, H-4′, H-6′), 2.22–2.02

(2H, m, H-2′), 1.80 (1H, br s, $CH_3$), 1.75 (1H, br s, $CH_3$), 0.87 (9H, s, $SiC(CH_3)_3$), 0.09 (6H, s, $Si(CH_3)_2$). $\delta_C$ (75 MHz, DMSO-$d_6$, $Me_4Si$) 164.3, 163.6 (C-4), 151.1, 150.4 (C-2), 142.3, 136.1 (C-6), 109.4, 107.7 (C-5), 87.1 (C-4′), 83.9 (C-1′), 73.2 (C-3′), 67.2 (C-5′), 50.7 (C-6′), 40.1 (C-2′), 25.6 ($C(CH_3)_3$), 17.6 ($C(CH_3)_3$), 12.3, 11.9 ($CH_3$), −4.8, −5.0 ($Si(CH_3)_2$). HiRes MALDI FT-MS *m/z* (M + Na) found/calc. 517.2082/517.2089.

### Preparation of 5′(*S*)-*C*-(thymin-1-ylmethyl)-5′-*O*-(4,4′-dimethoxytrityl)thymidine (8)

Compound **6** (100 mg, 0.202 mmol) was dissolved in $CH_3CN$ (5 mL) and pyridine (5 mL). The mixture was stirred at room temperature and DMTCl (0.25 g, 0.74 mmol) was added. The mixture was stirred for 24 h and another portion of DMTCl (0.47 g) was added. After another 24 h, DMTCl (0.53 g) together with another $CH_3CN$ (5 mL) and pyridine (5 mL) was added and after stirring for another 48 h, another portion of DMTCl (0.96 g). After stirring for one week in total, the mixture was concentrated under reduced pressure, and the residue was purified by column chromatography to give the crude product **7** (101 mg). This was dissolved in THF (1 mL) and a 1 M solution of TBAF in THF (0.1 mL) was added. The reaction mixture was stirred for 5 h and directly purified by column chromatography (0.2% pyridine and 50% diethyl ether in petroleum ether). The product was dissolved in $CH_2Cl_2$ and precipitated in *n*-hexane to give the product (41 mg, 41%) as a white compound. $\delta_H$ (300 MHz, $CDCl_3$, $Me_4Si$) 8.40 (1H, br s, NH), 8.34 (1H, br s, NH), 7.65 (1H, s, H-6), 7.45–7.15 (9H, m, Ph), 6.88–6.81 (4H, m, Ph), 6.74 (1H, s, H-6), 6.27 (1H, t, *J* 6.9 Hz, H-1′), 4.30 (1H, m, H-3′), 4.02 (1H, m, H-5′), 3.85–3.53 (9H, m, H-4′, H-6′, $OCH_3$), 2.33–2.23 (3H, m, H-2′, OH), 1.85 (1H, s, $CH_3$), 1.79 (1H, s, $CH_3$).

### Preparation of 5′(*S*)-*C*-(thymin-1-ylmethyl)-5′-*O*-(4,4′-dimethoxytrityl)-3′-*O*-(*P*-β-cyanoethoxy-*N*,*N*-diisopropylaminophosphinyl)thymidine (9)

Compound **8** was dissolved in $CH_2Cl_2$ (1.0 mL) and DIPEA (0.1 mL). *N*,*N*-Diisopropylamino-β-cyanoethylphosphinochloridite (0.03 mL, 0.14 mmol) was added and the mixture was stirred for 1.5 h and directly purified by column chromatography (0.25% $Et_3N$ in EtOAc). The residue was dissolved in $CH_2Cl_2$ and precipitated in *n*-hexane to give the product (21.9 mg, 58%). $\delta_P$ (121.5 MHz, $CDCl_3$) 151.19, 150.28. HiRes MALDI FT-MS *m/z* (M + Na) found/calc. 905.3609/905.9266.

### Preparation of 5′(*S*)-*C*-(thymin-1-ylmethyl)-5′-*O*-pixyl-3′-*O*-(*tert*-butyldimethylsilyl)thymidine (10)

Compound **6** (60 mg, 0.12 mmol) was dissolved in anhydrous pyridine (1.5 mL) and pixyl chloride (121 mg, 0.41 mmol) was added. The mixture was stirred for 12 h and concentrated under reduced pressure. The residue was purified by column chromatography (1% $Et_3N$ in EtOAc) to give the product (72 mg, 79%) as a white compound. $R_f$ 0.23 (EtOAc). $\delta_H$ (300 MHz, $CDCl_3$, $Me_4Si$) 9.14 (1H, m, NH), 8.60 (1H, m, NH), 7.76 (1H, d, *J* 1.2 Hz, H-6), 7.45–7.26 (9H, m, Ph), 7.11–6.97 (4H, m, Ph), 6.75 (1H, d, *J* 0.9 Hz, H-6), 6.25 (1H, dd, *J* 4.8 Hz, 9.3 Hz, H-1′), 3.77 (1H, m, H-5′), 3.53 (1H, d, *J* 4.8 Hz, H-3′), 3.47 (1H, br s, H-4′), 3.33–3.19 (2H, m, H-6′), 2.05–2.00 (4H, m, H-2′, $CH_3$), 1.88–1.79 (4H, m, H-2′, $CH_3$), 0.74 (9H, s, $SiC(CH_3)_3$), −0.12 (3H, s, $Si(CH_3)$), −0.23 (3H, s, $Si(CH_3)$. $\delta_C$ (75 MHz, $CDCl_3$, $Me_4Si$) 164.0, 163.9 (C-4), 152.2, 152.1 150.5 (C-2, Ph), 146.9, 141.2, 135.6, 131.2, 130.7, 130.5, 130.4, 128.1, 127.6, 123.8, 123.1, 122.6, 117.7, 117.3 (C-6, Ph), 111.3, 110.6 (C-5), 86.9, 85.1, 78.5, 73.5, 70.3 (C-1′, C-3′, C-4′, C-5′, $CPh_3$), 49.3 (C-6′), 41.2 (C-2′), 25.7 ($C(CH_3)_3$), 17.8 ($C(CH_3)_3$), 12.9, 12.4 ($CH_3$), −4.5, −4.8 ($Si(CH_3)_2$). HiRes MALDI FT-MS *m/z* (M + Na) found/calc. 773.2946/773.2977.

### Preparation of 5′(*S*)-*C*-(thymin-1-ylmethyl)-5′-*O*-pixylthymidine (11)

Compound **10** (70 mg, 0.093 mmol) was dissolved in THF (1 mL) and a 1 M solution of TBAF in THF (0.1 mL) was added. The mixture was stirred for 4 h and concentrated under reduced pressure. The residue was purified by column chromatography (0.5% $Et_3N$ in EtOAc) to give the product (32.2 mg, 54%). $R_f$ 0.16 (EtOAc). $\delta_H$ (300 MHz, $CDCl_3$, $Me_4Si$) 9.04 (2H, br s, NH), 7.49–7.27 (10H, m, Ph, H-6), 7.10–6.99 (4H, m, Ph), 6.68 (1H, d, *J* 0.9 Hz, H-6), 6.09 (1H, dd, *J* 6.3, 7.2 Hz, H-1′), 3.79 (1H, m), 3.52 (2H, br s), 3.34–3.18 (2H, m), (H-3′, H-4′, H-5′, H-6′), 2.14–1.90 (5H, m, H-2′, $CH_3$), 1.80 (3H, s, $CH_3$). $\delta_C$ (75 MHz, $CDCl_3$, $Me_4Si$) 164.3, 163.8 (C-4), 152.4, 152.2, 151.3, 150.5 (C-2, Ph), 146.6, 141.2, 131.2, 135.5, 130.8, 130.5, 130.4, 128.1, 127.7, 127.6, 123.9, 123.7, 123.1, 122.8, 117.7, 117.2 (C-6, Ph), 111.4, 110.6 (C-5), 85.2, 83.9, 78.1, 70.8, 69.5 (C-1′, C-3′, C-4′, C-5′, $CPh_3$), 49.4 (C-6′), 39.8 (C-2′), 13.0, 12.4 ($CH_3$).

### Preparation of 5′(*S*)-*C*-(thymin-1-ylmethyl)-5′-*O*-pixyl-3′-*O*-(*P*-β-cyanoethoxy-*N*,*N*-diisopropylaminophosphinyl)thymidine (12)

Compound **11** (32 mg, 0.051 mmol) was dissolved in $CH_2Cl_2$ (3 mL), DIPEA (0.04 ml, 0.23 mmol) and $CHCl_3$ (3 mL). The solution was stirred at room temperature and *N*,*N*-diisopropylamino-β-cyanoethylphosphinochloridite (0.03 mL, 0.13 mmol) was added. The reaction mixture was stirred for 3 h and concentrated under reduced pressure. The residue was purified by column chromatography (1% $Et_3N$ in EtOAc) as an oil. This was dissolved in $CH_2Cl_2$ (0.2 mL) precipitated in *n*-hexane (1.5 mL), cooled over night at −20 °C and centrifuged to give the product (27 mg, 64%) as a white solid. $R_f$ 0.40 (EtOAc). $\delta_P$ (121.5 MHz, $CDCl_3$) 151.98, 151.14. HiRes MALDI FT-MS *m/z* (M + Na) found/calc. 860.3257/860.3269.

### Preparation of 5′(*S*)-*C*-(thymin-1-ylmethyl)-3′,5′-*O*-(1,1,3,3-tetraisopropylsiloxan-1,3-diyl)thymidine (13)

Compound **8** (15 mg, 0.022 mmol) was dissolved in $CDCl_3$ (0.5 mL) and added 2 drops of DCA. The mixture was stirred for 2 min. and then added pyridine (1 mL) followed by an excess of TIPDSCl (0.2 mL, 0.7 mmol). The mixture was stirred for 48 h, added EtOH (0.1 mL) and concentrated under reduced pressure. The residue was purified by column chromatography (25% EtOAc in petrol ether and then 5% $CH_3OH$ in $CH_2Cl_2$) to give the product (5.1 mg, 39%) as a white foam. $\delta_H$ (300 MHz, $CDCl_3$, $Me_4Si$) 8.60 (1H, s, NH), 8.48 (1H, s, NH), 7.31 (1H, d, *J* 0.9 Hz, H-6), 7.09 (1H, d, *J* 0.9 Hz, H-6), 6.17 (1H, dd, *J* 1.5 Hz, 7.8 Hz, H-1′), 4.52–4.43 (2H, m, H-3′, H-5′), 3.99–3.84 (2H, m, H-6′), 3.63 (1H, dd, *J* 2.7 Hz, 8.4 Hz, H-4′), 2.51 (1H, m, H-2′), 2.23 (1H, m, H-2″), 1.94 (3H, d, *J* 0.9 Hz, $CH_3$), 1.92 (3H, d, *J* 0.9 Hz, $CH_3$),

1.12–0.97 (28H, m, $CH(CH_3)_2$). $\delta_C$ (75 MHz, $CDCl_3$, $Me_4Si$) 163.9, 163.4 (C-4), 150.8, 150.0 (C-2), 141.7, 134.7 (C-6), 111.2, 110.3 (C-5), 83.5, 83.1, 67.5, 65.8 (C-1′, C-3′, C-4′, C-5′), 51.5 (C-6′), 39.0 (C-2′), 17.40, 17.33, 17.30, 17.06, 16.95, 16.87, 16.78 ($CH(CH_3)_2$), 13.46, 12.98, 12.86, 12.75, 12.56, 12.25 ($CH(CH_3)_2$, T-$CH_3$). HiRes MALDI FT-MS *m/z* (M + Na) found/calc. 645.2724/645.2746. NOE-difference; (H-1′ irradiated) H-4′ (2%), H-2″ (4%), (H-2′ irradiated) H-1′ (1%), H-2″ (18%), H-3′ and/or H-5′ (7%), (H-2″ irradiated) H-1′ (10%), H-2′ (18%), H-4′ (4%), (H-4′ irradiated) H-1′ (2%), H-3′ and/or H-5′ (10%).

### Preparation of 5′(*R*)-*C*-(thymin-1-ylmethyl)-3′-*O*-(*tert*-butyldimethylsilyl)thymidine (14)

Compound **6** (27 mg, 0.055 mmol) was dissolved in $CH_2Cl_2$ (2 mL) and added DMAP (79 mg, 0.65 mmol). The mixture was stirred at 0 °C, and $Tf_2O$ (0.02 mL, 0.12 mmol) was added slowly. The mixture was stirred for 1.5 h at 0 °C, and water (5 mL) was added. The mixture was extracted with $CH_2Cl_2$ (20 mL) and concentrated under reduced pressure. The residue was dissolved in ethanol (2 mL) and a 1 M aqueous solution of NaOH (1 mL) was added. The solution was stirred for 3 h and then neutralised with acetic acid. The mixture was extracted with $CH_2Cl_2$ (4 × 5 mL), dried ($Na_2SO_4$) and concentrated under reduced pressure. The residue was purified by column chromatography (1% methanol in diethyl ether and then 10% EtOAc in $CH_2Cl_2$) to give the product (10.9 mg, 40%) as a white foam. $\delta_H$ (300 MHz, DMSO-$d_6$, $Me_4Si$) 11.34 (1H, s, NH), 11.26 (1H, s, NH), 7.54 (1H, d, *J* 1.2 Hz, H-6), 7.43 (1H, d, *J* 0.9 Hz, H-6), 6.22 (1H, dd, *J* 5.7 Hz, 9.3 Hz, H-1′), 5.61 (1H, d, *J* 6.0 Hz, OH), 4.54 (1H, d, *J* 3.6 Hz, H-3′), 4.05 (1H, dd, *J* 2.1 Hz, 13.8 Hz, H-6′), 3.71–3.62 (2H, m, H-4′, H-5′), 3.30 (m, H-6″), 2.12 (1H, m, H-2′), 2.00 (1H, m, H-2″), 1.83 (3H, s, $CH_3$), 1.75 (3H, s, $CH_3$), 0.88 (9H, s, $SiC(CH_3)_3$), 0.11 (6H, s, $Si(CH_3)_2$). $\delta_C$ (75 MHz, DMSO-$d_6$, $Me_4Si$) 164.4, 163.7 (C-4), 151.1, 150.5 (C-2), 143.0, 135.5 (C-6), 110.2, 107.5 (C-5), 88.4, 84.3, 72.7, 68.5 (C-1′, C-3′, C-4′, C-5′), 50.9 (C-6′), 38.5 (C-2′), 25.7 ($C(CH_3)_3$), 17.7 ($C(CH_3)_3$), 12.1, 12.0 ($CH_3$), −4.8 ($Si(CH_3)_2$). HiRes MALDI FT-MS *m/z* (M + Na) found/calc. 517.2071/517.2089.

### Preparation of 5′(*R*)-*C*-(thymin-1-ylmethyl)-3′,5′-*O*-(1,1,3,3-tetraisopropyldisiloxan-1,3-diyl)thymidine (15)

Compound **13** (6.7 mg, 0.016 mmol) was dissolved in $CH_2Cl_2$ (0.5 mL), and a 0.01 M solution of TBAF in THF (0.15 mL) was added. The mixture was concentrated under reduced pressure and redissolved in THF (0.4 mL) and a 1 M solution of TBAF in THF (0.02 mL) was added. The mixture was stirred at 40 °C for 2 h and the solvent was removed under reduced pressure. The residue was dissolved in pyridine (1.0 mL) and an excess of TIPDSCl (0.1 mL, 0.4 mmol) was added. The mixture was stirred for 48 h at room temperature. Ethanol (0.2 mL) was added, and the mixture was concentrated under reduced pressure. The residue was purified by column chromatography (5% methanol in $CH_2Cl_2$ and then 10% EtOAc in $CH_2Cl_2$) to give the product (4.5 mg, 53%) as a white foam. $\delta_H$ (300 MHz, $CDCl_3$, $Me_4Si$) 8.78 (1H, s, NH), 8.60 (1H, s, NH), 7.21 (1H, s, H-6), 7.07 (1H, s, H-6), 6.04 (1H, t, *J* 6.3 Hz, H-1′), 4.61 (1H, m, H-3′), 4.53 (1H, d, *J* 14.1 Hz, H-6′), 4.31 (1H, t, *J* 9.3 Hz, H-5′), 3.63 (1H, dd, *J* 5.1 Hz, 9.3 Hz, H-4′), 3.18 (1H, dd, *J* 9.3 Hz, 14.1 Hz, H-6″), 2.47–2.40 (2H, m, H-2′), 1.94 (3H, s, $CH_3$), 1.89 (3H, s, $CH_3$), 1.10–0.87 (28H, m, $CH(CH_3)_2$). $\delta_C$ (75 MHz, $CDCl_3$, $Me_4Si$) 164.0, 163.5 (C-4), 150.9, 150.0 (C-2), 142.7, 135.7 (C-6), 111.4, 109.5 (C-5), 87.5, 85.7, 74.3, 73.1 (C-1′, C-3′, C-4′, C-5′), 53.5 (C-6′), 41.0 (C-2′), 17.44, 17.36, 17.31, 17.06, 16.97, 16.85, 16.77 ($CH(CH_3)_2$) 13.19, 12.99, 12.74, 12.56, 12.41, 12.08 ($CH(CH_3)_2$, T-$CH_3$). HiRes MALDI FT-MS *m/z* (M + Na) found/calc. 645.2757/645.2746. NOE-difference; (H-1′ irradiated) H-4′ (3%), H-2′ and/or H-2″ (6%), (H-2′ and H-2″ irradiated) H-1′ (12%), H-3′ (9%), H-4′ (1%), (H-3′ irradiated) H-5′ (7%), H-4′ (2%), H-2′ and/or H-2″ (6%), (H-4′ irradiated) H-1′ (2%), H-3′ (1%), H-5′ (1%), H-6″ (1%), (H-5′ irradiated) H-3′ (7%), H-4′ (2%), H-6″ (1%), (H-6″ irradiated) H-6′ (24%), H-4′ (4%).

### Oligosynthesis and melting experiments

The oligodeoxynucleotides were synthesised using an automated Expedite 8909 Nucleic acid synthesis system following the phosphoramidite approach. Synthesis of oligonucleotides **16–30** was performed on a 0.2 μmol scale by using 2-cyanoethyl phosphoramidites of standard 2′-deoxynucleosides in combination with the modified phosphoramidites **9** and **12**. The synthesis followed the regular protocol employing standard CPG supports and 4,5-dicyanoimidazole as the activator except for the modified amidites, which were manually coupled using 0.05 M amidite (25 eq.) and 0.5 M pyridinium chloride (250 eq.) as activator in $CH_3CN$ for 10 min. before oxidised with *tert*-butyl hydroperoxide for 5 min. The following unmodified amidite was coupled under standard conditions expect for a prolonged coupling time of 15 min. The coupling yields for **9** and **12** in combination with the following unmodified amidite were in the range of 80–90%. The 5′-*O*-DMT-ON oligonucleotides were removed from the solid support by treatment with concentrated aqueous ammonia at 55 °C for 16 h, which also removed the protecting groups. The oligonucleotides were purified by reversed-phase HPLC on a Waters 600 system using a $X_{terra}$ prep MS $C_{18}$; 10 μm; 7.8 × 150 mm column; Buffer: 0.05 M triethylammonium acetate; 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 over night. After dissolution in double distilled water, the concentrations were determined spectrometrically at 260 nm and at 80 °C in the pH 7.0 buffer assuming an extinction coefficient for the modified double-headed nucleoside 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 10 and 80 °C after denaturation at 80 °C except for the experiments with $MgCl_2$, where precipitation was observed above 60 °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.

### Molecular dynamics simulation

MD simulation was computed using the amber 94/99 force field supplied with the AMBER 7.0 software[34] on an Intel 32 bit machine compiled with a g77compiler. The starting coordinates were generated using *nucgen* and the Arnott A-DNA parameters. The

modified thymidine residues were incorporated using an amber prep file based on the standard thymidine nucleoside. The *pro-S* H-5′ was replaced by a thymin-1-ylmethyl group setting the thymin-1-yl group in an *anti* periplanar conformation to the 5′-hydroxyl moiety to fit into the initial duplex. The internal coordinates for this moity was obtained from a QM optimisation in rhf/6-31G(d) using *Gaussian 03*.[35] Partial charges were calculated using *resp* with a two stage fit, first restraining the thymidine moiety to assume the force field charges and secondly also restraining the methyl and methylene protons to be equivalent, respectively. The electrostatic potential was generated from an rhf/6-31G(d) wave function.

The duplex was neutralised with 20 $Na^+$ ions and solvated with 3555 water molecules by placing a 8.0 Å TIP3P truncated octahedron around the duplex. Using *Sander* with PBC and PME, the system was then allowed to relax with 1000 minimisation steps using a 10 Å cutoff for the non-bonded potential terms and a 500 kcal $mol^{-1}$ $Å^{-2}$ restraint on the duplex. The hydrogens were then constrained using the SHAKE algorithm and the whole system relaxed by 2500 minimisation steps. MD simulation was then initiated. While restraining the duplex with a 10 kcal $mol^{-1}$ $Å^{-2}$ force constant, the solvent environment was allowed to warm up to 298 K for 10 ps at constant volume using the Berendsen thermostat before the system was allowed to equilibrate unrestrained for 50 ps at 298 K and 1 atm (with time constants 1 ps and 2 ps, respectively). Finally, the dynamics of the system was simulated for 1 ns with time steps of 0.2 ps. The coordinates from the final 50 ps were used to produce an average structure of the final duplex (Fig. 3).

Root mean squared deviation of the fit of the backbone to the initial Arnott A-DNA structure (ESI-Fig. 3†) shows the duplex conformation to be stabilised after *ca.* 500 ps. Similarly, the alignment of the extra nucleobases electric dipoles (ESI-Fig. 4†) indicates a very stable "sandwich" structure for the two rings after some 600 ps.

The coordinates for the TWJ was made by replacing DC5 and DC6 with **3**, replacing DG15 and DG28 with DA and DC16 with DT in the 1EKW.pdb file.[22] Then the structure was neutralised with 28 $Na^+$ and solvated by placing it in a truncated octahedron with 4022 TIP3P water molecules. The MD-simulation simulation protocol is identical to the simulation used for the duplex modelling. A total time of 100 ps was allowed for warm up and the final 10 ps used to generate an average structure (Fig. 2).

## 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

1 C. J. Leumann, *Bioorg. Med. Chem.*, 2002, **10**, 841.
2 J. Kurreck, *Eur. J. Biochem.*, 2003, **270**, 1628.
3 J. Wengel, *Org. Biomol. Chem.*, 2004, **2**, 277.
4 S. L. Pedersen and P. Nielsen, *Org. Biomol. Chem.*, 2005, **3**, 3570.
5 Before our paper (ref. 4), ODN's containing an acyclic achiral nucleoside analogue with two nucleobases (A or T) have been studied; T. Wu, M. Froeyen, G. Schepers, K. Mullens, J. Rozenski, R. Busson, A. Van Aerscot and P. Herdewijn, *Org. Lett.*, 2004, **6**, 51.
6 T. Wu, K. Nauwelaerts, A. Van Aerschot, M. Froeyen, E. Lescrinier and P. Herdewijn, *J. Org. Chem.*, 2006, **71**, 5423.
7 A. M. Costa, M. Faja, J. Farràs and J. Villarasa, *Tetrahedron Lett.*, 1998, **39**, 1835.
8 R. Fecher, K. H. Boswell, J. J. Wittick and T. Y. Shen, *J. Am. Chem. Soc.*, 1970, **92**, 1400.
9 M. W. Logue and N. J. Leonard, *J. Am. Chem. Soc.*, 1972, **94**, 2842.
10 B. Kašnar, V. Škarić, B. Klaić and M. Žinić, *Tetrahedron Lett.*, 1993, **34**, 4997.
11 J. Fensholdt and J. Wengel, *Acta Chem. Scand.*, 1996, **50**, 1157.
12 G. Wang and P. J. Middleton, *Tetrahedron Lett.*, 1996, **37**, 2739.
13 G. Wang and P. J. Middleton, *Nucleosides Nucleotides*, 1998, **17**, 1033.
14 V. Banuls and J.-M. Escudier, *Tetrahedron*, 1999, **55**, 5831.
15 A. M. Sørensen and P. Nielsen, *Org. Lett.*, 2000, **2**, 4217.
16 A. M. Sørensen, K. E. Nielsen, B. Vogg, J. P. Jacobsen and P. Nielsen, *Tetrahedron*, 2001, **57**, 10191.
17 H. Trafelet, E. Stulz and C. Leumann, *Helv. Chim. Acta*, 2001, **84**, 87.
18 P. K. Sharma, B. H. Mikkelsen, M. S. Christensen, K. E. Nielsen, C. Kirchhoff, S. L. Pedersen, A. M. Sørensen, K. Østergaard, M. Petersen and P. Nielsen, *Org. Biomol. Chem.*, 2006, **4**, 2433.
19 P. Børsting, K. E. Nielsen and P. Nielsen, *Org. Biomol. Chem.*, 2005, **3**, 2183.
20 N. B. Leontis, W. Kwok and J. S. Newman, *Nucleic Acids Res.*, 1991, **19**(4), 759.
21 B. Liu, N. B. Leontis and N. Seeman, *Nanobiology*, 1994, **3**, 177.
22 V. Thiviyanathan, B. A. Luxon, N. B. Leontis, N. Illangasekare, D. G. Donne and D. G. Gorenstein, *J. Biomol. NMR*, 1999, **14**, 209.
23 J.-L. Mergny and L. Lacroix, *Oligonucleotides*, 2003, **13**, 515.
24 P. G. Higgs, *Q. Rev. Biophys.*, 2000, **33**, 199.
25 N. G. A. Abrescia, C. González, C. Gouyette and J. A. Subirana, *Biochemistry*, 2004, **43**, 4092.
26 D. De Luchi, V. Tereshko, C. Gouyette and J. A. Subirana, *ChemBioChem*, 2006, **7**, 585.
27 E. Cubero, N. G. A. Abrescia, J. A. Subirana, F. J. Luque and M. Orozco, *J. Am. Chem. Soc.*, 2003, **125**, 14603.
28 C. H. Spink and J. B. Chaires, *Biochemistry*, 1999, **38**, 496.
29 T. E. Cheatham and P. A. Kollman, *J. Mol. Biol.*, 1996, **259**, 434.
30 P. J. Hrdlicka, B. R. Babu, M. D. Sørensen and J. Wengel, *Chem. Commun.*, 2004, 1478.
31 S. K. Singh, P. Nielsen, A. A. Koshkin and J. Wengel, *Chem. Commun.*, 1998, 455.
32 H. Trafelet, S. P. Parel and C. Leumann, *Helv. Chim. Acta*, 2003, **86**, 3671.
33 W. Sanger, *Principles of Nucleic Acid Structure*, 1984, Springer-Verlag, New York.
34 D. A. Case, D. A. Pearlman, J. W. Caldwell, T. E. Cheatham, III, J. Wang, W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Merz, R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, V. Tsui, H. Gohlke, R. J. Radmer, Y. Duan, J. Pitera, I. Massova, G. L. Seibel, U. C. Singh, P. K. Weiner and P. A. Kollman, 2002, *AMBER7*, University of California, San Francisco.
35 *Gaussian 03, Revision C.02*, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, Gaussian, Inc.: Wallingford CT, 2004.