Published in final edited form as:

Angew Chem Int Ed Engl. 1996 April 19; 35(7): 743–746. doi:10.1002/anie.199607431.

Formation of Stable DNA Loops by Incorporation of Nonpolar, Non-Hydrogen-Bonding Nucleoside Isosteres**

Xiao-Feng Ren, Barbara A. Schweitzer, Charles J. Sheils, and Eric T. Kool [Prof.]*

Department of Chemistry, University of Rochester, Rochester, NY 14627 (USA)

Keywords

base stacking; DNA; hydrogen bonds; nucleosides; triple helices

Hairpin loops are ubiquitous structures found in folded RNA and DNA sequences in nature. Recent reports have noted special stability for certain loop sequences. For example, RNA tetranucleotide loops (tetraloops) having the sequence GNRA and UUCG are more stable than other tetraloops and are highly conserved in nature.[1,2] DNA tetraloops having the sequence GAAA, in certain contexts, are unusually stable in duplex DNA;[3] in triplex structures the loop sequence CTTTG has been reported to be especially stable as well.[4] In most or all of these cases, structural study has either implicated or identified intraloop hydrogen bonds between bases and/or phosphates as important structural features. It has long been recognized, however, that hydrogen bonds in aqueous solution are weak. In this regard a detailed study of hydrogen bonding in the GAAA tetraloop in RNA concluded that single hydrogen bonds contribute relatively little to the overall stability.[5] While less well understood, base stacking is also known to be at least as important a contribution to nucleic acid stability as hydrogen bonding.[6–8] Studies of loop structures in RNA and DNA have not generally addressed the relative importance of base stacking and hydrogen bonding in stabilizing such loops.

We have undertaken a program to design and synthesize non-hydrogen-bonding nucleoside analogues to be used as probes of the biological noncovalent interactions of oligonucleotides and nucleic acids in general.[9] We now describe the incorporation of three such isosteres into DNA loops, and we have found that such substitutions can lead to significant stabilization of double- and triple-helical folded structures in DNA.

Nucleosides **1** (F), **2** (B), and **3** (D) were designed to act as nearly perfect isosteres for the natural nucleosides thymidine (T) and deoxyadenosine (A) (Fig. 1). However, in contrast to the natural nucleosides, analogues **1–3** have little or no hydrogen bonding capability.[9] Compounds **1–3** are considerably more hydrophobic and less polar than the natural structures, and evidence from dangling-base studies in our laboratory indicate that they are more efficient at base stacking as well.[9b,10] To test the effects of their properties on folded DNA structures, we incorporated **1–3** at loop positions into DNA hairpin-forming sequences and evaluated the thermal and thermodynamic stabilities of the folded structures.

Nucleosides **1–3** were synthesized by using previously described methods.[9] The primary products of coupling of the bases of **1** and **2** with the α -chlorodeoxyribose toluoyl ester synthon

**This work was supported by the National Institutes of Health (GM 52956). E. T. K also acknowledges a Camille and Henry Dreyfus Teacher-Scholar Award and an Alfred P. Sloan Foundation Fellowship.

© VCH Verlagsgesellschaft mbH, D-69451 Weinheim, 1996

*Fax: Int. code +(716)473-6889, etk@etk.chem.rochester.edu.

are α -anomers,[11] which were converted to β -anomers by acid-catalyzed epimerization. [11b] Compounds **1–3** were unknown previous to our reports, although two other substituted indole nucleoside derivatives have recently been described.[12] The coupling of **1–3** in oligodeoxynucleotides using standard phosphoramidite chemistry was efficient, with stepwise yields of >95 % (trityl cation monitoring). For comparison we also synthesized analogous sequences containing the corresponding natural nucleosides at the same positions (see Tables 1 and 2).

We first studied the effects of **1** (F), **2** (B), and **3** (D) in a duplex DNA hairpin loop, by using two self-complementary sequences of eight base pairs bridged by a tetranucleotide loop consisting of four identical residues. Previous studies of this sequence by Breslauer et al.[13] showed that, of sequences containing A₄, C₄, G₄, and T₄ loops, the most stable was the T₄ case, and the least stable, the A₄ case. We studied the properties of sequences containing F₄, B₄, and D₄ loops, and we synthesized the natural T₄ and A₄ analogues for comparison.

Table 1 shows data obtained for the duplex hairpin sequences. Buffer conditions were pH 7.0 (10 mM phosphate) with 100 mM NaCl. All five sequences gave melting transitions that were independent of concentration, confirming the expected hairpin structure as opposed to intermolecular complexes. The data show that the modified hairpin sequences containing **1** (F) and **2** (B) are significantly more stable than the T₄ analogue. The melting temperature (T_m) for the modified F₄ sequence, which contains difluorotoluene nucleosides, is 78.5 °C, or 10.8 °C higher than that of the T₄ loop structure, which is the most stable tetraloop containing four identical natural nucleosides.[13] The four hairpin structures with natural nucleosides (A₄, T₄, C₄, G₄) have T_m values lying in a narrow range between 62 and 68 °C.[13] Free energy values obtained by curve fitting show that the modified loop containing **1** is thermodynamically more stable than the most stable natural loop (the T₄ loop) (Table 1). Similarly, the B₄ hairpin sequence, with trimethylphenyl nucleosides, is almost as stable as the sequence with the F₄ loop.

Interestingly, just as the pyrimidine isosteres give more stable loops than a natural pyrimidine, we find that the purine analogue **3** is more stabilizing than its deoxyadenosine counterpart ($\Delta T_m = 4.5$ K). As was reported by Breslauer et al.,[13] we find that the T₄ loop is more stabilizing than the A₄ loop. Moreover, we find that the nonnatural pyrimidine isosteres **1** and **2** are more stabilizing than the purine isostere **3**. This difference in the natural hairpin structures has been attributed to cross-loop hydrogen bonding. Since in the nonnatural cases this difference cannot be due to hydrogen bonding, it seems likely that steric differences in the pyrimidine and purine analogues may play an important role both in the natural and nonnatural cases.

Because the nonpolar thymidine isostere **1** was found to be the better stabilizing of the two analogues and because it is the best isosteric substitution, we tested it in a triple helix bridging loop. Such folded triplex-forming oligonucleotides are finding increasing use as ligands for single-stranded nucleic acids.[14] The stability of the loops in such a complex depends largely on the nature of the first and last bases in the loop.[4] The nucleoside **1** was therefore placed at these positions in a pentanucleotide loop bridging the pyrimidine strands in an eight-base triplex. For comparison we synthesized the natural T₅-bridged triplex (see Table 2). We examined both 5'- and 3'-type loop orientations [4] at pH 7.0 with 100 mM NaCl and 10 mM MgCl₂.

The nonpolar nucleoside **1** gives complexes that are both thermally and thermodynamically more stable than structures containing the natural nucleosides (Table 2). For example, the hydrophobic nucleoside F gives complexes that are 4.8–5.4 K more thermostable than the T,T-substituted oligonucleotides, and free energies are 1.8–2.2 kcal mol⁻¹ more favorable. Previous

studies carried out with natural bases in this loop[4] have shown that the T,T case is intermediate in stability; the C,C loop is the least stable and the G,G loop the most stable. The G,G loop was found to be more stable than the T,T analogue by 1.5 kcal mol⁻¹. [4] Thus our F,F loop is more stable than all four cases doubly substituted with the natural bases.

Previous studies have described the use of simple nonnucleotide linkers as loop replacements in nucleic acids.[15] Such loops are often less stabilizing than natural nucleotide loops; this may be due both to the greater rigidity of nucleotide loops relative to the highly flexible nonnucleotide linkers and to favorable base stacking interactions in natural loops, which are missing in simple linkers. The nucleoside analogues in this study are designed to stabilize nucleic acid helices as a result of both rigidity and base stacking. In contrast to natural DNAs, loops containing nucleosides **1–3** have the added advantage of being resistant to degradation by nucleases (data not shown). In addition, such nonpolar nucleosides have been shown to pair with low affinity and low specificity with natural sequences [9b] and thus would not interfere with the expected pairing in biological systems.

The present results on the incorporation of hydrophobic nucleoside analogues in oligonucleotide loops show that in general these compounds are significantly stabilizing. Because they do not undergo measurable hydrogen bonding interaction,[9b] it would seem that the best explanation for this behavior is superior base stacking propensity. This suggests a general strategy for stabilizing nucleic acid complexes containing loops: the use of arenes instead of natural bases, especially at positions adjacent to the helix, where stacking interactions are expected to make the greatest contributions. Although the present compounds cannot form stabilizing cross-loop hydrogen bonds, they may have favorable hydrophobic or van der Waals interactions. Detailed structural and physical studies will be necessary to investigate the specific origins of this stability.

Experimental Procedure

Epimerization of α -anomers of **1a** and **2a** (bis-*p*-toluoyl esters) to β -anomers. **1a**: To a solution of the α -anomer of the bis-*p*-toluoyl ester of compound **1a** (synthesis described in ref. [9a]) (780 mg, 1.62 mmol) in toluene (50 mL) was added a catalytic amount of benzenesulfonic acid (ca. 10%), one drop of concentrated H₂SO₄, and two to four drops of H₂O. The reaction mixture was stirred vigorously at reflux for 4–6 h, and the mixture was then poured into 5% aqueous NaHCO₃ (50 mL) and extracted with EtOAc (3 × 50 mL). The combined organic layers were dried over anhydrous MgSO₄ and concentrated to dryness. Silica column chromatography (eluent 8/1 → 2/1 hexanes/EtOAc) of the crude mixture gave 430 mg of **1a** (β -epimer, 46% yield). ¹H NMR (CDCl₃): δ = 8.02 (2H, d, *J* = 8.0 Hz), 7.97 (2H, d, *J* = 8.0 Hz), 7.30–7.37 (1H, m, obscured), 7.31 (2H, d, *J* = 8.0 Hz), 7.25 (2H, d, *J* = 8.0 Hz), 6.75 (1H, dd, *J* = 8.0, 8.0 Hz), 5.64 (1H, brd, *J* = 5.8 Hz), 5.47 (1H, dd, *J* = 5.1, 10.8 Hz), 4.78 (1H, dd, *J* = 3.8, 11.8 Hz), 4.66 (1H, dd, *J* = 3.7, 11.8 Hz), 4.54 (1H, m), 2.64 (1H, m), 2.46 (3H, s), 2.43 (3H, s), 2.23 (1H, m), 2.17 (3H, s); ¹³C NMR (CDCl₃): δ = 13.8 (d), 22.0 (d), 40.1, 64.9, 74.9, 83.0, 103.0 (t), 120.1 (d), 124.5 (d), 127.3 (d), 128.6, 128.8 (d), 128.9 (d), 144.0 (d), 156.5 (d), 158.0 (d), 155.9 (d), 162.3 (d), 166.1 (d); HRMS (FAB, 3-nitrobenzyl alcohol matrix, *M* + H⁺): calcd for C₂₈H₂₆F₂O₅ 481.1827, found 481.1853.

2a (β -epimer, 54% yield). ¹H NMR (CDCl₃): δ = 8.02 (2H, d, *J* = 8.0 Hz), 7.97 (2H, d, *J* = 8.0 Hz), 7.30–7.37 (1H, m, obscured) 7.32 (2H, brd, *J* = 10.7 Hz), 7.31 (2H, br d, *J* = 11.2 Hz), 7.24 (1H, s), 6.95 (1H, s), 5.65 (1H, brd, *J* = 5.6 Hz), 5.42 (1H, dd, *J* = 5.0, 10.8 Hz), 4.76 (1H, d, *J* = 3.9, 11.8 Hz), 4.70 (1H, dd, *J* = 3.6, 11.8 Hz), 4.55 (1H, m), 2.56 (1H, m), 2.46 (3H, s), 2.43 (3H, s), 2.31 (3H, s), 2.23 (3H, s), 2.18 (1H, m), 2.13 (3H, s); ¹³C NMR (CDCl₃): δ = 18.2, 19.5, 22.1 (d), 41.0, 65.0, 82.5, 126.2, 127.0 (d), 128.6 (d), 128.8 (d), 132.1 (d), 135.5,

136.2, 141.4, 144.5, 165.5, 166.0; HRMS (FAB, 3-nitrobenzyl alcohol matrix, M^+): calcd for $C_{30}H_{32}O_5$ 472.2250, found 472.2234.

Conversion of nucleoside bis-*p*-toluoyl esters of **1a**, **2a** to phosphoramidite derivatives **1c**, **2c**: The hydrolysis of the bis-*p*-toluoyl esters, subsequent conversion to 5'-O-dimethoxytrityl derivatives, and 3'-O-phosphitylation in preparation for automated DNA synthesis were carried out by using the same methods previously described for the α -isomers [9b]. Analytical data for these compounds were in complete accord with their assigned structures [11b].

Oligonucleotide synthesis: DNA oligonucleotides were synthesized on an Applied Biosystems 392 instrument by using standard β -cyanoethylphosphoramidite chemistry [16]. Methods for deprotection, tritylation, and phosphitylation of nucleosides **3a–c** were as described previously [9b]. Oligomers were purified by preparative gel electrophoresis with 20% denaturing polyacrylamide (detection at 260 nm) Molar extinction coefficients for sequences containing the natural nucleosides were calculated by the nearest neighbor method [17]. The molar extinction coefficients measured for nucleosides **1**, **2**, and **3** were 1200, 851, and 6361, respectively; molar absorptivities for oligonucleotides containing these residues were derived as previously described [9b]. Oligodeoxynucleotides were obtained after purification as the sodium salts. Intact incorporation of residues **1–3** was confirmed by synthesis of short oligomers of sequence 5'-T-X-T (where X = **1**, **2**, or **3**). The ^1H NMR spectra (500 MHz) indicated the presence of the intact aromatic ring structures with the expected integration relative to thymine C-5 protons and methyl groups. Enzymatic degradation analysis could not be performed because the nonnatural residues were found to inhibit enzymatic cleavage.

Thermal denaturation studies: Buffers used for optical melting experiments were 100 mM NaCl, 10 mM Na-phosphate (for duplex hairpins) or 100 mM NaCl, 10 mM MgCl_2 , 10 mM Na-PIPES (1,4-piperazine(bisethanesulfonate)) (for triplexes). The buffer pH is that of a 1.4x stock solution at 25 °C containing the buffer and salts. After the solutions were prepared they were heated to 90 °C and allowed to cool slowly to room temperature prior to the melting experiments.

The melting studies were carried out in Teflon-stoppered quartz cells (1 cm path-length) under nitrogen atmosphere on a Varian Cary 1 UV/V is spectrophotometer equipped with a thermoprogrammer. Absorbance was monitored at 260 nm while temperature was increased at a rate of 0.5 °C min^{-1} ; a slower heating rate with this instrument does not affect the results. In all cases the complexes displayed sharp, apparently two-state melting transitions. Melting temperatures (T_m) were determined by computer fit of the first derivative of absorbance with respect to 1/T. Uncertainty in individual T_m values is estimated at ± 0.5 K. Free energy values were derived by computer-fitting the denaturation data with an algorithm employing linear sloping baselines by using the two-state approximation for melting[7]. Uncertainty in individual free energy measurements is estimated at ± 5 to 10%.

References

1. Groebe DR, Uhlenbeck OC. Nucleic Acids. Res 1988;16:11725–11735. [PubMed: 3211748]
2. Heus HA, Pardi A. Science 1991;253:191–194. [PubMed: 1712983]
3. Hirao I, Nishimura Y, Tagawa Y, Watanabe K, Miura K. Nucleic Acids. Res 1992;20:3891–3896. [PubMed: 1508675]
4. Wang S, Booher MA, Kool ET. Biochemistry 1994;33:4639–4644. [PubMed: 8161520]
5. SantaLucia J, Kierzek R, Turner DH. Science 1992;256:217–219. [PubMed: 1373521]
6. Cantor, CR.; Schimmel, PR. Biophysical Chemistry, Part III: The Behavior of Biological Macromolecules. W. H. Freeman: San Francisco; 1980. p. 1117–1133.
7. Petersheim M, Turner DH. Biochemistry 1983;22:256–263. [PubMed: 6824629]

8. Senior M, Jones RA, Breslauer KJ. *Biochemistry* 1988;27:3879–3885. [PubMed: 3408733]
9. a) Schweitzer BA, Kool ET. *J. Org. Chem* 1994;59:7238–7242. [PubMed: 20882116] b) *J. Am. Chem. Soc* 1995;117:1863–1872. [PubMed: 20882111]
10. Ren X-F, Schweitzer BA, Shiels CJ, Chaudhuri NC, Kool ET. unpublished.
11. a) Chaudhuri NC, Kool ET. *Tetrahedron Lett* 1995;36:1795–1798. *ibid.* **1995**, 36, 4910; b) Ren X-F, Chaudhuri NC, Kool ET. unpublished.
12. a) Coleman RS, Dong Y, Arthur JC. *Tetrahedron Lett* 1993;34:6867–6870. b) Loakes D, Brown DM. *Nucleic Acids Res* 1994;22:4039–4043. [PubMed: 7937128]
13. Senior MM, Jones RA, Breslauer KJ. *Proc. Natl. Acad. Sci. USA* 1988;85:6242–6246. [PubMed: 3413094]
14. a) Xodo LE, Manzini G, Quadrifoglio F, van der Marel GA, van Boom JH. *Nucleic Acids Res* 1991;19:5625–5631. [PubMed: 1945840] b) Kool ET. *J. Am. Chem. Soc* 1991;113:6265–6266. c) Prakash G, Kool ET. *J. Chem. Soc. Chem. Commun* 1991:1161–1162. *ibid.* **1994**, 646; d) Giovannangeli C, Montenay-Garestier T, Rougée M, Chassignol M, Thuong NT, Hélène C. *J. Am. Chem. Soc* 1991;113:7775–7776. e) Prakash G, Kool ET. *ibid* 1992;114:3523–3528. f) Salunkhe M, Wu T, Letsinger RL. *ibid* 1992;114:8768–8772. g) D'Souza DJ, Kool ET. *J. Biomol. Struct. Dyn* 1992;10:141–152. [PubMed: 1418737] h) Rumney S, Kool ET. *Angew. Chem* 1992;104:1686–1689. *Angew. Chem. Int. Ed. Engl.* 1992, 31, 1617–1619. i) Perkins TA, Goodman JL, Kool ET. *J. Chem. Soc. Chem. Commun* 1993:215–217. j) Giovannangeli C, Thuong NT, Hélène C. *Proc. Natl. Acad. Sci. USA* 1993;90:10013–10017. [PubMed: 8234249] k) Gryaznov SM, Lloyd DH. *Nucleic Acids Res* 1993;21:5909–5915. [PubMed: 8290352] l) Hudson RHE, Damha MJ. *Nucleic Acids Symp. Ser* 1993;29:97–99. [PubMed: 8247812] m) D'Souza DJ, Kool ET. *Bioorg. Med. Chem. Lett* 1994;4:965–970. n) Wang S, Kool ET. *Nucleic Acids Res* 1994;22:2326–2333. [PubMed: 7518582] o) Wang S, Kool ET. *J. Am. Chem. Soc* 1994;116:8857–8858. [PubMed: 20882109] p) Kandimal ER, Agrawal S. *Gene* 1994;149:115–121. [PubMed: 7525410] q) Noll DM, O'Rear JL, Cushman CD, Miller PS. *Nucleosides Nucleotides* 1994;13:997–1005. r) Trapane TL, Christopherson MS, Roby CD, Ts'o POP, Wang D. *J. Am. Chem. Soc* 1994;116:8412–8413. s) Bandaru R, Hashimoto H, Switzer C. *J. Org. Chem* 1995;60:786–788. t) Chaudhuri NC, Kool ET. *J. Am. Chem. Soc* 1995;117:10434–10442. [PubMed: 20882118]
15. a) Durand M, Chevrie K, Chassignol M, Thuong NT, Maurizot JC. *Nucleic Acids Res* 1990;18:6353–6359. [PubMed: 2243780] b) Durand M, Peloille S, Thuong NT, Maurizot JC. *Biochemistry* 1992;31:9197–9204. [PubMed: 1390706] c) Ma MY-X, Reid LS, Climie SC, Lin WC, Kuperman R, Sumner-Smith M, Barnett RW. *ibid* 1993;32:1751–1758. d) Doktycz MJ, Paner TM, Benight AS. *Biopolymers* 1993;33:1765–1777. [PubMed: 8268405] e) Thomson JB, Tuschl T, Eckstein F. *Nucleic Acids Res* 1993;21:5600–5603. [PubMed: 7506829] f) Benseler F, Fu D, Ludwig J, McLaughlin LW. *J. Am. Chem. Soc* 1993;115:8483–8484. g) Rumney S IV, Kool ET. *ibid* 1995;117:5635–5646.
16. a) Beaucage SL, Caruthers MH. *Tetrahedron Lett* 1981;22:1859–1862. b) Sinha ND, Biernat J, Köster H. *ibid* 1983;24:5843–5846.
17. Borer, PN. *Handbook of Biochemistry and Molecular Biology*. 3rd ed.. Fasman, GD., editor. Vol. Vol. 1. Cleveland, OH, USA: CRC Press; 1985. p. 589

Fig. 1. Structural formulas and space-filling models of nonpolar nucleoside isosteres **1** (F), **2** (B), and **3** (D). Also shown are the structures of the natural nucleosides thymidine and deoxyadenosine.

Table 1

Effects of natural and nonnatural nucleosides in tetranucleotide loops on the stability of double-stranded DNA hairpin structures at pH 7.0^[a].

Sequence	T_m [°C]	$-\Delta G_{60}^0$ [kcal mol ⁻¹]
thymidine and analogues		
5'-G C A A T T G C T T 3'-C G T T A A C G T T	67.7 (0.6)	1.3 (0.1)
5'-G C A A T T G C F F 3'-C G T T A A C G F F	78.5 (0.5)	3.6 (0.3)
5'-G C A A T T G C B B 3'-C G T T A A C G B B	78.4 (0.5)	3.1 (0.3)
deoxyadenosine and analogues		
5'-G C A A T T G C A A 3'-C G T T A A C G A A	62.9 (0.4)	0.4 (0.2)
5'-G C A A T T G C D D 3'-C G T T A A C G D D	67.2 (1.2)	1.1 (0.1)

^[a] Conditions: 100 mM NaCl, 10 mM phosphate (pH 7.0); DNA concentration: 5.0 μM. The sequences with the natural nucleotides are the same as those described in ref. [13]. Data shown are averages; error limits are given in parentheses.

Table 2

Effects of varying the first and last nucleotides in the loop on the stability of triplexes formed between 21-nucleotide pyrimidine DNAs and complementary 8-nucleotide purine DNA strands at pH 7.0^[a].

Complex with 5' loop	T_m [°C]	$-\Delta G_{25}^0$ [kcal mol ⁻¹]	Complex with 3' loop	T_m [°C]	$-\Delta G_{25}^0$ [kcal mol ⁻¹]
TCTTTTCCTT 3' T GAAAGAA 3' TCTTTTCCTT 5'	27.2	8.8	5'-TTCCTTTTCCTT 5'-AAGAAAGT 3'-TTCCTTTTCCTT	25.6	8.4
TCCTTTTCCTT 3' T GAAAGAA 3' TCCTTTTCCTT 5'	32.6	10.6	5'-TTCCTTTTCCTT 5'-AAGAAAGT 3'-TTCCTTTTCCTT	32.4	10.6

[a] Conditions: 100 mM NaCl, 10 mM MgCl₂, 10 mM Na-PIPES (pH 7.0); DNA concentration: 2.0 μM per strand. Errors: ±5 °C in T_m and ± 10% in free energy.