A novel N^3-functionalized thymidine linker for the stabilization of triple helical DNA

André H. Uddin, Miguel A. Roman, Janelle R. Anderson and Masad J. Damha*

Department of Chemistry, McGill University, Montreal, QC, Canada H3A 2K6

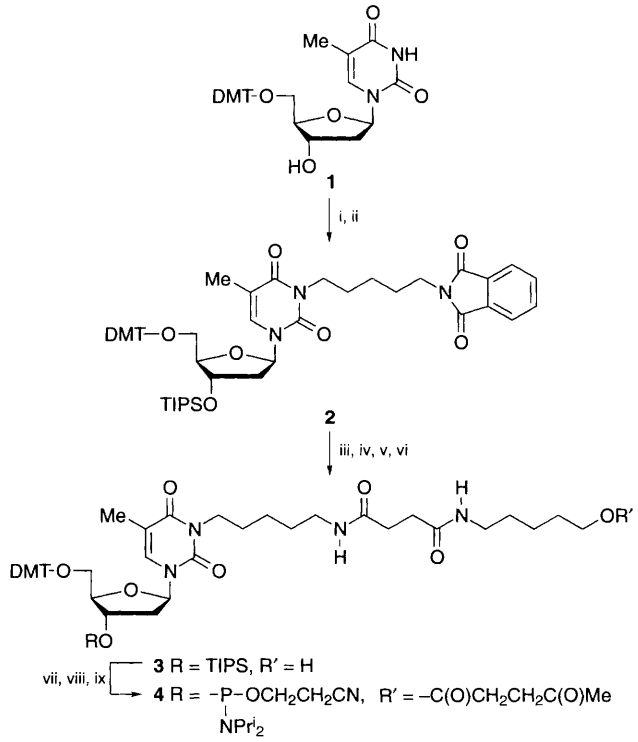
A novel method for regioselectively attaching a flexible linker to the N^3-position of thymidine is described; this nucleoside analogue served as a replacement for nucleotide loops in triplex forming DNA.

In recent years there has been a marked progression towards the study of nucleic acid triple helices both because of their roles in biological function and their potential use as therapeutics.[1] Due to the inability of DNA triplex structures to form under physiological conditions,[2] strategies are being developed to enhance the ability of oligonucleotides to form more stable triple helices: the use of minor group binders,[3] the covalent attachments of intercalators,[4] backbone and base modifications.[5] Our own studies relating to the synthesis of triplex forming branched 'V' oligonucleotide analogues[6] have been further motivated by our long-standing and continuing interest in the biological role of branched RNA.[7] The 'V' structures consisted of two parallel oligonucleotide chains joined to the sugar of a ribonucleoside *via* vicinal 2′-5′- and 3′-5′-phosphodiester bridges. Here we report the use of a N^3-functionalized thymidine as a nucleotide replacement for hairpin loops in triple helical DNA (Scheme 1). While the incorporation of natural nucleosides[8] or non-nucleoside bridges[9] as hairpin loops inside an oligonucleotide has been described, our new approach used a heterocyclic base branched to an aliphatic chain.† The main merit of this strategy is that this loop replacement provides a 5′-branching point for future substitutions or chain growth. Also of importance, this loop replacement affords a triple helix with comparable thermal stability to one with a natural loop.

The key branched synthon in our approach is compound **4**, which could be synthesized by the sequence depicted in Scheme 1. Regioselective N^3-alkylation of **1** with *N*-(5-bromopentyl) phthalimide and sodium hydride, followed by protection of the 3′-hydroxy with triisopropylsilyl chloride afforded **2** in 81% overall yield. Deprotection of the phthalamide moiety with methylamine afforded a primary amine, which was derivatized with a succinyl-aminopentanol linker (**2** → **3**; four steps: 60% overall). Levulination of **3** with levulinic (4-oxopentanoic) anhydride, followed by removal of the 3′-silyl group and 3′-*O*-phosphitylation led to desired product **4** in very good yields. In order to investigate the influence of the aliphatic linker on triplex stability, linkers of various lengths were also prepared by substituting *N*-(2-bromoethyl) and *N*-(4-bromobutyl)phthalimide for *N*-(5-bromopentyl)phthalimide in the synthetic scheme.[10]‡

The incorporation of the linker **4** into oligonucleotides is illustrated by the synthesis of oligomer **6** (Scheme 2). Thus, dT_{10} (1.0 μmol) was assembled on an automated DNA synthesizer using standard phosphoramidite protocols and controlled-pore glass (23 μmol dT g^{-1}) as the solid support. The branching synthon was introduced as a 0.15 mol dm^{-3} solution with an 80% efficiency (unoptimized) based upon the yield of the dimethoxytrityl cation released after coupling. The terminal 5′-hydroxy was then acetylated with capping solution to afford sequence **5**. The levulinyl protecting groups were removed manually with a solution of hydrazine hydrate using conditions which did not cause the cleavage of the oligomer from the support.[11] After thorough washing with dry MeCN, the column was re-installed on the synthesizer, and chain assembly of another dT_{10} strand was continued in the normal fashion. Quantitation of trityl cations released indicated that delevulination and subsequent coupling steps proceeded with *ca.* 98% yield. After deprotection with aq. ammonia (3 : 1, room temp., 24 h), preparative polyacrylamide gel electrophoresis (PAGE) on denaturing gels (24%/7 mol dm^{-3} urea) was employed for purification of the sequence. Desalting by size exclusion chromatography (Sephadex G-25) afforded 22 A_{260} units (35%) of oligomer **6**. Analysis of the sequence by analytical PAGE indicated that it was present as a single species, and its migration was comparable to that of unmodified sequences of similar chain lengths. The presence of the linker within the DNA sequence could be confirmed after hydrolysis with snake-venom phosphodiesterase and alkaline phosphatase, and analysis of the resulting mixture by HPLC.§

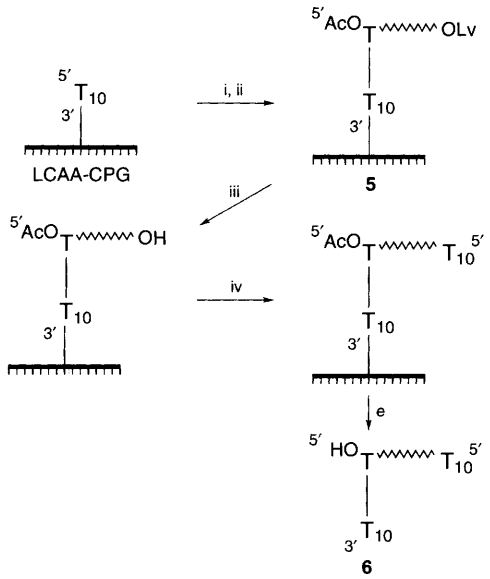
The binding affinities of **6** and the control sequence 5′-T_{10}-CCTC-T_{10}-3′ (**7**) for their single standard target dA_{10} were



Scheme 1 *Reagents and conditions*: i, *N*-(5-bromopentyl)phthalimide, sodium hydride, DMF, 24 h, room temp., 88%; ii, triisopropylsilyl chloride, imidazole, DMF, 24 h, room temp. (**2**, 92%); iii, 40% aqueous methylamine, benzene, ethanol, 34 h room temp. (78%); iv, succinic anhydride, 4-dimethylaminopyridine (DMAP), pyridine, 48 h, room temp. (quantitative); v, pentachlorophenol, DMAP, dicyclohexylcarbodiimide, DMF, 48 h, room temp. (quantitative); vi, 5 amino-pentanol, triethylamine, pyridine, 24 h, room temp. (**3**, 76%); vii, levulinic anhydride, pyridine, 24 h, room temp. (80%); viii, 1 mol dm^{-3} tetrabutylammonium fluoride in THF (quantitative); ix, 2-cyanoethyl *N,N*′-diisopropylphosphoramidous chloride, diisopropylethylamine, DMAP, THF (**4**, 72%)

measured by comparison of the melting temperatures of the complexes with 1 mol dm^{-3} NaCl, 10 mmol dm^{-3} Na_2HPO_4, pH 7.0 buffer, or 50 mmol dm^{-3} $MgCl_2$, 10 mmol dm^{-3} Tris pH 7.3. In both buffers, the melting curves at 260 nm showed a single transition from bound to unbound species. The same transition temperatures of both complexes (*ca.* 47 °C, in either Na^+ or Mg^{+2}) suggests that **6** and **7** bind to dA_{10} with the same affinity. By comparison, dT_{10} with its Watson–Crick complement dA_{10} forms a duplex which melts at 34 °C, demonstrating that **6** and **7** bind to dA_{10} much more strongly. Evidently, this is achieved by forming Hoogsteen and Watson–Crick hydrogen bonds, *i.e.* T·A/T base triplets, with the target dA_{10} strand (Fig. 1). Also supporting a triple helical structure are the observations that: (*a*) the melting transitions of the complexes of **6** and **7** could be followed at 284 nm, a wavelength at which pyr·pur/py triplexes composed entirely of T·A/T base triads display significant changes in absorbance, but at which duplex A/T pairs do not;[12] (*b*) A single mismatch resulting from the substitution of dT for dA in the target sequence, leads to a decrease in melting temperature of 13 degrees (T_m = 47 → 34 °C) whereas the corresponding decrease for the control duplex is more significant (T_m = 34 → 17 °C, or, ΔT_m = 17 °C); (*c*) circular dichroism spectra of **6**/dA_{10} and **7**/dA_{10}, is similar to that of the known triplex dT_{10}·dA_{10}/dT_{10} (data not shown).[12]

In conclusion, we have demonstrated an efficient and simple method for functionalizing thymidine with a linker, at the N^3-position. A hairpin loop inside of an oligonucleotide can be replaced with linker **4** without significant changes in its hybridization properties.¶ Such a linker may also enhance resistance of oligonucleotides to degradation in biological media, as well as its cellular uptake by increasing lipophilicity.[13] Of note, the present loop replacement **4** carries a 5′-hydroxy group that may be linked to a stabilizing group, or extended with a third ('guide') oligonucleotide to afford 'Y'-shaped structures.[6,7,14] The role of this 'guide' sequence and one of the pyrimidine sequences is to capture the target through specific Watson–Crick hydrogen bonding, while the role of the second pyrimidine strand is to 'fold' over and form a duplex/triplex chimera. Folding would be facilitated by the present flexible linker.



Scheme 2 *Reagents and conditions:* i, coupling with **4** (0.15 mol dm^{-3}) dissolved in THF–MeCN–CH_2Cl_2 (2 : 2 : 1), tetrazole, 15 min, room temp.; ii, Ac_2O–2,4,6-collidine–*N*-methylimidazole, 8 min, room temp.; iii, 0.5 mol dm^{-3} hydrazine hydrate in AcOH–pyridine, 5 min, room temp.; iv, 3′ → 5′ DNA synthesis; v, NH_3–EtOH (3 : 1), 24 h room temp.

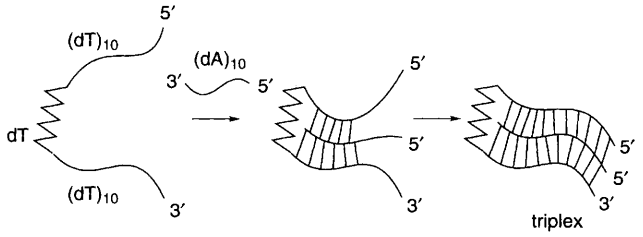


Fig. 1 Triplex formation of oligomer **6** with the complement dA_{10}

We thank Dr Robert H. E. Hudson for insightful discussions, and the NSERC (Canada) and FCAR (Quebec) for financial support of this work.

Footnotes

† 'Comb'-like nucleic acids have been prepared by a combination of a linear synthesis and subsequent orthogonal synthesis involving branching off cytosine bases.[14] However, there has been no reported use of these compounds for triplex helix formation.

‡ All compounds gave satisfactory MS (FAB) analyses, 2D-^{1}H, ^{13}C and where applicable, ^{31}P NMR, and complete protocols for their synthesis will be published elsewhere.[10]

§ These enzymes hydrolysed **6** to produce a mixture of dT and dT-linker (*i.e.* **3** lacking the dimethoxytrityl and TIPS groups) in the anticipated 20 : 1 ratio.

¶ As this work was being completed, Rumney IV and Kool reported the use of ethylene glycol oligomers as non-nucleoside replacements for nucleotide loops in triplex DNA.[15] The length of their optimum loop, a heptakis-(ethylene glycol) linker, is comparable to ours, *ca.* 30 Å. Also, they found that having one extra dT nucleotide as an extension adjacent to the flexible loop, as in this study, increases the stability of the complexes.

References

1 N. T. Thoung, and C. Hélène, *Angew. Chem., Int. Ed. Engl.*, 1993, **32**, 666.
2 M. Musso, and M. W. Van Dyke, *Nucleic Acids Res.*, 1995, **23**, 2320.
3 J. W. Lown, *ChemTracts-Organic Chemistry*, 1993, **6**, 205.
4 G. Duval-Valentin, T. Bizemont, M. Takasugi, J. Mergny, E. Bisagni, and C. Hélène, *J. Mol. Biol.*, 1995, **247**, 847; S. A. Cassidy, L. Strekoski, W. D. Wilson and K. R. Fox, *Biochemistry*, 1994, **33**, 15338.
5 R. L. Letsinger and T. Wu, *J. Am. Chem. Soc.*, 1995, **117**, 7323; J. Hunziker, S. Priestly, H. Brunar and P. B. Dervan, *J. Am. Chem. Soc.*, 1995, **117**, 2661.
6 R. H. E. Hudson, K. Ganeshan and M. J. Damha, in *Carbohydrate Modifications in Antisense Research*, ed. Y. S. Sanghvi, P. D. Cook, ACS Symposium Series, 1994, **580**, pp. 133–152; R. H. E. Hudson, A. H. Uddin and M. J. Damha, *J. Am. Chem. Soc.*, in the press.
7 R. H. E. Hudson and M. J. Damha, *J. Am. Chem. Soc.*, 1993, **115**, 2119; M. J. Damha, K. Ganeshan, R. H. E. Hudson and S. V. Zabarylo, *Nucleic Acids Res.*, 1992, **20**, 6565; M. J. Damha and K. K. Ogilvie, *Biochemistry*, 1988, **27**, 6403; M. J. Damha and K. K. Ogilvie, *J. Org. Chem.*, 1988, **53**, 3710.
8 E. R. Kandimalla and S. Agrawal, *Gene*, 1994, **149**, 115.
9 M. Salunkhe, T. Wu and R. L. Lestinger, *J. Am. Chem. Soc.*, 1992, **114**, 8768; S. Rumney and E. T. Kool, *Angew. Chem. Int. Ed. Engl.*, 1992, **31**, 1617; M. Y.-X. Ma, K. McCallum, S. C. Climie, R. Kuperman, W. C. Lin, M. Summer-Smith and R. W. Barnett, *Nucleic Acids Res.*, 1993, **21**, 2568; H. Gao, N. Chidambaram, B. C. Chen, D. E. Pelham, R. Patel, M. Yang, L. Zhou, A. Cook and J. S. Cohen, *Bioconjugate Chemistry*, 1994, **5**, 445.
10 M. Roman and M. J. Damha, manuscript in preparation.
11 K. K. Ogilvie, M. J. Nemer, G. H. Hakimelahi, Z. A. Proba and M. Lucas, *Tetrahedron Lett.*, 1982, **23**, 2615.
12 D. Pilch, C. Levenson and R. Shafer, *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 1942; R. D. Blake, J. Massoulie and J. R. Fresco, *J. Mol. Biol.*, 1967, **30**, 291.
13 M. Manoharan, L. K. Johnson, K. L. Tivel, R. H. Springer and P. D. Cook, *Bioorg. Med. Chem. Lett.*, 1993, **3**, 2765.
14 T. Horn and M. S. Urdea, *Nucleic Acids Res.*, 1989, **17**, 6959.
15 S. Rumney IV and E. T. Kool, *J. Am. Chem. Soc.*, 1995, **117**, 5635.

Received, 12th September 1995; Com. 5/06019B