Pergamon

SCIENCE DIRECT®

Tetrahedron Letters 44 (2003) 5065–5068

TETRAHEDRON LETTERS

Dual recognition of a C–G pyrimidine–purine inversion site: synthesis and binding properties of triplex forming oligonucleotides containing 2′-aminoethoxy-5-methyl-1*H*-pyrimidin-2-one ribonucleosides

Sabrina Buchini and Christian J. Leumann*

Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, CH-3012 Bern, Switzerland

Received 9 April 2003; revised 22 April 2003; accepted 5 May 2003

Abstract—The synthesis of a new ribonucleoside analogue, which combines two modifications, namely a 2′-aminoethoxy side-chain on the ribose and a 5-methyl-1*H*-pyrimidin-2-one (^{4H}T) unit as a base replacement, is presented. This building block was incorporated into triplex forming oligonucleotides and the binding properties to CG inversion sites in DNA duplex targets were studied. The data clearly show that the ^{4H}T base selectively recognizes the CG base-pair, while the aminoethoxy chain adds to the overall stability of the triple helix.

Specific recognition of double stranded DNA by triplex forming oligonucleotides (TFO) has been a field of intense investigations over recent years.[1,2] With such TFOs gene expression can be controlled on the level of transcription, which is of considerable interest for biotechnology and medicine.[3,4] Two triplex binding motifs are known in which the TFOs bind within the major groove of DNA either in a parallel or an antiparallel fashion to the purine strand of the target duplex. Sequence recognition is restricted to homopurine–homopyrimidine sequence tracts, and a general recognition code for any given DNA sequence is still elusive. Oligonucleotide modifications investigated to overcome this limitation included backbone modifications and the use of new bases designed to bind a whole Watson–Crick base pair, or to selectively recognize a pyrimidine base.[5] Since pyrimidine bases only present one free hydrogen bonding site within the major groove, their recognition is generally hampered by low affinity and low selectivity with respect to the bidentate recognition of purine bases.

We recently reported on the selective recognition of a CG base pair within the parallel DNA triple-helical binding motif by the base 5-methyl-1*H*-pyrimidin-2-one (^{4H}T, Fig. 1) in the context of deoxyribo-TFOs.[6] This base was designed to recognize cytosine via one conventional H-bond with the assistance of a non-conventional C–H–O hydrogen bond.[6,7] The general importance of such unconventional hydrogen bonds in (bio)molecular recognition has been demonstrated before.[8,9]

We found that CG recognition was highly selective, albeit at a lower level of affinity compared to a conventional pu–pu–py triplet. In order to increase affinity we decided to combine the ^{4H}T base with the 2′-aminoethoxy ribose modification which has been shown before to increase DNA target affinity of TFOs by additional salt bridge formation between the proto-

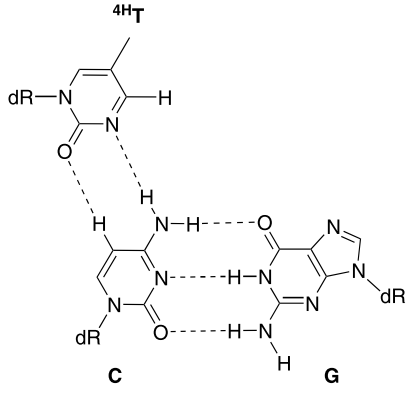
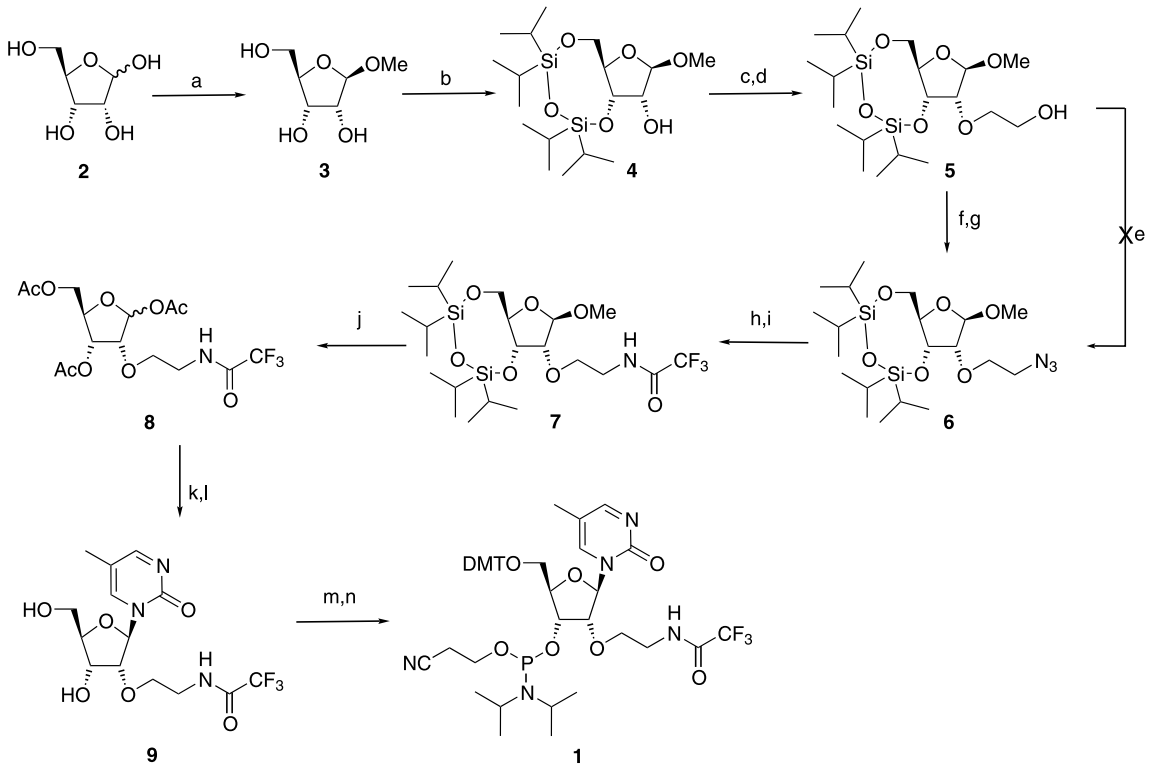


Figure 1. ^{4H}T-CG triplet in the parallel triplex binding motif.

* Corresponding author. Tel.: +41-31-631-4355; fax: +41-31-631-3422; e-mail: leumann@ioc.unibe.ch

0040-4039/03/$ - see front matter
doi:10.1016/S0040-4039(03)01135-3



Scheme 1. *Reagents and conditions*: (a) HCl 2% (w/w), MeOH, 3.5 h, rt, 63%; (b) TIPS-Cl (1.2 equiv.), pyridine, 1 h, rt, 83%; (c) $BrCH_2CO_2Me$ (5 equiv.), NaH (2.2 equiv.), DMF, 10 h, −10°C–rt, 77%; (d) $LiBH_4$ (4 equiv.), THF/MeOH (4/1), 1 h, 0°C, 94%; (e) Ph_3P (1.4 equiv.), LiN_3 (4 equiv.), CBr_4 (1.4 equiv.), DMF, 9 h, 0°C–rt; (f) TsCl (2 equiv.), NEt_3 (2.1 equiv.), DMAP (0.1 equiv.), CH_2Cl_2, overnight, rt, 93%; (g) LiN_3 (2 equiv.), DMF, 1.5 h, 90°C, 96%; (h) Lindlar catalyst (50%), H_2, THF, overnight, rt, quantitative; (i) Et_3N (1 equiv.), CF_3CO_2Et, 4 h, rt, 89%; (j) $Ac_2O/AcOH/H_2SO_4$, 14 h, rt, 91%; (k) (**11**, 1.05 equiv.), BSA (2.6 equiv.), $SnCl_4$ (1 equiv.), CH_3CN, 2 h, 0°C, 55%; (l) Na_2CO_3 anhyd. (10% w/w), MeOH, 30 min, rt, 90%; (m) DMT-Cl (2 equiv.), pyridine, 5 h, rt, 75%; (n) $i Pr_2NEt$ (3 equiv.), $[(iPr_2N)(NCCH_2CH_2O)P]Cl$ (1.5 equiv.), THF, 2 h, rt, 75%.

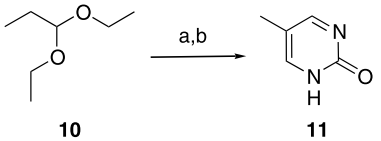
nated amino function of the side-chain and a phosphate group of the target duplex.[10]

Here we report on the synthesis of compound **1** (Scheme 1) and on the pairing properties of oligodeoxynucleotides containing this nucleoside analogue.

Methyl β-D-ribofuranoside **3** was synthesized starting from ribose, and was subsequently transformed into the TIPS protected methyl riboside **4**. Alkylation of **4** at the 2′-hydroxyl function with methyl bromoacetate, followed by reduction of the methyl ester, tosylation and substitution by azide yielded the azidoethyl derivative **6**. Direct azidation of **5** via Mitsunobu reaction proved to be unsuccessful. After reduction of the azide and protection of the resulting amine (→**7**) the coupling with the base **11** that was synthesised in two steps starting from propionaldehyde diethylacetal **10** (Scheme 2),[12,13] was attempted.

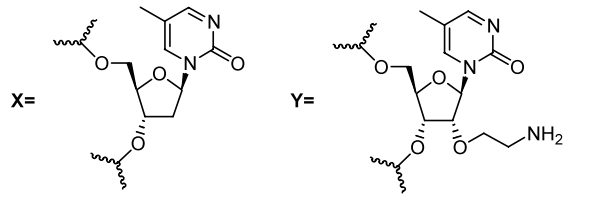
However, formation of the corresponding nucleoside could not be observed. The sugar was therefore converted into the triacetate **8** in order to facilitate the nucleosidation step.[11] Using **8** the nucleosidation could be achieved under Vorbrüggen conditions in 55% yield. Only the β-anomeric nucleoside **9** was isolated. Phosphoramidite **1** was then successfully obtained by standard procedures. Thus, intermediate **8**[14] may be a useful candidate for the introduction of other bases of interest.

Oligonucleotides were prepared on a DNA synthesizer using standard phosphoramidite chemistry (Table 1).[15] Sequence **12a** corresponds to the control containing thymine, sequence **12b** has previously been studied in our laboratories and differs from sequence **12c** only by the missing 2′-aminoethoxy chain. Modified oligonucleotide **12c** containing building block **1** was initially deprotected using standard conditions: NH_3 33%, 55°C, 16 h. However, under those conditions strand cleavage occurred at the site of base modification (HPLC and ESI-MS control), indicating instability of the base ^{4H}T



Scheme 2. *Reagents and conditions*: (a) (i) $POCl_3$ (2.2 equiv.), DMF (2.4 equiv.), 2 h, 70°C, (ii) H_2O, overnight, K_2CO_3, 36%; (b) (i) Na (2.3 equiv.), urea (1.2 equiv.), EtOH, 2 days, reflux, (ii) HCl 25% (pH 2), 1 day, reflux, 50%.

Table 1. Modified and natural TFOs as well as target duplexes investigated



12a	5′-TTTTT^{Me}CT**T**T^{Me}CT^{Me}CT^{Me}CT
12b	5′-TTTTT^{Me}CT**X**T^{Me}CT^{Me}CT^{Me}CT
12c	5′-TTTTT^{Me}CT**Y**T^{Me}CT^{Me}CT^{Me}CT
13	5′-GCTAAAAAGA**A**AGAGAGATCG
	3′-CTGTTTTTCT**T**TCTCTCTGAC
14	5′-GCTAAAAAGA**T**AGAGAGATCG
	3′-CTGTTTTTCT**A**TCTCTCTGAC
15	5′-GCTAAAAAGA**C**AGAGAGATCG
	3′-CTGTTTTTCT**G**TCTCTCTGAC
16	5′-GCTAAAAAGA**G**AGAGAGATCG
	3′-CTGTTTTTCT**C**TCTCTCTGAC
17	5′-TTTTT**Y**T^{Me}CT**Y**T^{Me}CT^{Me}CT
18	5′-TTTTT**Y**T**Y**T**Y**T^{Me}CT^{Me}CT
19	5′-TT**Y**TT**Y**T**Y**T**Y**T^{Me}CT**Y**T
20	5′-GCTAAAAA**C**AGA**C**AGAGATCG
	3′-CGATTTTT**G**TCT**G**TCTCTAGC
21	5′-GCTAAAAA**C**A**C**A**C**AGAGATCG
	3′-CGATTTTT**G**T**G**T**G**TCTCTAGC
22	5′-GCTAA**C**AA**C**A**C**A**C**AGA**C**ATCG
	3′-CGATT**G**TT**G**T**G**T**G**TCT**G**TAGC

towards extended exposure to conc. NH_3 at 55°C. Successful deprotection conditions were: NH_3 33%, rt, 2 h or $MeNH_2$ (40% aqueous solution), rt, 3 h.

In a first series of experiments, the binding efficiency of the TFOs **12a–c** to the target duplexes **13–16** was assessed by UV-melting curve analysis. Inspection of the T_m-data (Table 2) clearly shows that there is a preference for ^{4H}T-CG base-triple formation with both **12b** and **12c**. The loss of the carbonyl oxygen O(4) of thymine in ^{4H}T is responsible for the altered base-pair selectivity. Moreover the introduction of the 2′-aminoethoxy side-chain in **12c** leads to an increase in thermal stability of about 1.5°C/mod., relative to **12b**, without changing the selectivity. This indicates that at pH 6.5 the protonated aminoethoxy side-chain can interact specifically with a nearby phosphate group of the double helix much in the same way as reported before.[16]

Table 3. T_m (°C) data of third strand dissociation from UV-melting curves (260 nm)[a]

	T_m triplex (°C)
17/**20**	19.6
18/**21**	6.9
19/**22**	–[b]

[a] c(triplex) = 1.6 μM in 10 mM Na-cacodylate, 100 mM NaCl, 0.25 mM spermine, pH 6.5.
[b] No T_m detectable.

Triplex formation was also investigated at different pH values (Table 2). The data shows that there is no change in target selectivity for TFOs **12b** and **12c** within the pH range investigated (pH 6.5–7.4). At pH 7.0 the stability of the triplex **12c**/**15** is still higher than the stability of **12b**/**15**, while at pH 7.4 the two triplexes **12b**/**15** and **12c**/**15** are almost equally stable. This could either be due to incomplete protonation of the aminoethoxy side-chain, or reflect changes in the sugar conformation (preferred 2′-*endo* conformation for the deoxyribo- and 3′-*endo* conformation for the ribonucleoside) of the two ^{4H}T nucleotide units, that may be masked at low pH.

In order to explore the limits of this approach we investigated triplex formation of TFOs containing **1** with dsDNA targets containing multiple CG interruptions (Table 1, **17–22**). The TFOs **17** and **18**, containing two and three units of **1**, were found to still form triplexes with their respective targets **20** and **21** respectively at pH 6.5 (Table 3), however with significantly reduced T_m compared to TFOs targeted to single CG inversions.

The formation of a triplex with five **1**-CG triplets, as in **19–22** could not be observed, even by decreasing the pH to 5.5.

In order to further extend TFO binding also to dsDNA targets as e.g. **22**, we will prepare and investigate TFOs

Table 2. T_m (°C) data of third strand dissociation from UV-melting curves (260 nm)[a]

pH	**12a**				**12b**				**12c**			
	13 (A/T)	**14** (T/A)	**15** (C/G9	**16** (G/C)	**13** (A/T)	**14** (T/A)	**15** (C/G)	**16** (G/C)	**13** (A/T)	**14** (T/A)	**15** (C/G)	**16** (G/C)
6.5	61.0	29.3	38.0	32.3	27.4	29.7	38.9	29.5	29.8	27.6	40.5	31.1
7.0	45.0	21.1	29.7	23.3	19.0	17.0	30.1	18.2	21.1	20.0	31.4	20.4
7.4	38.4	13.8	22.6	17.1	13.8	13.3	23.9	12.5	14.8	13.2	24.2	12.6

[a] c(triplex) = 1.6 μM in 10 mM Na-cacodylate, 100 mM NaCl, 0.25 mM spermine, pH 6.5. UV-melting curves were recorded with a temperature gradient of 0.5°C/min. Data were obtained from the first derivative of the melting curve.

containing **1** that are uniformly 2′-aminoethoxy modified.

Acknowledgements

The authors wish to thank the Swiss National Science Foundation (Grant No. 20-63582.00) for financial support of this work.

References

1. Gowers, D. M.; Fox, K. R. *Nucleic Acids Res*. **1999**, *27*, 1569–1577.
2. Robles, J.; Grandas, A.; Pedroso, E.; Luque, F. J.; Eritja, R.; Orozco, M. *Curr. Org. Chem*. **2002**, *6*, 1333–1368.
3. Hèléne, C. *Anti-Cancer Drug Design* **1991**, *6*, 569–584.
4. Neidle, S. *Anti-Cancer Drug Design* **1997**, *12*, 433–442.
5. Fox, K. R. *Curr. Med. Chem*. **2000**, *7*, 17–37.
6. Prévot-Halter, I.; Leumann, C. J. *Bioorg. Med. Chem. Lett*. **1999**, *9*, 2657–2660.
7. Marfurt, J.; Leumann, C. J. *Angew. Chem., Int. Ed*. **1998**, *37*, 175–177.
8. Wahl, M. C.; Sundaralingam, M. *TIBS* **1997**, *22*, 97–102.
9. Desiraju, G. R. *Acc. Chem. Res*. **1996**, *29*, 441–449.
10. Cuenoud, B.; Casset, F.; Hüsken, D.; Natt, F.; Wolf, R. M.; Altmann, K. H.; Martin, P.; Moser, H. E. *Angew. Chem., Int. Ed*. **1998**, *37*, 1288–1291.
11. Sollogoub, M.; Dominguez, B.; Fox, K. R.; Brown, T. *Chem. Commun*. **2000**, *23*, 2315–2316.
12. Nair, V.; Vietti, D. E.; Cooper, C. S. *J. Am. Chem. Soc*. **1981**, *103*, 3030–3036.
13. Hildbrand, S.; Leumann, C.; Scheffold, R. *Helv. Chim. Acta* **1996**, *79*, 702–709.
14. Spectroscopical data of **8** (β anomer): ^{1}H NMR (300 MHz, $CDCl_3$): δ 2.10, 2.12, 2.14 (3s, 9H); 3.46–4.51 (m, 2H), 3.71–3.74 (m, 2H), 3.96–4.00 (dd, 1H); 4.11–4.18 (dd, 1H); 4.28–4.33 (dd, 1H); 4.48–4.52 (m, 1H), 5.18–5.21 (dd, 1H); 6.35 (d, 1H); 7.24 (s, 1H). ^{13}C NMR (75 MHz, $CDCl_3$): δ 21.3, 21.4, 21.7, 40.5, 64.1, 70.6, 71.1, 79.5, 81.8, 95.9, 114.7, 171.0, 171.2, 171.7.
15. *Oligonucleotides and Analogues: A Practical Approach*; Oxford University Press: Oxford, 1991.
16. A pK_a of 8.7 was determined for the 2′-aminoethoxy modified thymidine,[10] indicating that the 2′ side-chain is protonated under physiological conditions.