

Recognition of CG inversions in DNA triple helices by methylated 3*H*-pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-one nucleoside analogues†

Rohan T. Ranasinghe,^a David A. Rusling,^b Vicki E. C. Powers,^a Keith R. Fox^b and Tom Brown*^a

Received (in Cambridge, UK) 16th February 2005, Accepted 1st April 2005

First published as an Advance Article on the web 14th April 2005

DOI: 10.1039/b502325d

Substituted 3*H*-pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-one nucleoside analogues have been synthesised from 5-alkynyl-uridine derivatives, incorporated into triplex forming oligonucleotides (TFOs) and found to selectively bind CG inversions with enhanced affinity compared to T.

Triple helices¹ have been the subject of much interest in recent years due to their potential as antigene² and gene repair agents³ and as tools in other molecular biology applications.⁴ Mixed sequence recognition of duplex DNA is essential if triplex forming oligonucleotides (TFOs) are to be effective as tools in medicine and biotechnology. Pyrimidine-rich TFOs bind in a parallel orientation to the purine-rich strand of duplex DNA, forming hydrogen bonds to the purine bases. Although the natural bases T and C, and their analogues are effective at recognising AT and GC base pairs, Py.Pu pairs are more difficult targets, as the pyrimidine bases present fewer hydrogen bonding sites in the major groove.^{5,6} Despite this, TA can be selectively recognised by G, and synthetic nucleotides have been presented as alternatives.⁷ Although T binds to CG with intermediate affinity, this is the only base pair that cannot be recognised selectively by a natural base. As a result, many synthetic nucleotides have been designed to meet this molecular recognition challenge. Recently, 5-methyl-1*H*-pyrimidin-2-one (^{4H}T) has been presented as a selective partner for CG, forming one conventional N–H...N and one weak C–H...O hydrogen bond to C,⁸ and a 2'-aminoethoxy-modified ^{4H}T-nucleotide has been used in three base pair recognition of duplex DNA.⁹

We sought to design new nucleobase analogues for CG recognition that retain the hydrogen bonding residues from ^{4H}T, which are easily synthesised and allow incorporation of further functionality to improve recognition properties. To this end, we chose to synthesise two 3*H*-furo[2,3-*d*]pyrimidin-2-one nucleosides. The heterocycles are formed by base/CuI-catalysed cyclisation of 5-alkynylated uridine derivatives, and the versatility of the Sonogashira reaction¹⁰ allows introduction of a variety of substituents at the 6-position of the 3*H*-furo[2,3-*d*]pyrimidin-2-one. The appropriate phthalimide-protected alkynes **2a** and **2b** were prepared using Mitsunobu conditions.¹¹ These were reacted with 5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxy-5-iodouridine (**3**) using Pd-catalysed cross coupling to yield the protected nucleosides **4a** and **4b** (Scheme 1). Cyclisation to form the 3*H*-furo[2,3-*d*]pyrimidin-2-one heterocycles was accomplished using conditions

established by McGuigan *et al.*¹² Phosphitylation gave the phosphoramidites **6a** and **6b** for use in solid phase DNA synthesis.

The phosphoramidite monomers **6a**, **6b** and the commercially available **6c**† were incorporated into oligonucleotides using standard conditions. Deprotection was accomplished with 10% aqueous methylamine, which we have previously found to be an effective reagent for unmasking the phthalimide protecting group (unpublished data).

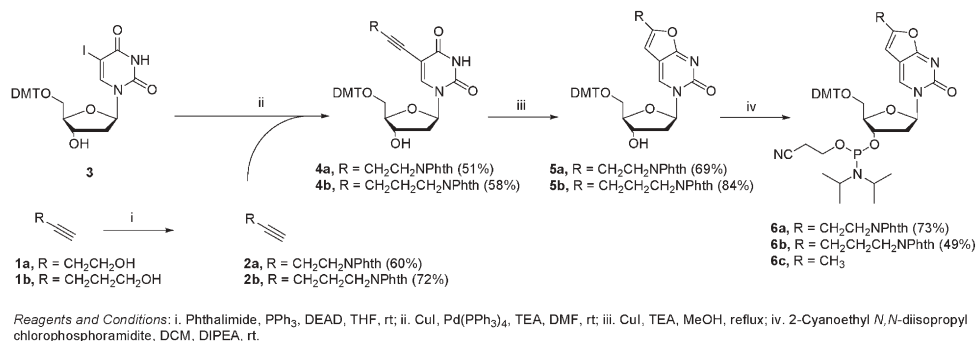
Despite some reports to the contrary,¹³ the 3*H*-furo[2,3-*d*]pyrimidin-2-one ring has been shown to be unstable to standard oligonucleotide deprotection conditions, with O7 being replaced with an NH from ammonia.¹⁴ The 3*H*-furo[2,3-*d*]pyrimidin-2-one oxygen has also been found to be unstable to other nucleophiles.¹⁵ We expected O7 to be replaced with an *N*-methyl group in the presence of methylamine, resulting in 6,7-dimethyl-3,7-dihydro-pyrrolo[2,3-*d*]pyrimidin-2-one, MP, and the 6-(aminoalkyl)-7-methyl-3,7-dihydro-pyrrolo[2,3-*d*]pyrimidin-2-ones, ^AEP and ^APP. The presence of the *N*-methyl group is essential to prevent the nucleobases acting as cytosine analogues, which would result in recognition of GC base pairs. The conversion to methylated 3*H*-pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-one nucleoside analogues was confirmed by mass spectrometry (Tables 1–3). Under these conditions, we observed no hydrolysis of the 3*H*-furo[2,3-*d*]pyrimidin-2-one heterocycles to generate uracil analogues, as has been recently reported.¹⁶

The three nucleobases studied (Fig. 1) allowed evaluation of several factors: MP could be used to assess the suitability of the 3*H*-pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-one core for CG recognition, and the protonated acyclic amino groups present in ^AEP and ^APP could provide extra stability by hydrogen bonding with N7 or O6 of G or by participating in electrostatic interactions with the anionic phosphate groups. Comparison of ^AEP with ^APP would demonstrate whether an ethylene or propylene spacer between the amino group and the nucleobase analogue is more effective.

Fluorescence melting allows rapid measurement of melting curves of nucleic acid structures.¹⁷ We used this technique to study the stability of triplexes containing a single CG inversion opposite the modified nucleotides MP, ^AEP, ^APP, and the natural nucleotide which has the greatest affinity for CG, T.¹⁸ The TFO containing the nucleotide for study is labelled with a quencher, Methyl Red, which causes release of a fluorescent signal from the fluorescein-labelled duplex upon triplex melting. We also used fluorescence melting to determine the selectivity of the modified nucleotides and T for each of the four base pairs.

The results are summarised in Table 1. All the substituted 3*H*-pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-one bases showed enhanced selectivity and affinity for the CG base pair compared with T. For all the 3*H*-pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-ones, the order of

† Electronic supplementary information (ESI) available: experimental protocols for DNA synthesis, fluorescence melting, DNase I footprinting, UV-melting and representative fluorescence melting curve data. See <http://www.rsc.org/suppdata/cc/b5/b502325d>
*tb2@soton.ac.uk



Scheme 1 Synthesis of 3*H*-furo[2,3-*d*]pyrimidin-2-one nucleoside phosphoramidites for incorporation into TFOs.

Table 1 T_m values (°C) determined for the fluorescence melting of intermolecular triplexes formed by the 18-mer TFO 5'-MR^a-TCTCTCTTXXCTCCTCC-3'^b, with the target duplexes 5'-F^c-AGAGAGAAAYAGGAGGAGG-3' / 5'-CCTCCTCTTCTCTCT-3'. T_m s were determined at pH 5.5 and 6.0 as indicated. The duplex concentration was 0.25 μ M, while the third strand was 3 μ M. Each T_m value is the average of three separate determinations, which typically differed by less than 0.5 °C

| X | pH | T_m (°C) | | | | |
|-----------------|-----|------------|-------------------|-------------------|-------------------|------|
| | | YZ = | AT | TA | GC | CG |
| T | 5.5 | | 54.5 | 39.2 | 43.7 | 44.5 |
| | 6.0 | | n.d. ^d | n.d. ^d | 27.7 | 28.3 |
| MP | 5.5 | | 42.7 | 38.7 | 41.4 | 46.4 |
| | 6.0 | | n.d. ^d | n.d. ^d | n.d. ^d | 31.5 |
| ^Δ EP | 5.5 | | 40.4 | 37.8 | 40.2 | 45.6 |
| | 6.0 | | n.d. ^d | n.d. ^d | n.d. ^d | 30.5 |
| ^Δ PP | 5.5 | | 42.1 | 37.6 | 41.4 | 47.1 |
| | 6.0 | | n.d. ^d | n.d. ^d | n.d. ^d | 32.1 |

^a MR = Methyl Red (fluorescence quencher). ^b MALDI-TOF MS of modified TFOs: X = MP found m/z 5720.5 [M + H]⁺ (expected 5720.9); X = ^ΔEP found m/z 5749.6 [M + H]⁺ (expected 5749.9); X = ^ΔPP found m/z 5764.1 [M + H]⁺ (expected 5764.0). ^c F = fluorescein (fluorescent reporter group). ^d n.d. – triplex formation was not observed above 28 °C.

stability of triplets at pH 5.5 and 6.0 is X.CG > X.AT > X.GC > X.TA. The order of affinity for CG of the nucleotides determined by fluorescence melting is ^ΔPP ≈ MP > ^ΔEP > T.

Quantitative DNase I footprinting with singly-modified 12-mer TFOs was used to confirm these results (Table 2). A 50% reduction

Table 2 C_{50} values (μ M) determined from quantitative DNase I footprinting experiments for the interaction of triplex forming oligonucleotides 5'-TCTCTTXXCTCCTCC-3'^a with target duplexes containing the sequence 5'-AGAGAAAYAAAG-3' / 5'-TCTTCTTCTCTCT-3' (YZ = AT, TA, GC or CG). The experiments were performed in 50 mM sodium acetate buffer (pH 5.0) containing 200 mM NaCl and 10 mM MgCl₂

| X | YZ | | | |
|-----------------|-----------|-------------------|------------|-----------|
| | AT | TA | GC | CG |
| T | 0.2 ± 0.1 | n.d. ^b | 2.1 ± 0.3 | 0.7 ± 0.4 |
| ^Δ EP | 2.0 ± 1.0 | n.d. ^b | 12.0 ± 5.1 | 0.8 ± 0.4 |
| ^Δ PP | 1.2 ± 0.5 | n.d. ^b | 4.8 ± 1.5 | 0.3 ± 0.1 |
| MP | 1.6 ± 0.2 | n.d. ^b | 6.0 ± 1.7 | 0.5 ± 0.1 |

^a MALDI-TOF MS of modified TFOs: X = MP found m/z 3581.8 [M + H]⁺ (expected 3581.4); X = ^ΔEP found m/z 3610.3 [M + H]⁺ (expected 3610.5); X = ^ΔPP found m/z 3624.5 [M + H]⁺ (expected 3624.5). ^b n.d. – triplex formation was not observed.

Table 3 T_m values (°C) determined for the UV-melting of singly substituted 15-mer TFOs 5'-TTTTT^mCTXT^mCT^mCT^mCT-3'^{a,b} with the 21 base pair duplex target 5'-GCTAAAAAGACA-GAGAGATCG-3' / 5'-CGATCTCTCTGTCTTTTATGC-3'. Melting curves were determined in 10 mM sodium cacodylate, 100 mM NaCl, 0.25 mM spermine buffer, at pH 7.0 with a heating rate of 0.5 °C/minute. The concentration of each strand was 0.53 μ M. Each T_m value is the average of three separate determinations and is reported to the nearest 0.5 °C

| X = | T_m (°C) | | | |
|-----|------------|------|-----------------|-----------------|
| | T | MP | ^Δ EP | ^Δ PP |
| | 27.0 | 28.0 | 26.5 | 28.0 |

^a ^mC = 5-methyl-2'-deoxycytidine. ^b MALDI-TOF MS of modified TFOs: X = MP found m/z 4535.7 [M + H]⁺ (expected 4535.1); X = ^ΔEP found m/z 4564.1 [M + H]⁺ (expected 4564.1); X = ^ΔPP found m/z 4577.8 [M + H]⁺ (expected 4578.2).

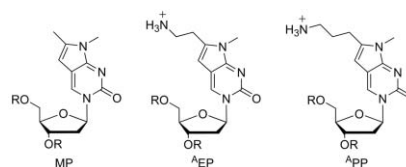


Fig. 1 3*H*-Pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-one nucleobase analogues studied.

in band intensity within the footprint was attained at third strand concentrations of 0.5 ± 0.1, 0.8 ± 0.4 and 0.3 ± 0.1 μ M for MP, ^ΔEP and ^ΔPP-containing TFOs respectively. ^ΔPP and MP therefore compare favourably with T (C_{50} = 0.7 ± 0.4 μ M). The order of triplet stability was again X.CG > X.AT > X.GC > X.TA. Footprints were only observed at the expected target site within the 110 bp fragment, confirming the selectivity of third strand binding (Fig. 2).

Finally, to allow comparison with ⁴H^T and to establish affinity of the three novel nucleobase analogues for duplexes containing a single CG inversion at pH 7.0, UV-melting was performed using TFO and duplex sequences analogous to those previously used by Buchini and Leumann to study ⁴H^T and its 2'-aminoethoxy-congener.¹⁹

Triplex formation at pH 7.0 was observed for all TFOs. Although the T_m observed for the T-containing TFO with the target duplex was not precisely that observed by Buchini and Leumann (27.0 °C in this study, compared to 29.7 °C), this is not surprising given the very large dependence on pH of the stability of

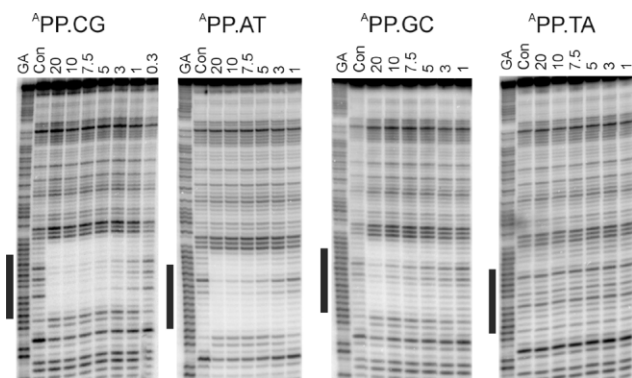


Fig. 2 DNase I footprinting experiments showing the interaction of oligonucleotide 5'-TCTCTT^APPTTTCT-3' with DNA fragments derived from *tyrT*(43–59) containing each base pair at the centre of the oligopurine tract, generating the central triplets ^APP.CG, ^APP.AT, ^APP.GC and ^APP.TA. The lanes labelled 'GA' are Maxam–Gilbert markers specific for purines while 'con' indicates DNase I cleavage in the absence of added oligonucleotide. The oligonucleotide concentration (μM) is shown at the top of each gel lane. The experiments were performed in 50 mM NaOAc buffer at pH 5.0 containing 10 mM MgCl₂ and the complexes were left to equilibrate overnight at 20 °C. The filled boxes show the position of the triplex target site.

triplexes containing ^mC⁺.GC triplets. We found that reducing the pH to 6.75 increased the T_m by > 5 °C (data not shown). Nevertheless, the difference between T_m s observed for TFOs containing modified nucleotides and that obtained for the T-containing TFO (ΔT_m) would allow evaluation of their relative stability. ΔT_m s previously reported for ⁴H^T and 2'-aminoethoxy-⁴H^T-modified TFOs at pH 7.0 were +0.4 and +1.7 °C. We determined ΔT_m s of +1, +1 and –0.5 °C for TFOs modified with MP, ^APP and ^AEP respectively. Triplets formed by MP and ^APP with a CG base pair are therefore very close in stability to those formed by ⁴H^T and its 2'-aminoethoxy- analogue.

In summary, we have synthesised three substituted 3*H*-pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-ones and studied their properties in triplexes by fluorescence melting, quantitative DNase I footprinting and UV-melting, and found them to bind CG base pairs between pH 5.0 and 7.0. We have therefore established a new heterocyclic nucleobase core for binding CG inversions. Recognition is presumably achieved by the N–H···N and C–H···O hydrogen bonds present in ⁴H^T (Fig. 3), but this requires confirmation from a thorough structural study by NMR or X-ray crystallography.

When all data are taken into account, the order of affinity for CG of the substituted 3*H*-pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-ones is ^APP ≈ MP > ^AEP ≈ T. It can therefore be concluded that although the 3*H*-pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-one heterocyclic core is effective in recognising CG base pairs, our initial attempts to enhance the thermodynamic stability of triplexes by the inclusion of pendant protonated amino groups have not been successful, although the propylene spacer present in ^APP was found to be more effective than the ethylene spacer of ^AEP. However, it has been reported that protonated primary amino groups enhance binding kinetics of TFOs,²⁰ so ^APP may be important in this regard. This requires further study.

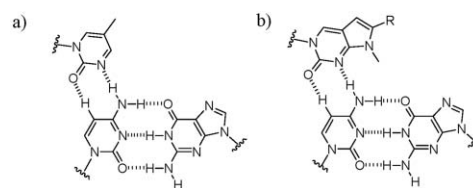


Fig. 3 Proposed triplets for previously studied CG recognition bases a) ⁴H^T, b) substituted 3*H*-pyrrolo[2,3-*d*]pyrimidin-2(7*H*)-ones.

Most importantly, we have shown that the use of different alkyne derivatives in the Sonogashira reaction provides easy access to nucleobases with a variety of substituents at the 6-position, which will allow us to synthesise a range of second generation compounds that may have enhanced binding affinities.

This research was supported by the European Union (EU) and Cancer Research UK. We are grateful to the EPSRC (DAR) and the BBSRC (VECP) for PhD studentships.

Rohan T. Ranasinghe,^a David A. Rusling,^b Vicki E. C. Powers,^a Keith R. Fox^b and Tom Brown^{*a}

^a*School of Chemistry, University of Southampton, Highfield, Southampton, UK SO17 1BJ. E-mail: tb2@soton.ac.uk; Fax: 44 2380 592974; Tel: 44 2380 592991*

^b*School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton, UK SO16 7PX. E-mail: krfl@soton.ac.uk; Fax: 44 2380 4459; Tel: 44 2380 594374*

Notes and references

‡ Furano-dT Phosphoramidite (commercial name of compound **6c**) was a gift from the Glen Research Corporation, Sterling, Virginia, USA.

- G. Felsenfeld, D. R. Davies and A. Rich, *J. Am. Chem. Soc.*, 1957, **79**, 2023.
- C. Hélène, *Anti-Cancer Drug Des.*, 1991, **6**, 569.
- M. M. Seidman and P. M. Glazer, *J. Clin. Invest.*, 2003, **112**, 487.
- H. E. Moser and P. B. Dervan, *Science*, 1987, **238**, 645.
- K. R. Fox, *Curr. Med. Chem.*, 2000, **7**, 17.
- D. M. Gowers and K. R. Fox, *Nucleic Acids Res.*, 1999, **27**, 1569.
- D. Guianvarc'h, R. Benhida, J. L. Fourrey, R. Maurisse and J. S. Sun, *Chem. Commun.*, 2001, 1814.
- I. Prevot-Halter and C. J. Leumann, *Bioorg. Med. Chem. Lett.*, 1999, **9**, 2657.
- S. Buchini and C. J. Leumann, *Angew. Chem., Int. Ed.*, 2004, **43**, 3925.
- K. Sonogashira, *J. Organomet. Chem.*, 2002, **653**, 46.
- O. Mitsunobu, *Synthesis*, 1981, 1.
- C. McGuigan, H. Barucki, S. Blewett, A. Carangio, J. T. Erichsen, G. Andrei, R. Snoeck, E. De Clercq and J. Balzarini, *J. Med. Chem.*, 2000, **43**, 4993.
- C. J. Yu, H. Yowanto, Y. J. Wan, T. J. Meade, Y. Chong, M. Strong, L. H. Donilon, J. F. Kayyem, M. Gozin and G. F. Blackburn, *J. Am. Chem. Soc.*, 2000, **122**, 6767.
- J. S. Woo, R. B. Meyer and H. B. Gamper, *Nucleic Acids Res.*, 1996, **24**, 2470.
- D. Loakes, D. M. Brown, S. A. Salisbury, M. G. McDougall, C. Neagu, S. Nampalli and S. Kumar, *Helv. Chim. Acta*, 2003, **86**, 1193.
- G. S. Jiao and K. Burgess, *Chem. Commun.*, 2004, 1304.
- R. A. J. Darby, M. Sollogoub, C. McKeen, L. Brown, A. Risitano, N. Brown, C. Barton, T. Brown and K. R. Fox, *Nucleic Acids Res.*, 2002, **30**, e39.
- K. Yoon, C. A. Hobbs, J. Koch, M. Sardaro, R. Kutny and A. L. Weis, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**, 3840.
- S. Buchini and C. J. Leumann, *Tetrahedron Lett.*, 2003, **44**, 5065.
- N. Puri, A. Majumdar, B. Cuenoud, P. S. Miller and M. M. Seidman, *Biochemistry*, 2004, **43**, 1343.