COMMUNICATION

ChemComm
www.rsc.org/chemcomm

An analogue of adenine that forms an "A:T" base pair of comparable stability to G:C

James Booth,[a] W. Jon Cummins[b] and Tom Brown*[a]

[a] *School of Chemistry, University of Southampton, Highfield, Southampton, UK SO17 1BJ. E-mail: tb2@soton.ac.uk*

[b] *Jon Cummins, GE Healthcare, Grove Center, White Lion Road, Amersham, Bucks, UK HP7 9LL*

Received (in Cambridge, UK) 17th June 2004, Accepted 19th July 2004
First published as an Advance Article on the web 20th August 2004

The heterocyclic base 7-aminopropargyl-7-deaza-2,6-diaminopurine (D) has been incorporated into oligodeoxynucleotides. D:T has similar thermodynamic stability to G:C and is a stable analogue of A:T.

Parallel high-throughput oligonucleotide probe technologies are used extensively in mutation detection and SNP (single nucleotide polymorphism) analysis. In all such systems it is important that fully-matched probe–target duplexes anneal at the same temperature, and that sequences containing a single mismatch are strongly disfavoured. This is not fully attainable with canonical Watson–Crick base pairs, as AT-rich sequences must be longer than their GC-rich counterparts in order to have the same melting temperature (T_m). Consequently the relative instability of a single base mismatch is sometimes too small to be of practical value. If this undesirable sequence dependence of duplex stability could be eliminated it would be much easier to design and develop highly efficient massively parallel DNA analysis systems. In this context it has recently been shown that analogues of A and T can be used to increase the efficiency of hybridisation probes.[1] In this paper we describe the synthesis and thermodynamic properties of a new 7-aminopropargyl-7-deazapurine (**D**) that has been designed to increase the stability of AT-rich DNA duplexes.

5-Aminopropargyl-modified deoxyuridine derivatives have previously been used as thymidine analogues to stabilise DNA duplexes and triplexes.[2–5] The origin of the stabilising effect is twofold; the aminopropargyl group, which is protonated at physiological pH, reduces charge repulsion between the phosphate groups of the opposing DNA chains[6] and the alkynyl moiety improves base stacking.[7] There is also a need for an adenine analogue for modified oligonucleotide probes to stabilise A:T as well as T:A base pairs. It is well established that the 2,6-diaminopurine:T pairing, with three inter-base hydrogen bonds, is intermediate in stability between G:C and A:T base pairs[8] and Seela *et al.* have recently demonstrated that 7-propynyl-8-aza-7-deaza-2,6-diaminopurines significantly stabilise DNA duplexes.[9]

It was therefore reasoned that 7-aminopropargyl-7-deaza-2,6-diaminopurine (**D**) might give rise to an A:T base pair analogue with similar stability to G:C (Fig. 1). In order to further investigate the effects of the 7-aminopropargyl and 2-amino groups on base pair stability, work was also undertaken to investigate 7-aminopropargyl-7-deaza-2′-deoxyinosine (**I**) when base-paired to cytosine (Fig. 1). It has recently been shown that the 7-aminopropargyl-7-deaza derivatives of 2′-deoxyadenosine (**A**) and 2′-deoxyguanosine (**G**) yield more stable DNA duplexes than their unmodified counterparts[9–11] so these phosphoramidites were synthesised for comparison.

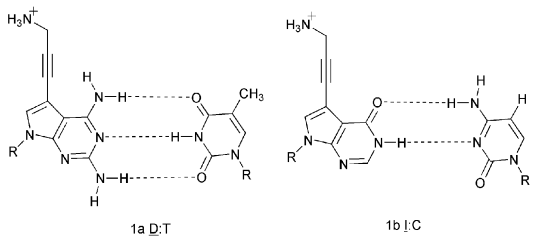
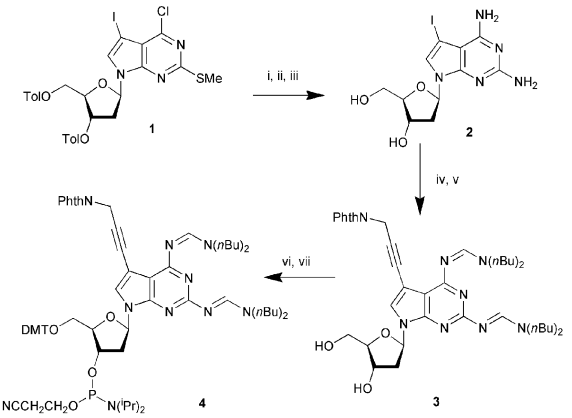


Fig. 1 Pseudosymmetric Watson–Crick base pair analogues.

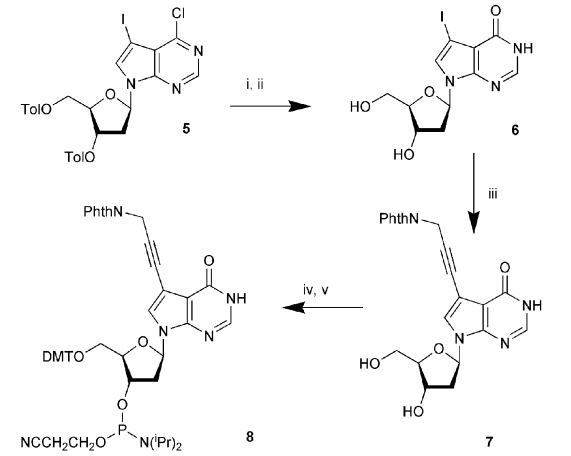
The phosphoramidite (**4**) of the 7-deazadiaminopurine analogue **D** was prepared[12] according to Scheme 1.

The deoxyinosine analogue (**I**) was synthesised using similar methodology to give phosphoramidite **8** (Scheme 2).

All four modified nucleosides (**A**, **D**, **I** and **G**) were incorporated into 12-mer oligonucleotides by standard solid-phase phosphoramidite chemistry† and thermodynamic parameters for duplex formation were determined by UV-melting studies in 0.1 M sodium chloride at pH 7.0‡ The low salt concentration was chosen so that the data would be relevant to PCR and physiological conditions. Entries 1 to 5 in Table 1 show that both the 7-aminopropargyl



Scheme 1 i, MCPBA, CH_2Cl_2, 0 °C to room temp., 2 h, ii, NH_3, 1,4-dioxane, 140 °C, 9 h, iii, NH_3, MeOH, 150 °C, 24 h, 43% over 3 steps; iv, $[(Ph_3P)_4Pd^{(0)}]$, CuI, TEA, DMF, 12 h, 94%; v, $(MeO)_2NCHN(nBu)_2$, DMF, 12 h, 37%; vi, pyridine, DMTCl, 16 h, 42%; vii, 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite, DIPEA, CH_2Cl_2, 1 h, 47%.



Scheme 2 i, *syn*-2-pyridinealdoxime, 1,1,3,3-tetramethylguanidine, DMF, 1,4-dioxane, 24 h, 95%; ii, NH_3, MeOH, 140 °C, 2 h, 76%; iii, $[(Ph_3P)_4Pd^{(0)}]$, CuI, TEA, DMF, 12 h, 94%; iv, DMTCl, pyridine, 16 h, 41%; v, 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite, DIPEA, DCM, 1 h, 45%.

DOI: 10.1039/b409155h

Table 1 Thermodynamic data for oligonucleotide duplexes containing the 7-aminopropargyl-7-deaza derivatives of 2′-deoxyadenosine (**A**), 2,6-diaminopurine (**D**), deoxyguanosine (**G**), and deoxyinosine (**I**)

	Duplex 5′–3′/3′–5′	T_m/°C[a]	ΔT_m/°C[b]	ΔH°/kJ mol^{-1}	ΔS°/J mol^{-1} K^{-1}	ΔG°_{310}/kJ mol^{-1}
1	GCTATCTATCTG CGATAGATAGAC	37.60	—	−354.3	−1019.5	−38.08
2	GCTATCT**A**TCTG CGATAGATAGAC	40.60	+2.96	−346.5	−984.64	−41.09
3	GCT**A**TCT**A**TCTG CGATAGATAGAC	44.30	+3.35	−370.4	−1045.7	−46.11
4	GCTATCT**D**TCTG CGATAGATAGAC	42.25	+4.65	−363.6	−1032.7	−43.10
5	GCT**D**TCT**D**TCTG CGATAGATAGAC	47.72	+5.06	−378.4	−1059.7	−49.70
6	GCTATCTATCTG CGATA**G**ATAGAC	41.22	+3.62	−367.7	−1048.8	−42.39
7	GCTATCTATCTG CGATA**G**ATA**G**AC	43.51	+2.95	−351.7	−991.2	−44.31
8	GCTATCTATCTG C**G**ATA**G**ATA**G**AC	46.30	+2.90	−371.9	−1043.9	−48.11
9	GCTATCTATCTG CGATA**I**ATAGAC	36.03	−1.57	−343.0	−988.99	−36.20
10	GCTATCTATCTG CGATA**I**ATA**I**AC	33.91	−1.85	−332.2	−961.02	−34.11
11	GCTATCTATCTG C**I**ATA**I**ATA**I**AC	28.54	−3.02	−306.5	−890.38	−30.35

[a] T_m for 2 μM total single strand concentration. [b] Change in T_m compared to unmodified oligonucleotide/number of modifications.

Table 2 Thermodynamic data for modified and unmodified oligonucleotide duplexes in which the modified bases in Table 1 have been substituted with guanine or adenine

	Unmodified sequence Positions of substitution lower case	T_m/°C [a]	Modified oligonucleotide T_m/°C	ΔT_m/°C [b]
1	GCTATCTgTCTG CGATAGAcAGAC	42.97	**A:**T = 40.56	**A:**T = −2.41
			D:T = 42.25	**D:**T = −0.72
2	GCTgTCTgTCTG CGAcAGAcAGAC	47.77	**A:**T = 44.30	**A:**T = −3.47
			D:T = 47.72	**D:**T = −0.05
3	CGATAaATAGAC GCTATtTATCTG	35.01	**I:**C = 36.03	+1.02
4	CGATAaATAaAC GCTATtTATtTG	31.59	**I:**C = 33.91	+2.32
5	CaATAaATAaAC GtTATtTATtTG	26.97	**I:**C = 28.54	+1.57

[a] T_m for 2 μM total single strand concentration. [b] Difference in T_m between modified and unmodified oligonucleotide.

group and the 2-amino group of the 7-deazadiaminopurine increased the stability of the "A:T" base pair. Consequently **D** is a strongly stabilising analogue of A:T whereas **A** is less strongly stabilising. The very favourable enthalpy for the **D**:T base pair (entry 5) is an indication of good base stacking. A clear trend was also found for G:C base pair analogues (entries 6–11). **G**:C was strongly stabilising whereas **I**:C was significantly less stable than G:C.

Table 2 entries 1 and 2 compare **D**:T and **A**:T with **G**:C, and give an indication of the potential of these analogues for harmonising the melting temperatures of G:C and A:T-rich oligonucleotide probes. Clearly **A** is not useful in this context, but **D** is a close thermodynamic analogue of G and the **I**:C base pair is slightly more stable than A:T. The judicious use of **I** and **D** in place of G and A at specific loci in oligonucleotides would permit the synthesis of probes with harmonised melting temperatures. Substitution with **I** or **G** would also prevent aggregation, a common problem in G-rich oligonucleotides that occurs through mini-quadruplex formation, that requires the presence of the purine N(7)-atom.[13] Table 1 entries 6–8 show that **G**:C is more stable than G:C and Table 2 entries 1 and 2 indicate that **D**:T is similar in stability to G:C. It can therefore be concluded that **G**:C is more stable than **D**:T and that **G**:C is the most stable base pair in this study. Clearly neither **D** nor **I** are close structural analogues of A or G respectively in terms of hydrogen bonding capacity (Fig. 1). In the minor groove **D** is G:C-like and **I** is A:T-like and the lack of a purine N(7)-atom dictates that neither analogue accurately mimics its Watson–Crick counterpart in the major groove, which is partly blocked by the aminopropargyl moiety. This will influence protein and drug binding characteristics, and oligonucleotides containing these analogues may be valuable probes for exploring such interactions.

Notes and references

† All oligonucleotides were synthesised on an ABI 394 DNA synthesiser using a standard 0.2 μmole phosphoramidite cycle. Deprotection was by 10% methylamine in water (2 mL) containing phenol (2.5 mg) for 36 hours at room temp. Purification was by reversed phase HPLC (column: ABI C8 (octyl) 8 mm × 250 mm, pore size = 300 Å) (Buffer A: 0.1 M ammonium acetate, pH 7.0. Buffer B: 0.1 M ammonium acetate with 20% acetonitrile, pH 7.0). Gradient: time in min (% Buffer B); 0 (0); 3 (0); 5 (20); 21 (100); 25 (100); 27 (0); 30 (0). UV detection at 290 nm. Mass data were obtained for all oligonucleotides on a MALDI-TOF ThermoBioAnalysis Dynamo mass spectrometer in positive ion mode using a 3-hydroxypicolinic acid–picolinic acid (4 : 1) matrix with 50% aqueous acetonitrile solvent). All were accurate to 1 mass unit.

‡ Thermodynamic parameters for duplex formation were determined by UV melting on a Cary 400 UV/visible spectrometer in 100 mM NaCl, 1 mM EDTA and 10 mM Na-phosphate buffer at pH 7.0. Duplexes were melted and annealed and the mean T_m was determined in triplicate for 10 different oligonucleotide concentrations from 0.4 to 16 μM. Each melting curve was fitted to a non-self-complementary two-state model using Meltwin 3.5 software.[14,15] Thermodynamic parameters for duplex formation were determined by averaging data from individual melt curves. The errors were less than 2.1%, 2.4% and 0.8% for ΔH°, ΔS° and ΔG° respectively. Concentration dependent van't Hoff analysis gave the same trends for T_m and ΔG°.

1 M. J. Heller, *Ann. Rev. Biomed. Eng.*, 2002, **4**, 129–153.
2 D. Graham, J. A. Parkinson and T. J. Brown, *Chem. Soc., Perkin Trans. 1*, 1998, 1131–1138.
3 R. A. J. Darby, M. Sollogoub, C. McKeen, L. J. Brown, A. Risitano, N. Brown, C. Barton, T. Brown and K. R. Fox, *Nucleic Acids Res.*, 2002, **30**, e39.
4 D. M. Gowers, J. Bijapur, T. Brown and K. R. Fox, *Biochemistry*, 1999, **38**, 13747–13758.
5 J. Bijapur, M. D. Keppler, S. Bergqvist, T. Brown and K. R. Fox, *Nucleic Acids Res.*, 1999, **27**, 1802–1809.
6 L. E. Heystek, H. Q. Zhou, P. Dande and B. Gold, *J. Am. Chem. Soc.*, 1998, **120**, 12165–12166.
7 R. W. Wagner, M. D. Matteucci, J. G. Lewis, A. J. Gutierrez, C. Moulds and B. C. Froehler, *Science*, 1993, **260**, 1510–1513.
8 B. L. Gaffney, L. A. Marky and R. A. Jones, *Tetrahedron*, 1984, **40**, 3–13.
9 H. Rosemeyer, N. Ramzaeva, E. M. Becker, E. Feiling and F. Seela, *Bioconjugate Chem.*, 2002, **13**, 1274–1285.
10 N. Ramzaeva, C. Mittelbach and F. Seela, *Helv. Chim. Acta*, 1999, **82**, 12–18.
11 F. Seela, N. Ramzaeva, P. Leonard, Y. Chen, H. Debelak, E. Feiling, R. Kroschel, M. Zulauf, T. Wenzel, T. Frohlich and M. Kostrzewa, *Nucleosides Nucleotides Nucleic Acids*, 2001, **20**, 1421–1424.
12 C. A. Buhr, R. W. Wagner, D. Grant and B. C. Froehler, *Nucleic Acids Res.*, 1996, **24**, 2974–2980.
13 D. Sen and W. Gilbert, *Nature*, 1988, **334**, 364–366.
14 C. E. Longfellow, R. Kierzek and D. H. Turner, *Biochemistry*, 1990, **29**, 278–285.
15 J. A. McDowell and D. H. Turner, *Biochemistry*, 1996, **35**, 14077–14089.