COMMUNICATION

ChemComm
www.rsc.org/chemcomm

Oligonucleosides with a nucleobase-including backbone; synthesis and self-association of novel dinucleotide analogues

Andrew J. Matthews, Punit K. Bhardwaj and Andrea Vasella*
Laboratory for Organic Chemistry, ETH-Hönggerberg, HCI, CH-8093 Zürich, Switzerland.
E-mail: vasella@org.chem.ethz.ch; Fax: +41 1632 1136; Tel: +41 1632 5130

Received (in Cambridge, UK) 23rd January 2003, Accepted 26th February 2003
First published as an Advance Article on the web 12th March 2003

The synthesis and self-association of protected oxymethylene-bridged UA analogues are described.

We are studying oligonucleotide analogues with a nucleobase-including backbone (Fig. 1, **A**), to determine whether the structural differentiation between nucleobase and backbone in DNA, RNA, and their analogues (Fig. 1, **B**), is a prerequisite for the formation of stable homo- and/or heteroduplexes.[1]†

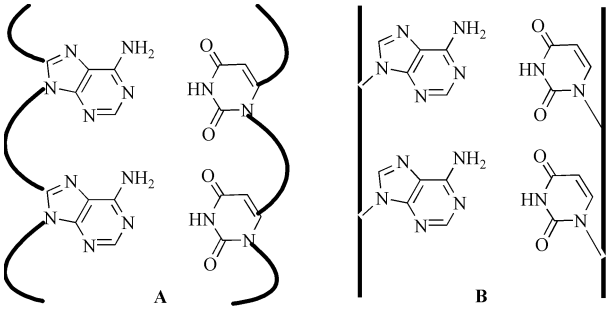


Fig. 1 Schematic representation of (**A**) oligonucleotide analogues with a nucleobase including backbone, and (**B**) oligonucleotides and analogues with a structural differentiation of nucleobase and backbone.

Tetrameric analogues of type **A**, derived from ethynediyl-linked adenosine and uridine, showed no evidence for homo-pairing, and a similar uridine hexamer did not hetero-pair with a complementary RNA strand.[2] Modeling suggested that an *anti* conformation of these analogues is a prerequisite for pairing, whilst NMR analysis of an adenosine dimer showed that a *syn* conformation is preferred.[3] Modeling studies also suggested that oxymethylene-bridged oligomers (Fig. 2), should pair in the *syn* conformation (Watson–Crick type hydrogen bonding), so far only known to occur in *Z*-DNA.[4] We have hence prepared the corresponding self-complementary UA dimer.

2′,3′-*O*-Isopropylideneuridine (**1**) was protected as the TIPS ether **2** (Scheme 1). Deprotonation with LDA,[5] followed by formylation with DMF and reduction with $NaBH_4$[6] gave C(6)-hydroxymethylated **3**. The C(8)-hydroxymethylated adenosine **7** was prepared *via* a similar route from protected adenosine **5**. Treatment of **3** with mesyl chloride gave the chloromethylated **4** (64% from **1**). 4,4′-Dimethoxytritylation of **7** yielded **8**, which was then desilylated to **9** (59% from **5**). The ether **10** (Scheme 2) was prepared by alkylation of alcohol **9** with the chloride **4**; *N*-debenzoylation gave the amine **11**. Detritylation or desilylation of **11** led to the monoalcohols **12** and **13**, respectively, which were further deprotected to give diol **14**. Acid hydrolysis[2] of **14** yielded the oxymethylene dimer **15**.‡

According to the chemical shift for H–C(2′),[7] the protected uridine **2** prefers an *anti* conformation (δH–C(2′) = 4.70 ppm), whilst the C(6)-substituted derivatives **3**, **4** and **10**–**13** prefer a *syn* conformation (δ = 5.19–5.40 ppm). Likewise, the protected adenosine **6** prefers the *anti* conformation (δH–C(2′) = 5.32 ppm), whilst the C(8)-substituted derivatives **7–13** prefer a *syn* conformation (δ = 5.70–6.03 ppm).

The ^{1}H-NMR spectra of **11–13** in $CDCl_3$ are characterised by a concentration dependent downfield shift of the uridine H–N(3), evidencing an intermolecular hydrogen bond. Association constants K_a (Table 1),[8] and the thermodynamic parameters $\Delta H°$ and $\Delta S°$ in $CDCl_3$ were calculated for **11–13** from the δ(H–N(3)) concentration and temperature dependence (Fig. 3).§ The δ(H–N(3)) of diol **14** in $CDCl_3$ was almost concentration independent (12.96–12.65 ppm from 33–1 mM), whilst the fully deprotected dimer **15** was insufficiently soluble to evidence association. Other solvent systems are under investigation. The data for **11** and **12** highlight the contribution of the lipophilic dimethoxytrityl group, but even the K_a value obtained for the

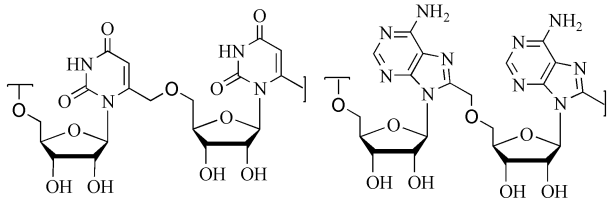
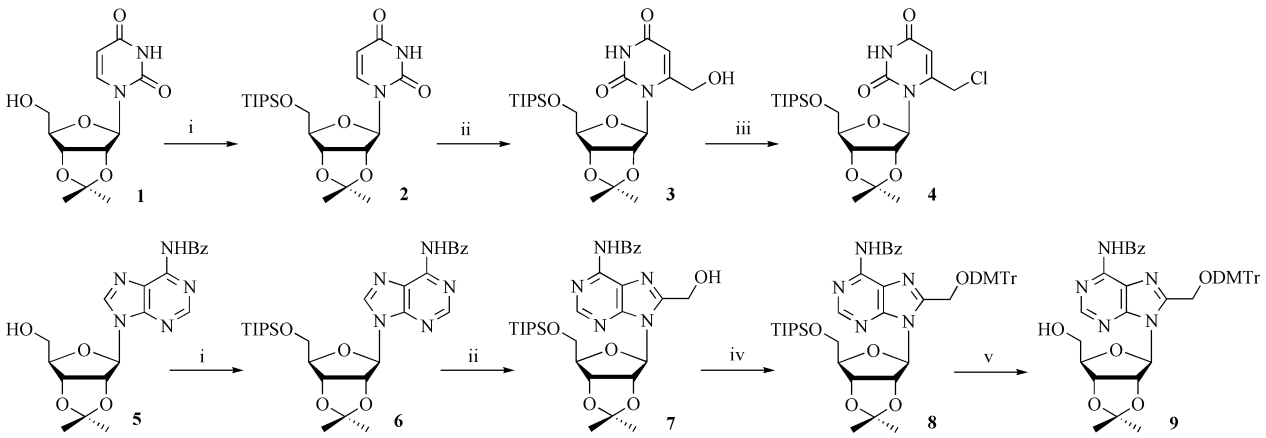


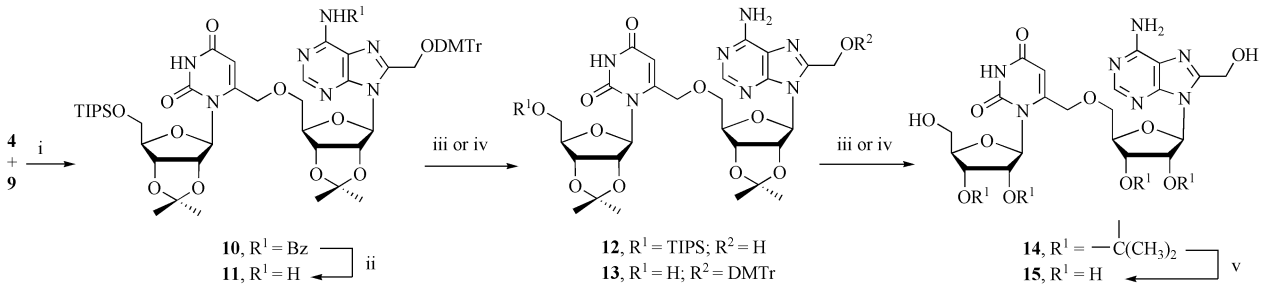
Fig. 2 Oxymethylene bridged oligonucleotide analogues.



Scheme 1 *Reagents and conditions*: i. TIPSCl, imidazole, CH_2Cl_2, rt, 98% (**2**), 96% (**6**); ii. a) LDA (5 eq.), THF, −78 °C; b) DMF, −78 °C to rt; c) AcOH, EtOH, $NaBH_4$, rt, 70% (**3**), 74% (**7**); iii. MsCl, C_5H_5N, 0 °C to rt, 94%; iv. DMTrCl, $EtN(iPr)_2$, CH_2Cl_2, 50 °C, 94%; v. TBAF, THF, rt, 89%.

DOI: 10.1039/b300932g

This journal is © The Royal Society of Chemistry 2003



Scheme 2 *Reagents and conditions*: i. NaH, DMF:THF (2:1), 0 °C, 60%; ii. NH_4OH, MeOH, 95%; iii. HCO_2H, $MeNO_2$, rt, 84% (**12**), 82% (**14**); iv. TBAF, THF, rt, 91% (**13**), 85% (**14**); v. HCO_2H, H_2O (8:2), rt, 75%.

Table 1 Association constants and thermodynamic parameters for the dimers **11–13** in $CDCl_3$

	K_a (M^{-1})[a]	$-\Delta H°$ (kcal mol^{-1})	$-\Delta S°$ (e.u.)[b]
11	966	15.8	40.4
12	277	21.8	63.7
13	3222	24.4	64.8

[a] Determined at 22 °C, uncertainty in K_a estimated at 15%. [b] e.u. = entropy units (1 e.u. = 1 cal per (mol.K)).

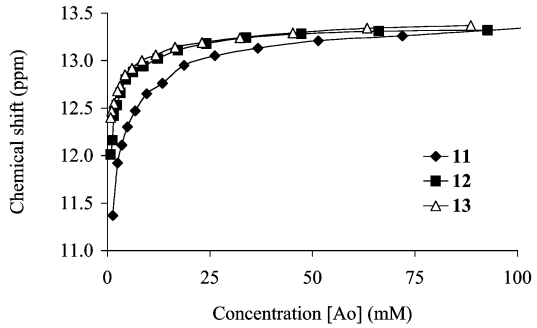


Fig. 3 Concentration dependence of δ(H–N(3)) for dimers **11**, **12** and **13** in $CDCl_3$ at 295 K.

detritylated **12** compares favourably with that determined for 3′,5′-di-*O*-acetyl-2′-deoxyuridine with a 2′-deoxyadenosine derivative (70 M^{-1}).[9] The high K_a value for **13**, and the inability to dissociate **14** appreciably in $CDCl_3$ correlate with a downfield shift of H–O(5′) as compared to **1** ($\Delta\delta \approx 1.0$ ppm), and are rationalised by the formation of a C(5′)O–H···O=C(2) intramolecular hydrogen bond.

Watson–Crick type base pairing is suggested by a cross-peak between the hydrogen bonded imino H–N(3) and the adenine H–C(2) in a 2D-NOESY experiment on associated dimer **11**.[9]

These results support the contention that a structural differentiation of nucleobases and backbone is not required for pairing. We are now investigating the details of base pairing, stacking, and hydrogen bonding.

We thank The Royal Society (A. J. M.), the Swiss National Science Foundation and *F*. Hoffmann-La-Roche AG, Basel for generous support.

Notes and references

† For the sake of simplicity, we have designated these analogues as 'oligonucleotide analogues with a nucleobase-including backbone', while, strictly speaking, these systems do not possess a 'backbone'.

‡ All new compounds showed satisfactory NMR, IR, and MS data. **11**: I = adenosyl unit; II = uridyl unit. ^{1}H-NMR (500 MHz, $CDCl_3$): 13.09 (br s, H–N(3/II)); 8.37 (s, H–C(2/I)); 8.37–7.49 (m, 2 arom. H); 7.47–7.38 (m, 4 arom. H); 7.32–7.21 (m, 3 arom. H); 6.90 (br s, 2 H–N(6/I)); 6.86–6.83 (m, 4 arom. H); 6.22 (d, J = 1.3, H–C(1′/I)); 5.88 (dd, J = 1.3, 6.3, H–C(2′/I)); 5.75 (d, J = 1.0, H–C(1′/II)); 5.36 (s, H–C(5/II)); 5.34 (dd, J = 6.3, 3.8, H–C(3′/I)); 5.27 (dd, J = 1.0, 6.3, H–C(2′/II)); 4.87 (dd, J = 6.3, 4.5, H–C(3′/II)); 4.55, 4.41 (AB, J = 11.8, 2 H–C(10/I)); 4.44, 4.03 (AB, J = 13.3, 2 H–C(7/II)); 4.31 (ddd, J = 3.8, 5.3, 4.9, H–C(4′/I)); 4.15 (ddd, J = 4.4, 5.4, 7.1, H–C(4′/II)); 3.85 (dd, J = 10.5, 5.4, H–C(5′a/II)); 3.83 (dd, J = 10.5, 7.1, H–C(5′b/II)); 3.79 (s, MeO); 3.65 (dd, J = 10.5, 5.3, H–C(5′a/I)); 3.63 (dd, J = 10.5, 4.9, H–C(5′b/I)); 1.55, 1.55, 1.42, 1.41 (4s, Me_2C); 1.01–0.96 (m, $(Me_2CH)_3$–Si). ^{13}C-NMR (75 MHz, $CDCl_3$): 164.07 (s, C(2/II)); 158.50 (s, C(8/I)); 155.61 (s, C(6/I)); 151.91 (d, C(2/I)); 150.74 (s, C(6/II)); 150.24 (s,C(4/II)); 150.10 (s, C(4/I)); 148.81 (s, arom. C); 144.00 (s, arom. C); 135.08 (s, arom. C); 130.02 (d, arom. C); 128.10 (d, arom. CH); 127.81 (d, arom. CH); 126.90 (d, arom. CH); 118.45 (s, C(5/I));113.59 (s, *C*$(Me)_2$/I); 113.26 (s, *C*(Me)2/II); 113.17 (d, arom. CH);103.71 (d, C(5/II)); 91.39 (d, C(1′/II)); 90.00 (d, C(1′/I)); 89.65 (d, C(4′/II)); 87.50 (s, *C*Ar_3);86.74 (d, C(4′/I)); 84.37 (d, C(2′/II)); 83.57 (d, C(2′/I)); 82.26 (d, C(3′/II)); 81.67 (d, C(3′/I)); 69.80 (t, C(5′/I)); 68.08 (t, C(7/II)); 64.50 (t, C(5′/II)); 59.27 (t, C(10/I); 55.23 (q, 2 × MeO); 27.36, 27.36, 25.76, 25.76 (4q, *Me*$_2$C); 17.95 (q, *Me*$_2$CH)$_3$Si); 11.99 (d, Me_2*C*H)$_3$Si). HR-MALDI-MS: 303 (100%, $[DMTr]^+$); 1114.492 (23%, $[M + Na]^+$; calc. 1114.4936). IR ($CHCl_3$): 3488w, 3185w, 2993m, 2943m, 2866m, 2840w, 1712s, 1636m, 1608m, 1509s, 1446m, 1383m, 1157m, 1068s, 1036m, 882m, 831m.

§ NMR was performed at 295 K on a Varian Gemini300 spectrometer (300 MHz) in $CDCl_3$ passed through aluminium oxide immediately prior to use. Experiments started at the highest concentration, with stepwise replacement of 0.2 ml of the 0.7 ml solution with 0.2 ml pure $CDCl_3$. The data were analysed graphically and by nonlinear least-squares fitting.[8]

Thermodynamic parameters were determined by van't Hoff analysis. The uridyl δH–N(3) was monitored between 50 and −30 °C at a fixed concentration (between 20–80% of saturation). Linear fits of data collected below 0 °C were poor.

1 S. Eppacher, N. Solladié, B. Bernet and A. Vasella, *Helv. Chim. Acta*, 2000, **83**, 1311.
2 H. Gunji and A. Vasella, *Helv. Chim. Acta*, 2000, **83**, 1331; H. Gunji and A. Vasella, *Helv. Chim. Acta*, 2000, **83**, 2975; H. Gunji and A. Vasella, *Helv. Chim. Acta*, 2000, **83**, 3229.
3 P. K. Bhardwaj and A. Vasella, *Helv. Chim. Acta*, 2002, **85**, 699.
4 W. Saenger, *Principles of Nucleic Acid Structure*, Springer Verlag, Berlin, 1984, p. 21; T. Wada, N. Minamimoto, Y. Inaki and Y. Inoue, *J. Am. Chem. Soc.*, 2000, **122**, 6900; S. M. Gryaznov, D. H. Lloyd, J. K. Chen, R. G. Schultz, L. A. DeDionisio, L. Ratmeyer and W. D. Wilson, *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 5798; O. Almarsson and T. C. Bruice, *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 5542.
5 H. Hayakawa, K. Haraguchi, H. Tanaka and T. Miyasaka, *Chem. Pharm. Bull.*, 1987, **35**, 72.
6 W. Czechtizky and A. Vasella, *Helv. Chim. Acta*, 2001, **84**, 594; W. Czechtizky and A. Vasella, *Helv. Chim. Acta*, 2001, **84**, 1000; W. Czechtizky, X. Daura, A. Vasella and W. van Gunsteren, *Helv. Chim. Acta*, 2001, **84**, 2132; W. Czechtizky, Dissertation ETH No. 14239, ETH-Zürich, 2001.
7 D. Gani and A. W. Johnson, *J. Chem. Soc., Perkin Trans. 1*, 1982, 1197; S. Nieble, M. R. Sanderson, A. Subbiah, J. B. Chattopadhyaya and C. B. Reese, *Biochim. Biophys. Acta*, 1979, **565**, 379; L. Dudycz, R. Stolarski, R. Pless and D. Shugar, *Z. Naturforsch.*, 1979, **34**, 359.
8 K. A. Conners, *Binding Constants*, John Wiley & Sons, New York, 1987; W. C. Luo and J. S. Chen, *Z. Phys. Chem.*, 2001, **215**, 447; *Associate v. 1.6*, B. R. Peterson, Ph.D Thesis, University of California at Los Angeles, 1994.
9 A. Dunger, H-H. Limbach and K. Weisz, *J. Am. Chem. Soc.*, 2000, **122**, 10109; G. M. Nagel and S. Hanlon, *Biochemistry*, 1972, **11**, 823.