The chemical nature of the 2′-substituent in the pentose-sugar dictates the pseudoaromatic character of the nucleobase (pK_a) in DNA/RNA†

Subhrangsu Chatterjee, Wimal Pathmasiri, Oleksandr Plashkevych, Dmytro Honcharenko, Oommen P. Varghese, Mohitosh Maiti and Jyoti Chattopadhyaya*

Received 31st January 2006, Accepted 3rd March 2006
First published as an Advance Article on the web 30th March 2006
DOI: 10.1039/b601460g

We here show that the pK_a (error limit: 0.01 to 0.03 pK_a unit) of a nucleobase in a nucleotide can be modulated by the chemical nature of the 2′-substituent at the sugar moiety. This has been evidenced by the measurement of nucleobase pK_a in 47 different model nucleoside 3′,5′-bis- and 3′-mono-ethylphosphates. The fact that the electronic character of each of the 2′-substituents (Fig. 1) alters the chemical shift of the H2′ sugar proton, and also alters the pK_a of the nucleobase in the nucleotides has been evidenced by a correlation plot of pK_a of *N3* of pyrimidine (T/C/U) or pK_a of *N7* of 9-guaninyl with the corresponding δH2′ chemical shifts at the neutral pH, which shows linear correlation with high Pearson′s correlation coefficients ($R = 0.85$–0.97). That this modulation of the pK_a of the nucleobase by a 2′-substituent is a through-bond as well as through-space effect has been proven by *ab initio* determined pK_a estimation. Interestingly, experimental pK_as of nucleobases from NMR titration and the calculated pK_as (by *ab initio* calculations utilizing closed shell HF 6-31G** basis set) are linearly correlated with $R = 0.98$. It has also been observed that the difference of ground and protonated/de-protonated HOMO orbital energies (ΔHOMO, a.u.) for the nucleobases (A/G/C/T/U) are well correlated with their pK_as in different 2′-substituted 3′,5′-bis-ethylphosphate analogs suggesting that only the orbital energy of HOMO can be successfully used to predict the modulation of the chemical reactivity of the nucleobase by the 2′-substituent. It has also been demonstrated that pK_a values of nucleobases in 3′,5′-bis-ethylphosphates (Table 1) are well correlated with the change in dipole moment for the respective nucleobases after protonation or de-protonation. This work thus unambiguously shows that alteration of the thermodynamic stability (T_m) of the donor–acceptor complexes [ref. 20], as found with various 2′-modified duplexes in the antisense, siRNA or in triplexes by many workers in the field, is a result of alteration of the pseudoaromatic character of the nucleobases engineered by alteration of the chemical nature of the 2′-substitution.

Introduction

Modification at C2′ of the sugar moiety is widely used to make oligonucleotides thermodynamically and nucleolytically stable as well as to recruit RNase H[1,2] (through the well-known gapmer or mixmer strategies) in the antisense[3–6] approach for the down-regulation of gene expression.[7] These 2′-modified nucleosides have been also found to be extremely useful in the design of unique binding properties of aptamers[8] to a specific ligand by *in vitro* evolution (SELEX)[9] as well as to understand the general mechanism of RNA catalysis[10–13] (ribozyme),[8] DNAzyme[14] and small interfering RNAs (RNAi).[15–17] Clearly, any change of the pseudoaromatic character[18,19] of the nucleobase has a profound consequence in terms of its hydrogen-bonding ability in the formation of donor–acceptor complex in general. Every pK_a unit increase for T or U or G aglycon in the antisense strand should result in the destabilization of A–T or A–U or C–G basepairing contribution in ΔG°_{25} by 5.8 kJ mol^{-1}, whereas every pK_a unit increase in C or A aglycon in the antisense strand would result in 5.8 kJ mol^{-1} stabilization of G–C or U–A basepairing in the sense–antisense duplexes. The exact basepairing contribution in ΔG°_{25} in a duplex will, however, be modulated by the sequence-context specific modulation of the pseudoaromaticity,[19] owing to the nearest-neighbor stacking interactions. We have earlier shown that the aglycons in 2′-deoxyribonucleotides[20] are in general more basic than those in the ribonucleotide counterpart because of the electron-withdrawing effect of the 2′-OH group[20] in the latter: Thus, 9-adeninyl in dA is more basic by 0.11 pK_a unit than that in rA ($\Delta\Delta G^\circ_{pKa} = 0.6$ kJ mol^{-1}), 9-guaninyl in dG is more basic by 0.3 pK_a unit than that in rG ($\Delta\Delta G^\circ_{pKa} = 1.7$ kJ mol^{-1}), 1-uracilyl in dU is more basic by 0.33 pK_a unit than that in rU ($\Delta\Delta G^\circ_{pKa} = 1.8$ kJ mol^{-1}), and 1-cytosinyl in dC is more basic by 0.1 pK_a unit than that in rC ($\Delta\Delta G^\circ_{pKa} = 0.8$ kJ mol^{-1}). Similarly, the net stabilization[20] of r(C–G) basepair over d(C–G) basepair and r(A–T) basepair over d(A–T) basepair are 0.20 and 0.76 kJ mol^{-1} respectively. Thus, any modulation observed in the relative acidity/basicity of the nucleobase by the chemical nature of the 2′-substituent (Table 1) should affect the donor–acceptor properties, and accordingly should influence the strength for H-bonding[20,21] as well as the stacking with the nearest-neighbor in a potential duplex or triplex. We here present a general picture

Department of Bioorganic Chemistry, Box 581, Biomedical Center, Uppsala University, SE-75123 Uppsala, Sweden. E-mail: jyoti@boc.uu.se; Fax: +46-18554495

† Electronic supplementary information (ESI) available: Fig. S1–S5 and Tables S1–S3. See DOI: 10.1039/b601460g

Table 1 The experimental pK_a of nucleoside phosphates (see Figure 1 for compound numbers) obtained by NMR titration

pK_a [a]			pK_a [a]			pK_a [a]			pK_a [a]	
N1/N3	N7		N1/N3	N7		N1/N3	N7		N3	1′,2′-Azetidine *or* 2′-NH_2
(1a) 3.69 (0.01) H8/H2	—	**(2a)**	3.77 (0.02) H8/H2	—	**(3a)**	3.83 (0.01) H8/2		**(4a)**	3.68 (0.01) H8/H2	—	—	—
(1b) 9.27 (0.02) H8	2.21 (0.02)	**(2b)**	9.73 (0.02) H8	2.21 (0.01)	**(3b)**	9.59 (0.02) H8	2.46 (0.01)	**(4b)**	9.74 (0.03) H8	1.82 (0.02)	—	—
(1c) 4.24 (0.01) H6/H5	—	**(2c)**	4.26 (0.01) H6[b]	—	**(3c)**	4.35 (0.01) H6/H5		**(4c)**	3.54 (0.01) H6/H5	—	**(5c)** 3.24 (0.02) H6/5	6.08 (0.02) H6
(1d) 9.78 (0.01) H6/CH_3	—	**(2d)**	9.94 (0.02) H6/CH_3	—	**(3d)**	10.12 (0.01) H6/CH_3		**(4d)**	9.51 (0.02) H6/CH_3	—	**(5d)** 9.60 (0.02) H6/CH_3	5.88 (0.02) H6
(1e) 9.26 (0.02) H6/H5	—		—		**(3e)**	9.59 (0.02), H6/H5		**(4e)**	8.93 (0.02) H6/H5		**(5e)** 9.11 (0.03) H6/H5	5.90 (0.03) H6
(1f) 3.11 (0.02) H8/H2	—	**(2f)**	3.46 (0.01) H8/H2	—	**(3f)**	3.35 (0.01) H8/H2		**(4f)**	3.59 (0.02) H8/H2	—	—	—
(1g) 9.27 (0.01) H8	1.90 (0.01)	**(2g)**	9.31 (0.01) H8	1.94 (0.02)	**(3g)**	9.40 (0.01) H8	2.17 (0.01)	**(4g)**	9.64 (0.01) H8	1.61 (0.01)	—	—
(1h) 3.84 (0.01) H6/H5	—	**(2h)**	4.03 (0.02) H6/H5	—	**(3h)**	4.12 (0.01) H6/H5		**(4h)**	3.48 (0.01) H6/H5	—	**(5h)** 3.09 (0.02) H6/H5	5.99(0.01) H6
(1i) 9.67 (0.02) H6/CH_3	—	**(2i)**	9.72 (0.02) H6/CH_3	—	**(3i)**	9.93 (0.01) H6/CH_3		**(4i)**	9.32 (0.02) H6/CH_3	—	**(5i)** 9.33 (0.03) H6/CH_3	5.74 (0.02) H6
(1j) 9.21 (0.01) H6/H5			—		**(3j)**	9.35 (0.01) H6/5		**(4j)**	8.74 (0.01) H6/H5		**(5j)** 8.93 (0.03) H6/H5	5.61(0.02) H6
											(6c) 3.51 (0.03) H6/H5	6.35 (0.02) H1′
											(6d) 9.76 (0.01) H6/CH_3	5.87(0.02) H1′
											(6i) 9.70 (0.01) H6/CH_3	5.69(0.02) H6

[a] Reporter ^{1}H and ^{31}P used for the pH-dependent titration are shown by aromatic H5/H6/H8/H2/CH_3(T) and 5′- and 3′-phosphodiester-phosphorus respectively. The error for the curve-fitting/Hill Plot analysis for pK_a determination is shown in parenthesis (see Experimental Section). [b] H5 is buried with H1′ in **2c**. We observe no significant 5′-phosphate protonation (pK_a 1.6–1.8) as evident from the negligible $\Delta\delta^{31}$P(5′) shift between pH 1 to 7 (0.051 ppm upfield for **1b**, 0.014 ppm upfield for **2b**, 0.052 ppm upfield for **3b** and 0.043 ppm upfield for **4b**). On the other hand, we see some definite 3′-phosphate protonation in **1b** and **2b**, but not in **3b** and **4b** between pH 1 to 7 ($\Delta\delta^{31}$P(3′): 0.230 ppm upfield for **1b**, 0.215 ppm upfield for **2b**, 0.085 ppm upfield for **3b** and 0.033 ppm upfield for **4b**).

of how a change of the electronic character of the 2′-sugar-substitution affects the electronic properties of the nucleobase basing on 47 different model nucleoside 3′,5′-bis- and 3′-mono-ethylphosphates as shown Fig. 1.

We here provide unambiguous pK_a evidence (Table 1) showing that the electronic character of the nucleobase is dynamically modulated as the substitution pattern in the pentose-sugar moiety changes in the model 47 different analogs of nucleoside 3′,5′-bis-ethylphosphates and nucleoside 3′-mono-ethylphosphates[20] (series **1–6** in Fig. 1). Since these model compounds are monomers and all the measurements have been made at a low concentration (1 mM), we can safely rule out any stacking interactions or any effect of the nearest-neighbor promoted modulation[18,19] of the electronic properties.[18, 19]

It has been known for some time now that due to the electrostatic interactions between the negatively charged 5′-phosphate and the nucleobase, the pK_a value of the nucleobase in nucleoside 5′-phosphate increases by *ca.* 0.2–0.5 pK_a unit. Hence all differences in pK_a found for nucleobases between nucleoside 3′,5′-bis-ethylphosphates (**1a–e/2a–d/3a–e/4a–e/5c–e/6d**) and the corresponding 3′-mono-ethylphosphate analogs with 5′-OH group (**1f–j/2f–i/3f–j/4f–j/5h–j/6i**) can be attributed to the electrostatic effect of the 5′-phosphate group.

Results and discussion

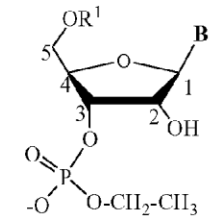
The error for the pK_a estimation reported in this work using pH-dependent ^{1}H chemical shift is found to be 0.01–0.03 [Table 1 and the Supplementary information†]. The following is a summary of our observations, based on the pK_a determinations (Table 1) of 47 compounds (Fig. 1), on how the alteration of the chemical nature of 2′-substituent alters the nucleobase pK_a:

(1) Substituent effect of 2′-OMe *versus* 2′-OH or 2′-deoxy in the alteration of the nucleobase pK_a

Comparison of the pseudoaromatic properties of 9-guaninyl in guanosine 3′,5′-bis-ethylphosphates shows that the –OMe substitution at C2′ (compare **1b/2b/3b**) drives the pK_a of N1 in 9-guaninyl to be more basic than that in nucleotides with 2′-OH or in the 2′-deoxy counterpart in contradistinction to those in the corresponding adenine series (compare **1a/2a/3a**), cytosine series (compare **1c/2c/3c**), or thymine series (compare **1d/2d/3d**). On the other hand, the pK_a of N1 in 9-guaninyl in guanosine 3′-mono-ethylphosphates is almost identical for both 2′-OH and 2′-OMe analogs, **1g** and **2g**, respectively. In contrast, the N1 of 9-adeninyl moiety in 2′-O-methyladenosine 3′-mono-ethylphosphate **2f** is more basic than those of compound with free 2′-OH **1f** and 2′-deoxy counterparts **3f**, and also in comparison with the corresponding guanosine, cytosine and thymine derivatives (compare **1g/2g/3g**, **1h/2h/3h**, **1i/2i/3i**, Table 1).

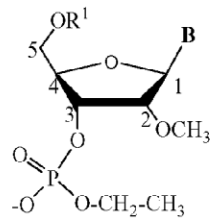
(2) Comparison of the 2′-OMe effect with 1′,2′-conformationally-constrained oxetane systems

(i) Comparison of the pK_a of N1, N3 or N7 in 2′-O-methyl-ribonucleoside 3′,5′-bis-ethylphosphates **2a–2d** with those of the North-East conformationally-constrained oxetane[22] counterparts **4a–4d** show that the 1′,2′-*cis*-fused oxetane moiety is probably



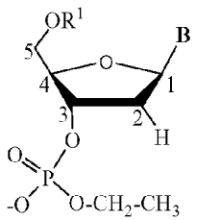
1a: B = A, $R^1 = PO_2^-$-OEt
1b: B = G, $R^1 = PO_2^-$-OEt
1c: B = C, $R^1 = PO_2^-$-OEt
1d: B = T, $R^1 = PO_2^-$-OEt
1e: B = U, $R^1 = PO_2^-$-OEt

1f: B = A, R^1 = H
1g: B = G, R^1 = H
1h: B = C, R^1 = H
1i: B = T, R^1 = H
1j: B = U, R^1 = H



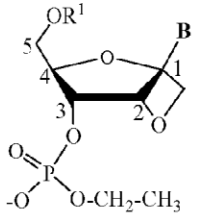
2a: B = A, $R^1 = PO_2^-$-OEt
2b: B = G, $R^1 = PO_2^-$-OEt
2c: B = C, $R^1 = PO_2^-$-OEt
2d: B = T, $R^1 = PO_2^-$-OEt

2f: B = A, R^1 = H
2g: B = G, R^1 = H
2h: B = C, R^1 = H
2i: B = T, R^1 = H



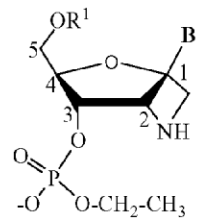
3a: B = A, $R^1 = PO_2^-$-OEt
3b: B = G, $R^1 = PO_2^-$-OEt
3c: B = C, $R^1 = PO_2^-$-OEt
3d: B = T, $R^1 = PO_2$-OEt
3e: B = U $R^1 = PO_2^-$-OEt

3f: B = A, R^1 = H
3g: B = G, R^1 = H
3h: B = C, R^1 = H
3i: B = T, R^1 = H
3j: B = U, R^1 = H



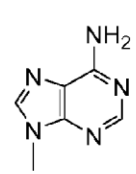
4a: B = A, $R^1 = PO_2^-$-OEt
4b: B = G, $R^1 = PO_2^-$-OEt
4c: B = C, $R^1 = PO_2^-$-OEt
4d: B = T, $R^1 = PO_2^-$-OEt
4e: B = U, $R^1 = PO_2^-$-OEt

4f: B = A, R^1 = H
4g: B = G, R^1 = H
4h: B = C, R^1 = H
4i: B = T, R^1 = H
4j: B = U, R^1 = H

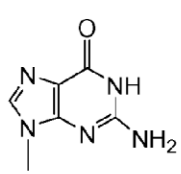


5c: B = C, $R^1 = PO_2^-$-OEt
5d: B = T, $R^1 = PO_2^-$-OEt
5e: B = U, $R^1 = PO_2^-$-OEt

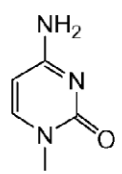
5h: B = C, R^1 = H
5i: B = T, R^1 = H
5j: B = U, R^1 = H



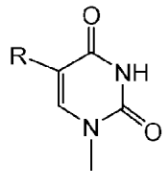
9-Adeninyl (A)



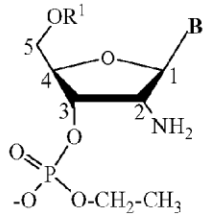
9-Guaninyl (G)



1-Cytosinyl (C)



R = H [1-Uracilyl (U)]
R = Me [1-Thyminyl (T)]



6c: B = C, $R^1 = PO_2^-$-OEt
6d: B = T, $R^1 = PO_2^-$-OEt
6i: B = T, R^1 = H

Fig. 1 Chemical structure of the 3′,5′-bis-ethylphosphates (**a–e**) and 3′-ethylphosphates (**f–j**) ribo-(series **1**), methoxy-(series **2**), deoxy-(series **3**), oxetane-(series **4**), azetidine-(series **5**) and 2′-amino-(series **6**) nucleotides investigated.

exerting more electron-withdrawing effect on N7 of the 9-guaninyl derivative [compare **2b** with **4b**: $\Delta pK_a(^{N7}G) = 0.39$] and N3 of the pyrimidine nucleobases [compare **2c**/4c$\Delta pK_a(^{N3}C) = 0.72$ and **2d**/**4d** $\Delta pK_a(^{N3}T) = 0.43$]. (ii) In contradistinction, a comparison of the pK_a of N1, N3 or N7 amongst the corresponding 3′-mono-ethylphosphate analogs with free 5′-OH group shows [**2f–2i** *versus* **4f–4i**] that the basicity of N1 (A/G) of 2′-OMe analogs in the purine series (**2f**/**2g**) is less than that in the oxetane constrained purine analogs (**4f**/**4g**) [$\Delta pK_a(N1)_{OMe\text{-}oxetane} = -0.13$ and -0.33 pK_a unit for A and G residues, respectively], whereas in the oxetane constrained pyrimidine (T/C) analogs (**4h**/**4i**) the pK_as (N3) have been found to be less compared to the 2′-OMe analogs (**2h**/**2i**) [$\Delta pK_a(N3)_{OMe\text{-}oxetane} = 0.40$ pK_a unit for T residues and 0.55 pK_a unit for C residues].

We see here an interplay of three competing effects: (i) the 5′-phosphate effect[23,24] near the nucleobase, (ii) the anomeric effect[23,25] [$n_{O4'} \rightarrow \sigma^*_{(C1'-N1/N9)}$], and (iii) the distance-dependent inductive effect between the 2′-substituent and the protonation/deprotonation site in the pyrimidine and purine nucleobases.

While the pK_a of N1 in the adenosine and guanosine 3′-monophosphates (**1f**/**1g**) is more acidic (3.11/9.27, respectively) compared to that of the corresponding oxetane (**4f**/**4g**: 3.59/9.64, respectively) and the 2′-OMe-(**2f**/**2g**: 3.46/9.31, respectively) counterparts, the pK_as of N7 in the guanosine 3′-monophosphates [**1g** (1.90)/**2g** (1.94)/**4g** (1.61)] follow a different trend with the conformationally constrained oxetane derivative (**4g**) being the most acidic (1.61) and conformationally free ribo- and 2′-OMe-nucleotides (**1g**/**2g**) being similarly more basic (1.90/1.94). This suggests that the oxetane group in **4g** has a profound effect in withdrawing the charge density from the imidazole part of the 9-guaninyl base compared to that of the 2′-OMe and 2′-OH substituents in **2g**/**1g**. The examination of the oxetane effect *vis-a-vis* 2′-OMe effect on the pK_a of N3 in pyrimidine nucleosides (**4h**/**2h** and **4i**/**2i**) shows that the 1′,2′-*cis*-fused oxetane group has indeed a more electron-withdrawing influence at N3 in 1-cytosinyl-oxetane derivative **4h** by 0.55 pK_a unit compared to that of 1-cytosinyl-2′-OMe analog **2h**. Similarly, the N3 in 1-thyminyl-oxetane derivative **4i** is 0.40 pK_a unit less basic compared to that of 1-thyminyl-2′-OMe derivative **2i**. Hence, these comparisons of relative basicity of purines and pyrimidine nucleotides with 2′-OMe substituent (**2b–d**/**2g–i**) with those of fused-oxetane counterparts (**4b–d**/**4g–i**) show that the electron-withdrawing effect of the oxetane works most effectively on the reduction of the electron-density in the pyrimidine moieties in **4c**/**4d**/**4h**/**4i** as well as on the imidazole part of the purine system (for example, compare pK_a of N7 in oxetane-guanine analogs **4b**/**4g** with that of the 2′-OMe analog **2b**/**2g** in Table 1), whereas the back-donation of the charge by the anomeric effect[23,25] works best for the pyrimidine part of oxetane constrained 9-adeninyl (**4a**/**4f**) and 9-guaninyl (**4b**/**4g**) (evidenced by an increase of the N1 basicity) because of the North-East constrained nature of the sugar in which the anomeric effect is overwhelmingly preferred.[23,25]

(3) Effect of the conformationally-constrained 1′,2′-*cis*-fused azetidine systems

A comparison of the pK_as amongst the 1′,2′-*cis*-fused azetidine-nitrogen[26] in the pyrimidine nucleotides (**5c–e**) uniquely shows that the azetidine-nitrogen in the 1-cytosinyl nucleotide (**5c**) is ~0.2 pK_a unit more basic compared to that of the azetidine constrained 1-thyminyl (**5d**) or 1-uracilyl (**5e**) nucleotides, thereby showing the donation of the charge from 1-cytosinyl moiety to the azetidine part owing to the relatively high electron-rich aromatic character of the 1-cytosinyl residue in **5c** compared to the 1-thyminyl **5d** or 1-uracilyl **5e** moieties. A very similar comparison of azetidine-nitrogen pK_a in azetidine constrained[26] cytidine 3′,5′-bis-ethylphosphate **5c** with that of its 3′-mono-ethylphosphate analog **5h** shows that the former is more basic compared to the latter. In a similar way, azetidine nitrogens in the 1-thyminyl analog **5d** and 1-uracilyl analog **5e** are more basic compared to that in their corresponding 3′-mono-ethylphosphate analogs **5i** and **5j** by *ca.* 0.2 pK_a unit.[27] An identical pK_a was observed for the azetidine-nitrogen for 3′,5′-bis-ethylphosphate azetidine constrained 1-thyminyl and 1-uracilyl residues (**5d**/**5e**), whereas the azetidine-nitrogen in 3′-mono-ethylphosphate analog of 1-uracilyl **5j** is 0.13 pK_a unit more acidic compared to that of corresponding 1-thyminyl analog **5i**. This suggests that the 5-methyl group in 1-thyminyl derivative **5i** can increase the constituent azetidine-nitrogen basicity through its electron donating effect which is absent in the 1-uracilyl derivative **5j**.

(4) Comparison of the effect of conformationally-constrained 1′,2′-*cis*-fused-azetidine and -oxetane systems on the modulation of the nucleobase pK_a

The larger reduction of basicity at *N3* of the 1-cytosinyl residue in the conformationally-constrained azetidine blocks (**5c**/**5h**) compared to those of the oxetane counterparts (**4c**/**4h**) is owing to the fact that azetidine-nitrogen becomes almost fully protonated when the pH of the medium is equal to the pK_a of the 1-cytosinyl-*N*3, thereby exerting a reduced *N3* basicity by 0.30 (3′,5′-bis-ethylphosphate) and 0.39 (3′-mono-ethylphosphate) pK_a units respectively.

(5) Comparison of nucleobase pK_a modulation in conformationally-constrained azetidine modified nucleotide with that of 2′-amino analogs

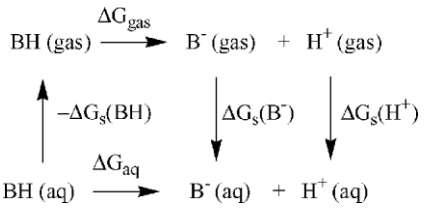
In order to understand the effect of conformational constraints introduced by the azetidine ring in 1-cytosinyl **5c** and 1-thyminyl **5d** derivatives, we have compared the pK_a of the azetidine-nitrogen with the corresponding 2′-amino group in analogs **6c** and **6d**: (a) Interestingly, the pK_a of the 2′-amino group in the 1-cytosinyl analog **6c** (6.35) shows that it is more basic compared to that of the azetidine-nitrogen in **5c** (6.08) [$\Delta pK_a = 0.27$]. On the other hand in the 1-thyminyl analogs, the pK_a of 2′-amino group in **6d** and that of the azetidine-nitrogen in **5d** were identical. (b) In contradistinction, the more electron withdrawing nature of the azetidine ring compared to the $-NH_2$ group is reflected from the relative pK_a of *N3* of 1-cytosinyl-azetidine (3.24) *versus* 1-cytosinyl-2′-amino (3.51) analogs (**5c**/**6c**: $\Delta pK_a = 0.27$) compared to 1-thyminyl-azetidine (9.60) *versus* 1-thyminyl-2′-amino (9.76) analogs (**5d**/**6d**: $\Delta pK_a = 0.16$). It shows that the 1′,2′-*cis*-fused azetidine ring has a more profound electron-withdrawing effect compared to the freely rotating 2′-amino group.

(6) Correlation of the chemical shifts of the H-2′ in different 2′-substitueted derivatives with the modulated pK_a of the nucleobases

Since our above observations show that the chemical nature of the 2′-substituent affects the electronic properties of the imidazole part (*N7*) of guanine base and the *N3* of the pyrimidines, we argued that the magnitude of the electronegativities of different 2′-substituents and also of the 1′,2′-conformationally constrained system (oxetane/azetidine) probably affects the chemical shift of H2′-sugar proton in various nucleotides (Fig. 1). Hence, we have plotted pK_a of *N7* of 9-guaninyl for different 2′-substituted and 1′,2′-oxetane constrained 3′,5′-bis-ethylphosphate analogs [Fig. 2D] as well as the pK_a of *N3* of 3′,5′-bis-ethylphosphate pyrimidine (T/U/C) derivatives [Fig. 2A–C] with their corresponding δH2′ at neutral pH. Indeed, we find a straight correlation with high Pearson's correlation coefficients (0.85–0.97), thereby demonstrating that the chemical nature of 2′-substituent which alters the chemical shift of the H2′ also directly alters the pseudoaromaticity of the nucleobases (Fig. 2).

Theoretical pK_a of nucleoside

In order to understand the mechanism of modulation of the pK_a in the constituent nucleobases by different 2′-substituents, we have performed a set of *ab initio* calculations utilizing the closed shell Hartree-Fock (HF) method and 6-31G** basis set to calculate pK_a values from the traditional thermodynamic cycle (1):



(1)

A key problem of the absolute pK_a determination using the thermodynamic circle above are the values to use for $\Delta G_s(H^+)$ and $\Delta G(H^+)$ which in the literature[28] vary in the range from 258.32 kcal mol^{-1} to −264 kcal mol^{-1} and from −6.04 to −7.76 kcal mol^{-1}, respectively. Since we have aimed to use the theoretical pK_as not as absolute values but as a tool to understand the mechanistic basis of the observed modulation of the experimental pK_as, the choice of these energies introduces only systematic error which is taken care of by the correlation parameters, thus allowing to make an arbitrary choice which was $\Delta G_s(H^+) = -262.5$ kcal mol^{-1},[29] and $\Delta G(H^+) = -6.28$ kcal mol^{-1}.[30]

The theoretical pK_a values obtained for different nucleosides have been compared with the experimental pK_as of respective nucleosides 3′,5′-bis-ethylphosphate. Due to hardware limitations the explicit phosphates have not been included in the simulations, but a good correlation between the NMR titration based experimental pK_a values and *ab initio* based theoretical pK_as (Fig. 3, Table 2) has however been observed for N3, N1 and N7 protonation/de-protonation of 2′-deoxy-, ribo-, methoxy-, amino-, oxetane- and azetidine-nucleosides as well as for N3 and

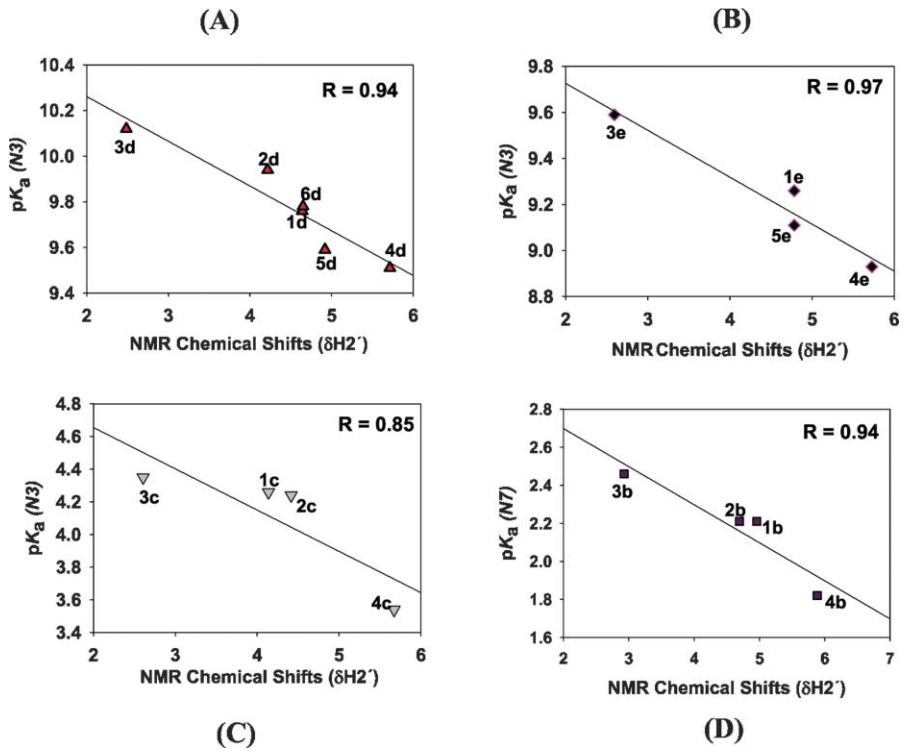


Fig. 2 Correlation plots of pK_as of *N3* of 3′5′bis-ethyl-phosphates of T/U/C, (see Fig. 1 for structures and Table 1 for pK_a). [panels A, B, C] and 3′,5′bis-ethyl phosphates of *N7*(G) [panel D] with different 2′-substituents *versus* their δH2′ (neutral pH) show that the pK_as of pyrimidine *N3* and imidazole *N7* in 9-guaninyl derivatives show that they are linearly correlated, giving high Pearson's correlation coefficients (*R*) (shown in the top right hand corners of the plots).

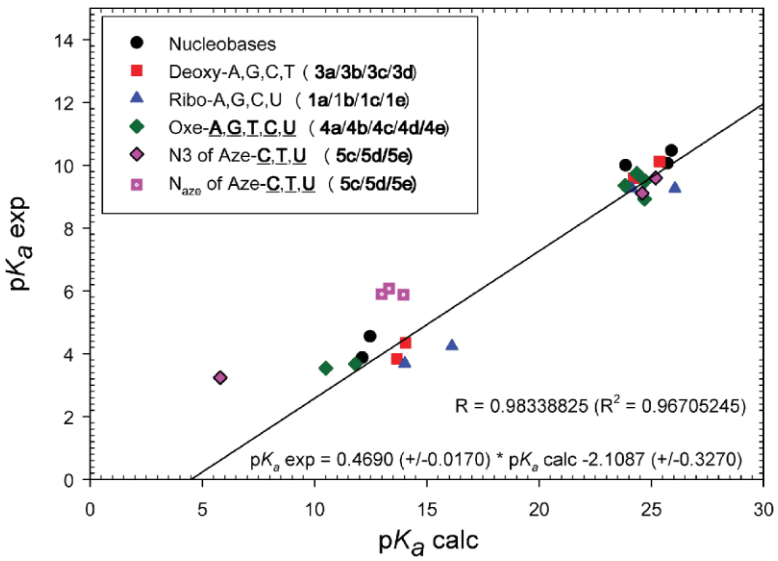


Fig. 3 Correlation between experimental pK_as obtained by NMR titration (Table 1) and the calculated pK_as (Table 2, Table S1 in the Supporting Information). Compound numbers (Fig. 1) of the corresponding 3′,5′-bis-ethylphosphates are shown in parenthesis.

N7 of cytosine, guanine, thymine and uracil nucleobases. Although the phosphate effect can clearly be seen in the experimental data (compare pairwise pK_a values in 3′,5′-bis-ethylphosphate-*versus* 3′-mono-ethylphosphate with free 5′-OH in the ribo-series **1**: **1a/1f**, **1b/1g**, **1c/1h**, **1d/1i**, **1e/1j** as well as in the other series of compounds in **2–6** in Fig. 1 and Table 1), the linear correlation ($R = 0.98$, pK_a (exp) = 0.4690 (±0.0170) * pK_a (calc) −2.1087 (±0.3270)) between experimental pK_a values for the bis-ethylphosphates and calculated pK_a values for nucleosides (Fig. 4) suggests that the phosphate effect is very similar for the protonation/de-protonation at pK_as far away from the phosphate pK_a of 1.3–2.1, which probably leads to a systematic error in the calculated pK_as. Thus, only calculations of pK_a corresponding to N3 protonation in cytosines (compounds **1c/1h**, **2c/2h**, **3c/3h**, **4c/4h**, **5c/5h**) and N7 protonation in guanines (compounds **1b/1g**, **2b/2g**, **3b/3g**,**4b/4g**) as well as probably the azetidine nitrogen pK_as in compounds **5c–5j** should be sensitive to exclusion of the explicit phosphates in the model compounds. This linear correlation also indicates that even such a restricted model allows us to calculate the absolute values of nucleobase pK_as in various nucleotide analogs (corrected using above mentioned parameters of the correlation line) from the *ab initio* simulation with up to 0.5 pK_a unit accuracy).

Method of calculations

Molecular geometries of all protonated and de-protonated nucleosides have been optimized in the gas phase and the effect of solvation has been estimated using Baron and Cossi's implementation of the polarizable conductor CPCM model[31] as implemented

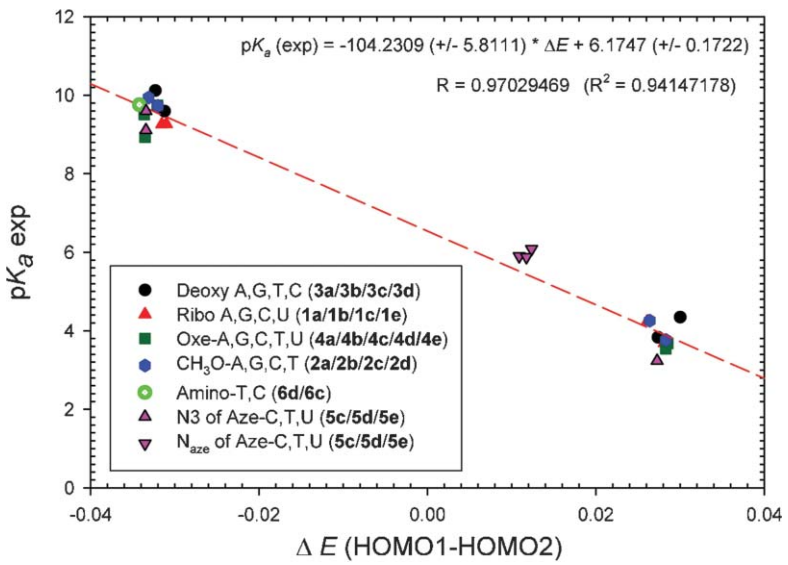


Fig. 4 Correlation between experimental pK_a (Table 1) and the difference of ground and protonated (de-protonated) HOMO orbital energies ΔE (HOMO1-HOMO2, a.u.) (Table 3, Table S2 in the Supporting Information). Compound numbers (Fig. 1) of the corresponding 3′,5′-bis-ethylphosphates are shown in parentheses.

Table 2 Theoretical proton affinities (PA), Gibbs free energies (gas phase and solvation), and the theoretical pK_a values of the nucleobases in ribo-, 2′-deoxy-, 2′-amino-, 2′-methoxy-, oxetane and azetidine nucleosides as well as the experimental pK_a values for the corresponding bis-ethylphosphate nucleotides (Table 1). Complete table is provided in the Supporting Information (Table S1)†

Nucleobase/nucleoside[a]	PA (gas)/kcal mol^{-1}	$\Delta\Delta G_{gas+}$ $\Delta\Delta G_s$/kcal mol^{-1}	$\Delta\Delta G_s$ kcal mol^{-1}	pK_a (calc)	pK_a (exp)[b]
Adenine: N1	231.61	285.53	53.46	12.12	3.88
Guanine: N1	352.26	301.54	−51.02	23.85	10.00
Thymine: N3	362.65	304.32	−58.33	25.89	10.47
Cytosine: N3	236.87	286.01	48.80	12.47	4.56
Uracil: N3	362.70	304.08	−58.58	25.71	10.06
Deoxy-A: N1	238.18	287.65	49.92	13.67	3.83 (**3a**)
Deoxy-G: N1	349.59	302.05	−47.78	24.22	9.59 (**3b**)
Deoxy-C: N3	247.64	288.18	41.21	14.06	4.35 (3c)
Deoxy-T: N3	358.99	303.62	−55.53	25.37	10.12 (**3d**)
Ribo-A: N1	236.32	288.13	52.62	14.02	3.69 (**1a**)
Ribo-G: N1	349.89	301.93	−48.41	24.13	9.27 (**1b**)
Ribo-C: N3	242.93	291.00	48.69	16.12	4.24 (**1c**)
Ribo-U: N3	350.05	304.54	−46.03	26.05	9.26 (**1e**)
Oxe-A: N1	233.02	285.13	52.31	11.82	3.68 (**4a**)
Oxe-G: N1	344.36	302.23	−42.71	24.35	9.74 (**4b**)
Oxe-C: N3	238.46	283.33	44.98	10.50	3.54 (**4c**)
Oxe-T: N3	352.40	302.72	−50.25	24.72	9.51 (**4d**)
Oxe-U: N3	352.34	302.68	−50.18	24.69	8.93 (**4e**)
Aze-A: N1	232.94	285.54	52.98	12.12	—
:N_{aze}	247.97	285.99	40.16	12.45	
Aze-G: N1	348.28	301.25	−47.51	23.63	
N_{aze}	231.55	290.00	57.28	15.39	
Aze-C: N3	241.65	276.91	35.42	5.80	3.24 (**5c**)
N_{aze}	232.96	287.18	53.74	13.32	6.08 (**5c**)
Aze-T: N3	355.13	303.37	−52.32	25.19	9.60 (**5d**)
N_{aze}	227.41	288.04	59.41	13.96	5.88 (**5d**)
Aze-U: N3	357.30	302.56	−55.16	24.59	9.11 (**5e**)
N_{aze}	230.45	286.74	55.90	13.00	5.90 (**5e**)
CH_3O–A: N1	237.98	287.61	50.73	13.64	3.77 (**2a**)
CH_3O–G: N1	346.56	304.05	−43.700	25.69	9.73 (**2b**)
CH_3O–C: N3	249.41	287.74	39.84	13.74	4.26 (**2c**)
CH_3O–T: N3	355.55	305.32	−50.98	26.62	9.94 (**2d**)

*Calculated pK_as of phosphates-free nucleosides are compared to the corresponding experimental pK_as of 3′,5′-bis-ethylphosphate nucleotides.[a] The protonation or de-protonation sites are shown by atom name and numbering. [b] Compound numbers (Fig. 1) of the 3′,5′-bis-ethylphosphates are shown in parenthesis.

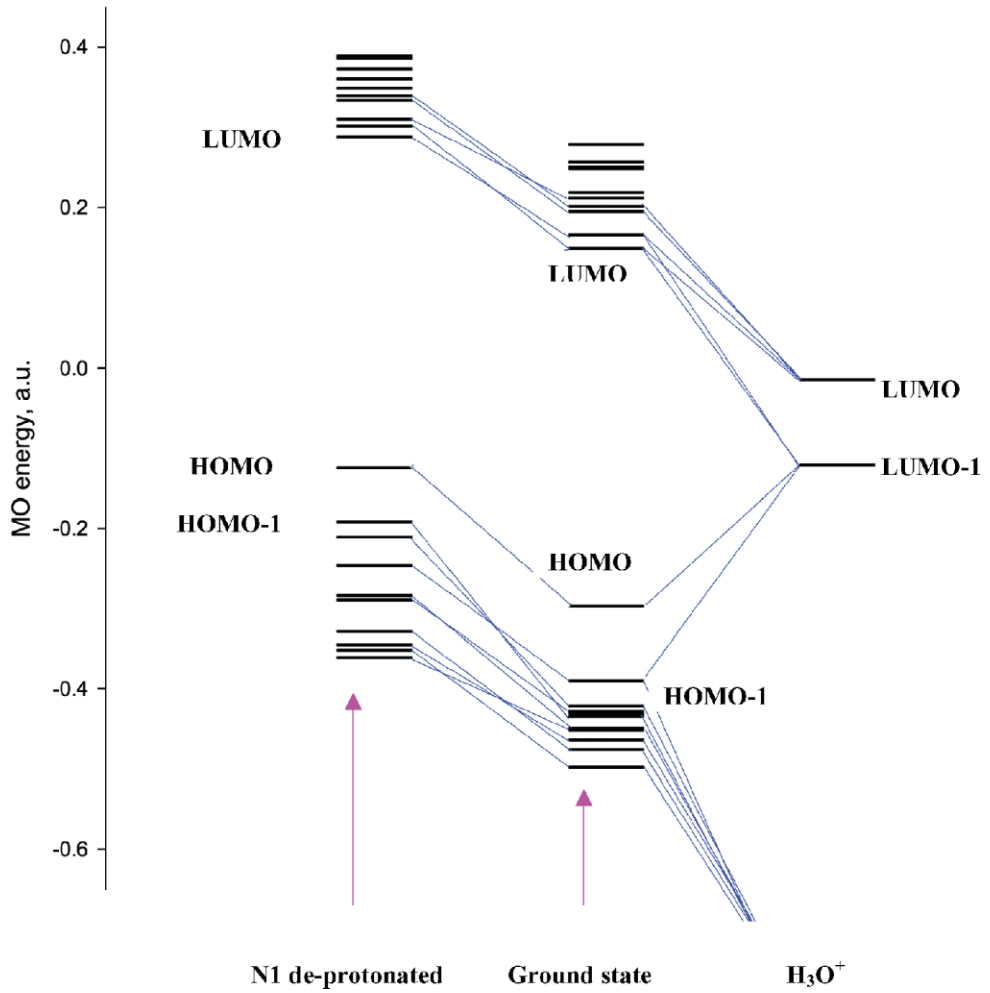
in the Gaussian 98[32] program package. The accurate prediction of absolute pK_a values presents a considerable challenge to the theoretical chemistry community, even to date, due to the fact that relatively small errors in the gas-phase thermodynamics and the solvation energy calculations lead to often large errors in absolute pK_a values. For example, an error of 1.36 kcal mol^{-1} in the total free Gibbs energy results in error of 1 pK_a unit (for discussion see Liptak and Shields[28] and references therein). We therefore have adopted an empirical correction scheme similar to that of Klicic and co-workers.[33] This allows us to compensate for deficiencies in both *ab initio* and solvation models and with the "right" set of training compounds (we have used a full set of A, G, C, T, U native nucleobases and deoxy-A,G,C,T and ribo-A,G,C,U nucleosides; 24 compounds altogether in the dataset) it provides reasonably good correlation (with errors of up to 0.5 pK_a unit) to experimental pK_a values *via* the following correction scheme: pK_a (scaled) = A pK_a (calc) + B, where A and B are constants determined from the correlation of the training set of compounds. Generally, these coefficients are basis set and theoretical method dependent; they can as well vary drastically for different functional groups[33] and thus cannot be applied to calculate absolute pK_as in a diverse groups of chemical compounds. This correlation scheme has also assumed a systematic error obeying linear free-energy relationship which can generally be doubtful. However, the scope of achieving 0.5 pK_a accuracy with reasonable computational efforts and within well-grounded theoretical (*ab initio*) model, overweighs the deficiencies of the empirical correction *per se* and prompts for a cautious use of this model.

Relationship between pK_a and frontier orbital structure and energy

Clearly, any rationalization and prediction of the sugar substitution effect at C2′ on the nucleobase pK_a in the form of Hammett constant-like correlation, as in substituted aromatic systems, is not possible as the experimental pK_as exhibit different trends for the different sets of nucleotides (see Fig. 1 and Table 1). Thus, for the adeninyl (A) group the pK_a follow the following trend: oxetane-(**4a**) ≈ ribo-(**1a**) < MeO-(**2a**) < deoxy-(**3a**); for the guaninyl (G) group it is ribo-(**1b**) < deoxy-(**3b**) < MeO-(**2b**) ≈ oxetane-(**4b**); for the cytosinyl (C) group it is azetidine-(**5c**) < amino-(**6c**) ≈ oxetane-(**4c**) < ribo-(**1c**) ≈ MeO-(**2c**) < deoxy-(**3c**); and for the thyminyl (T) group, the trend is oxetane-(**4d**) < azetidine-(**5d**) < amino-(**6d**) ≈ ribo-(**1d**) < MeO-(**2d**) < deoxy-(**3d**), which is similar to the trend found for Us: oxetane-(**4e**) < azetidine-(**5e**) < ribo-(**1e**) < deoxy-(**3e**). The explanation of these trends and the intrinsic mechanism behind the 2′-sugar modification effect on the nucleobase pK_a can

be derived from the molecular orbital description of the electronic structure of corresponding nucleosides (Table 3 and Table S2 in Supporting Information†). The frontier orbital theory[34–36] suggests that only the highest occupied (HOMO) and lowest unoccupied (LUMO) molecular orbitals (MOs) should be taken into account to predict the chemical reactivity. We have cautiously adopted this approach, but have also taken into account the lower occupied valence and higher unoccupied virtual MOs.

According to the frontier orbitals approach,[34–36] the main interactions to consider are overlap between HOMO of nucleoside and LUMO of hydroxonium ion H_3O^+ and *vice versa.* Since the protonation (de-protonation) process does not change the number of electrons in a given compound, and the HOMO of hydroxonium ion is much lower in energy than LUMO of the nucleoside (Scheme 1, off the scale), the overlap between the HOMO of the nucleosides and LUMO of the hydroxonium ion should play a decisive role. Since the hydroxonium ion's LUMO energy is constant for all the compounds in question, it is reasonable to expect that experimental pK_as correlate only with the energy of nucleoside's HOMO orbitals which change upon protonation/de-protonation. Indeed a good correlation ($R = 0.97$) between energy difference of the ground state HOMO and protonated/de-protonated HOMO (ΔHOMO) molecular orbitals, obtained using CPCM solvation model, and the experimental pK_a (Fig. 4) has been observed. The gas phase ΔHOMO energies *versus* experimental pK_as have also shown somewhat lower correlation ($R = 0.95$). Although the ΔHOMO energy represents only part of the total energy change in the system due to protonation/de-protonation, it appeared to show a good correlation even for the nitrogen pK_as of azetidine modified nucleosides (abs error of 0.3 pK_a units). This somewhat better correlation probably reflects variation of an error in estimation of solvation energy by the CPCM model. Despite a good correlation between HOMO orbital energies and pK_as, the HOMO orbital itself appeared to give only limited clues to the protonation (de-protonation) mechanism as it was found to be a typical π orbital 100% localized on the nucleobase for all purine nucleosides (protonated and de-protonated) and 55–98% localized on the nucleobase for the pyrimidine nucleosides (Table 3, Table S2 in the Supporting Information†). As the MO energy effectively includes (through HF Hamiltonian) all interactions for the particular electron(s) on the HOMO, its energy indeed reflects properties (such as pK_a) attributed to these electron(s), but the orbital structure does not reflect total electron distribution which is greatly influenced by the valence electrons



Scheme 1 MO diagram (gas phase, MO energy shown in Table 3) of the deoxy-G in the N1 de-protonated and ground states shown with the LUMO and LUMO + 1 orbitals of the hydroxonium ion. Occupied MOs with the orbital energies below −0.6 a.u. are not shown.

Table 3 Frontier orbitals and their respective energies (a.u.) of the ground, protonated and de-protonated states of deoxy A,G,C,T and ribo-U nucleosides. Complete table of the frontier orbitals of the 2′-ribo, 2′-deoxy, 2′-amino-, 2′-methoxy-, oxetane- and azetidine-nucleosides is provided in the Supporting Information (Table S2). MOs have been visualized using gOpenMol[41,42]

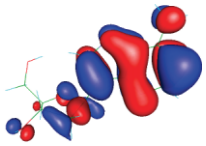
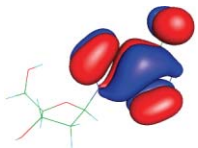
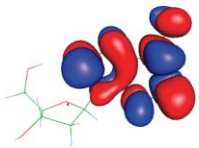
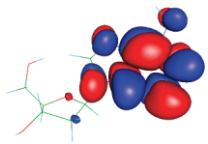
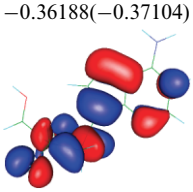
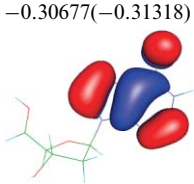
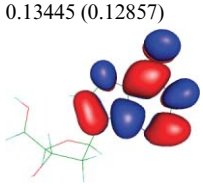
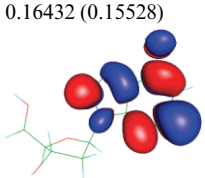
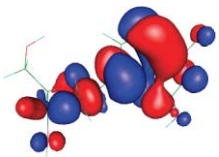
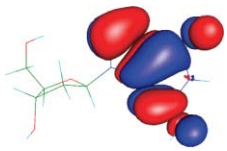
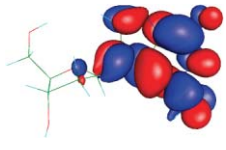
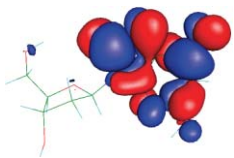
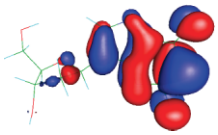
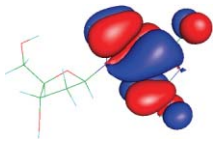
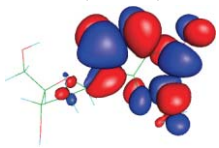
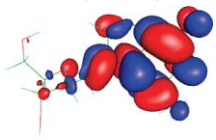
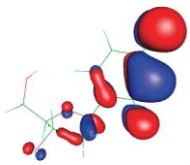
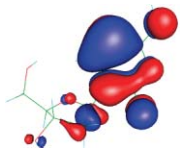
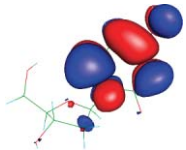
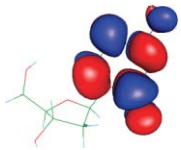
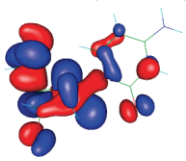
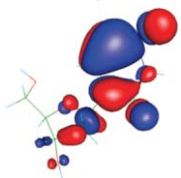
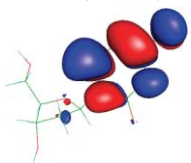
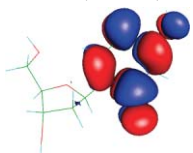
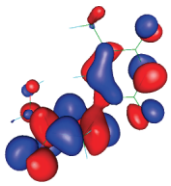
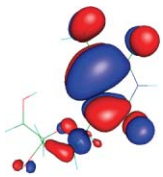
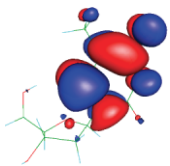
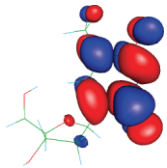
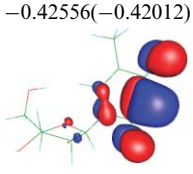
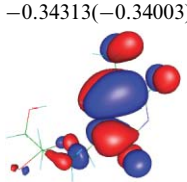
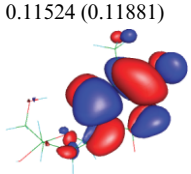
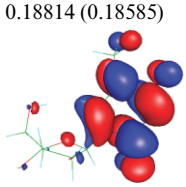
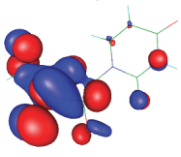
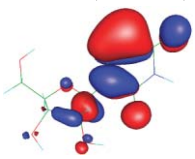
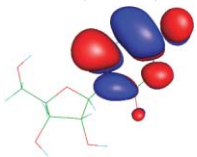
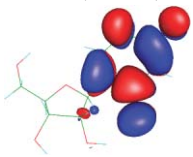
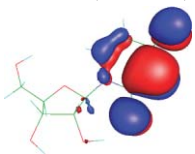
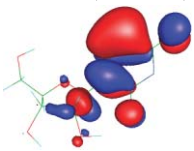
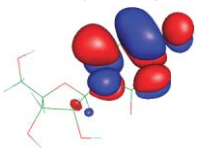
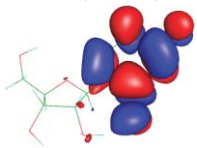
	HOMO-1	HOMO	LUMO	LUMO + 1
dA ground state				
MO energy (a.u)	−0.36188(−0.37104)	−0.30677(−0.31318)	0.13445 (0.12857)	0.16432 (0.15528)
dA N1 protonated				
MO energy (a.u)	−0.51492(−0.40391)	−0.46701(−0.34055)	−0.04077(0.09709)	−0.02470(0.10780)
dG ground state				
MO energy (a.u)	−0.38989(−0.39480)	−0.29724(−0.30490)	0.14929 (0.14697)	0.16615 (0.15953)
dG N1 deprotonated				
MO energy (a.u)	−0.19183(−0.34728)	−0.12450(−0.27368)	0.28839 (0.17518)	0.30231 (0.19678)
dC ground state				
MO energy (a.u)	−0.37048(−0.37725)	−0.32956(−0.33362)	0.12346 (0.12421)	0.19234 (0.18366)
dC N3 protonated				
MO energy (a.u)	−0.49830(−0.43517)	−0.05996(−0.36364)	0.02430 (0.08135)	0.04809 (0.15654)
dT ground state				
MO energy (a.u)	−0.42556(−0.42012)	−0.34313(−0.34003)	0.11524 (0.11881)	0.18814 (0.18585)
dT N3 deprotonated				

Table 3 (*Cont.*)

	HOMO-1	HOMO	LUMO	LUMO + 1
MO energy (a.u)	−0.17866(−0.35020)	−0.17533(−0.30774)	0.28690 (0.15558)	0.29572 (0.21448)
rU ground state				
MO energy (a.u)	−0.42612(−0.41494)	−0.36243(−0.34880)	0.10560 (0.11912)	0.16963 (0.18040)
rU N3 deprotonated				
MO energy (a.u)	−0.19997(−0.35654)	−0.18637(−0.31726)	0.28485 (0.15633)	0.30898 (0.20640)

occupying energy levels below HOMO. The electrons on HOMO-1 are expected to give highest contributions to the substitution effect as they are the closest in energy to HOMO electrons.

Clearly, the HOMO electrons contributions to the observed changes in p*K*a are negligible for the purines and the main effect comes from lower lying HOMO-1 π orbital (Table 3, Table S2 in the Supporting Information). In fact, the chemical nature of this orbital is changed upon protonation as HOMO-1 of the pyrimidines in the de-protonated state is stabilized so much that HOMO-2 or HOMO-3 (exact orbital number depends on the substituent) takes its place (see example of MO diagram in Scheme 1 shown for deoxy-G protonation/de-protonation which qualitatively describes the situation for all the nucleosides reported here). On the other hand, the observed changes in pK_a of pyrimidines are apparently dependent on the electrons occupying both the HOMO and HOMO-1 orbitals in the protonated and de-protonated states (Table 3, Table S2 in the Supporting Information). Similar to the effect observed for purines, the nature of HOMO-1 orbital is also changing upon protonation (Scheme 1). For both the purines and pyrimidines, the aglycon part of the HOMO-1 orbitals remains essentially the same in the particular protonated state of the nucleoside within the substitution series while the major variations are observed in the sugar moiety (Table 3, Table S2 in the Supporting Information).

Relation between dipole moment and pK_a

Due to the essentially electrostatic nature of the protonation effect, one could expect it to be dependent on the total charge distribution changing upon protonation (de-protonation) which manifests in a first approximation as a change in the dipole moment. Indeed, the acid–base difference of the respective dipole moments correlates linearly with the values of pK_as (Fig. 5). However, this dependence is not universal and the smaller changes in dipole moments for the aromatic systems like nucleobases lead to bigger changes in the pK_as (green triangles and squares in Fig. 5).

Conclusions and implications

(1) The *cis*-fused 1′,2′-oxetane or -azetidine moiety (shown in Fig. 1) has a stronger electron-withdrawing effect than that of 2′-OMe substituent, which has been evidenced from the pK_a of N3 of 1-cytosinyl, 1-thyminyl, 1-uracilyl and pK_a of N7 of 9-guaninyl analogs (Table 1). On the other hand, the *cis*-fused 1′,2′-oxetane moiety has a poorer electron-withdrawing effect on the N1 of 9-adeninyl and 9-guaninyl system.

(2) Comparison of nucleobase pK_a for 3′,5′-bis-phosphates with that of 3′-monophosphate with 5′-OH group shows that the 5′-phosphate electrostatically enhances the basicity of the nucleobase.

(3) Comparison of pK_a for N1 in 9-adeninyl (**4f**) and in 9-guaninyl (**4g**) in the 3′-monophosphate series shows that the oxetane modification exerts a stronger anomeric effect in sugar moiety [$n_{O4'} \rightarrow \sigma^*_{(C1'\text{-}N1/N9)}$][23,25] compared to that of the corresponding 2′-OMe analogs of 9-adeninyl **2f** and 9-guaninyl **2g**., probably because of its fully-constrained North-East-type conformation.[22]

(4) Correlation plots of the pK_a of N3 of pyrimidine (T/C/U) or pK_a of N7 of 9-guaninyl with the corresponding δH2′ at the neutral pH (Fig. 2) shows that these properties are linearly correlated with high Pearson's correlation coefficients (0.85–0.97) which reflects that the pseudoaromatic character of the nucleobases can be tuned depending upon the chemical nature of the 2′-substituent,[37] which also explains why the T_m of a duplex can be modulated[22,26,38–40] by the chemical nature of the 2′-substituent. The high correlation of δH2′ with the pK_a of the constituent nucleobase clearly suggests that the comparison of the chemical shifts for H2′ in the sugar moiety in a series of 2′-modified nucleosides or oligonucleotides can provide neat information (bypassing direct pK_a measurement) for interrogation of how the pseudoaromatic character of the genetic alphabets has been altered as a function of the electronegativity of the 2′-substituent.

(5) The 5′-phosphate group not only enhances the pK_a of nucleobases but also enhances the pK_a of the 2′-substituent (compare the pK_a of amine or azetidine nitrogen protonation for 3′,5′-bis- and 3′-mono-ethylphosphate derivatives in Table 1).

Experimental section

(A) pH-dependent ^{1}H NMR measurement

All NMR experiments were performed in Bruker DRX-500 and DRX-600 spectrometers. The NMR samples of all 2′-OH

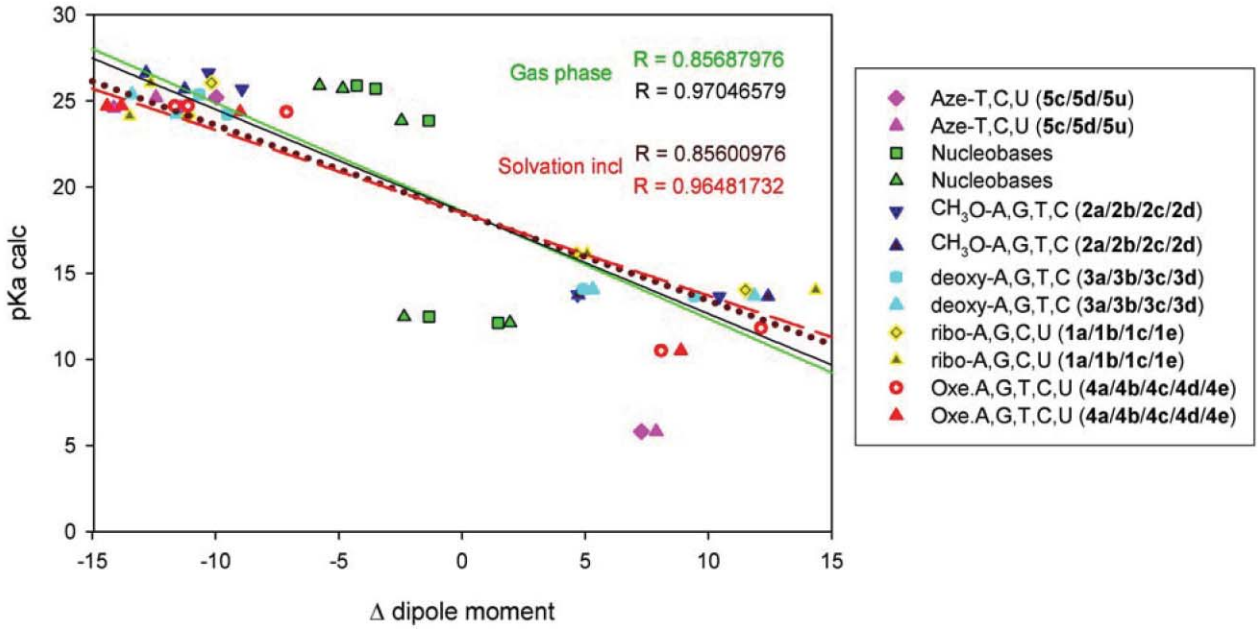


Fig. 5 Correlation of the experimental NMR-titration derived pK_as (Table 1) of the 3′,5′-bis-ethylphosphate nucleotides with the acid–base difference of the calculated dipole moments (Table S3 in Supporting Information) of the respective nucleosides. Compound numbers (Fig. 1) of the corresponding 3′,5′-bis-ethylphosphates are shown in parenthesis.

(**1a–1j**), 2′-OMe (**2a–2i**), 2′-deoxy (**3a–3j**), oxetane constrained (**4a–4j**) and azetidine constrained (**5c–5j**) series of nucleoside 3′,5′-bis-ethylphosphates, Etp(2′-OH/2′-OMe/2′-deoxy/Oxe/Aze)**B** pEt (**B** = Nucleobase), and nucleoside 3′-mono-ethylphosphates, (2′-OH/2′-OMe/2′-deoxy/Oxe/Aze)**B** pEt (**B** = Nucleobase), were prepared in D_2O solution (concentration of 1 mM in order to rule out any chemical shift change owing to self-association) with $\delta_{DSS} = 0.015$ ppm as internal standard. All pH-dependent NMR measurements have been performed at 298 K. The pH values [with the correction of deuterium effect] correspond to the reading on a pH meter equipped with a calomel microelectrode (in order to measure the pH inside the NMR tube) calibrated with standard buffer solutions (in H_2O) of pH 4, 7 and 10. The pD of the sample has been adjusted by simple addition of micro liter volumes of NaOD solutions (0.5M, 0.1M and 0.01M). The pH values are obtained by the subtraction of 0.4 from corresponding pD values [pH = pD − 0.4]. All ^{1}H spectra have been recorded using 128 K data points and 64 scans.

(B) The pH titration of aromatic protons and phosphorus of 3′ and 5′ phosphates and pK_a determination from Hill plot analysis

The pH titration studies were done over the range of pH (1.8 < pH < 12.2), with [0.2–0.3] pH interval for all 2′OH (**1a–j**), 2′-OMe (**2a–i**), 2′-deoxy (**3a–j**), oxetane constrained (**4a–j**) and 1′,2′ azetidine constrained (**5c–j**) series of nucleoside 3′,5′-bis-ethyl-phosphates, Etp(2′-OH/2′ OMe/2′-deoxy/Oxe/Aze)pEt, and (2′-OH/2′ OMe/2′-deoxy/Oxe/Aze)pEt for nucleoside 3′-ethylphosphates. All pH titration studies consist of ~20–33 data points and the corresponding sigmoidal pH metric titration curves for 2′-OMe (**2a–i**), (**4a–j**), (**5c–j**), and (**6c–i**) compounds are given in the Supporting Information (Fig. S2)†. For 2′ OH (**1a–j**) and 2′-deoxy (**3a–j**) pH metric titration plots (see ref. 20). The pH-dependent [over the range of pH 1.8 < pH < 12.2, with an interval of pH 0.2–0.3] ^{1}H chemical shifts (δ, with error ±0.001 ppm) for all compounds show a sigmoidal behavior. The δ^{31}P also shows sigmoidal behavior during the protonation of nitrogens of azetidines and amines in (**5c–j**) and (**6c–i**) [Fig. S4 in Supporting Information]. The pK_a determination is based on the Hill plot analysis using equation: pH = log ((1 − α)/α) + pK_a, where α represents fraction of the protonated species. The value of α is calculated from the change of chemical shift relative to the de-protonated (D) state at a given pH ($\Delta_D = \delta_D - \delta_{obs}$ for de-protonation, where δ_{obs} is the experimental chemical shift at a particular pH), divided by the total change in chemical shift between neutral (N) and de-protonated (D) state (Δ_T). So the Henderson–Hasselbach type equation can then be written as pH = log [($\Delta_T - \Delta_D$)/Δ_D] + pK_a. The pK_a is calculated from the linear regression analysis of the Hill plot.

Acknowledgements

Generous financial support from the Swedish Natural Science Research Council (Vetenskapsrådet), the Swedish Foundation for Strategic Research (Stiftelsen för Strategisk Forskning) and the EU-FP6 funded RIGHT project (Project no. LSHB-CT-2004-005276) is gratefully acknowledged.

References

1 M. Nowotny, S. A. Gaidamakov, R. J. Crouch and W. Yang, *Cell*, 2005, **121**, 1005.
2 E. Zamaratski, D. Ossipov, P. I. Pradeepkumar, N. Amirkhanov and J. Chattopadhyaya, *Tetrahedron*, 2001, **57**, 593.
3 J. Kurreck, Antisense technologies, *Eur. J. Biochem.*, 2003, **270**, 1628.
4 D. Kirn, *Oncogene*, 2000, **19**, 6660.
5 G. G. Carmichael, *Nat. Biotechnol.*, 2003, **21**, 371.

6 I. Lebedeva and C. A. Stein, *Annu. Rev. Pharmacol. Toxicol.*, 2001, **41**, 403.
7 C. H. Thomas, J. H. Collier, C. S. Sfeir and K. E. Healy, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 1972.
8 R. A. Stull and F. C. Szoka, *J. Pharm. Res.*, 1995, **12**, 465.
9 W. J. Dower and L. C. Mattheakis, *Curr. Opin. Chem. Biol.*, 2002, **6**, 390.
10 W. G. Scott, *Curr. Opin. Struct. Biol.*, 1998, **8**, 720.
11 D. M. Lilley, *ChemBioChem*, 2001, **2**, 729.
12 D. M. Lilley, *ChemBioChem*, 2001, **2**, 31.
13 D. M. Lilley, *Trends Biochem. Sci.*, 2003, **28**, 495.
14 M. Levy and A. D. Ellington, *Bioorg. Med. Chem.*, 2001, **9**, 2581.
15 G. Ramaswamy and F. J. Slack, *Chem. Biol.*, 2002, **9**, 1053.
16 B. Bartel, *Nat. Struct. Mol. Biol.*, 2005, **12**, 569.
17 P. Y. Lu, F. Y. Xie and M. C. Woodle, *Trends Mol. Med.*, 2005, **11**, 104.
18 S. Acharya, P. Acharya, A. Foldesi and J. Chattopadhyaya, *J. Am. Chem. Soc.*, 2002, **124**, 13722.
19 S. Acharya, J. Barman, P. Cheruku, S. Chatterjee, P. Acharya, J. Isaksson and J. Chattopadhyaya, *J. Am. Chem. Soc.*, 2004, **126**, 8674.
20 P. Acharya, P. Cheruku, S. Chatterjee, S. Acharya and J. Chattopadhyaya, *J. Am. Chem. Soc.*, 2004, **126**, 2862.
21 J. L. Hougland, S. K. Deb, D. Maric and J. A. Piccirilli, *J. Am. Chem. Soc.*, 2004, **126**, 13578.
22 P. I. Pradeepkumar, P. Cheruku, O. Plashkevych, P. Acharya, S. Gohil and J. Chattopadhyaya, *J. Am. Chem. Soc.*, 2004, **126**, 11484.
23 C. Thibaudeau, P. Acharya and J. Chattopadhyaya, *Stereoelectronic Effects in Nucleosides and Nucleotides and their Structural Implications*, 2nd ed., Uppsala University Press, Uppsala, 2005.
24 W Saenger, *Principles of Nucleic Acid Structure*, Springer-Verlag, New York, 1983.
25 I. Luyten, C. Thibaudeau and J. Chattopadhyaya, *J. Org. Chem.*, 1997, **62**, 8800.
26 D. Honcharenko, O. P. Varghese, O. Plashkevych, J. Barman and J. Chattopadhyaya, *J. Org. Chem.*, 2006, **71**, 299–314.
27 S Chatterjee, W Pathmasiri and J. Chattopadhyaya, *Org. Biomol. Chem.*, 2005, **3**, 3911.
28 M. D. Liptak and G. C. Shield, *J. Am. Chem. Soc.*, 2001, **123**, 7314.
29 G. J. Tawa, I. A. Topol, S. K. Burt, R. A. Caldwell and A. A. Rashin, *J. Chem. Phys.*, 1998, **109**, 4852.
30 I. A. Topol, G. J. Tawa, S. K. Burt and A. A. Rashin, *J. Chem. Phys.*, 1999, **111**, 10998.
31 V. C. M. Barone, *J. Phys. Chem. A*, 1998, **102**, 1995.
32 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle and J. A. Pople, *GAUSSIAN 98 (Revision A.6)*, Gaussian, Inc., Pittsburgh, PA, 1998.
33 J. J. F. R. A. Klicic, S.-Y. Liu and W. C. Guida, *J. Phys. Chem. A*, 2002, **106**, 1327.
34 K. Fukui, T. Yonezawa and C. Nataga, *Bull. Jpn. Chem. Soc.*, 1954, **27**, 423.
35 K. Fukui, T. Yonezawa and C. Nataga, *J. Chem. Phys.*, 1957, **26**, 831.
36 K. Fukui, T. Yonezawa, C. Nataga and H. Shingu, *J. Chem. Phys.*, 1954, **22**, 1433.
37 H. Ford, Jr., F. Dai, L. Mu, M. A. Siddiqui, M. C. Nicklaus, L. Anderson, V. E. Marquez and J. J. Barchi, Jr., *Biochemistry*, 2000, **39**, 2581.
38 S. M. Freier and K.-H. Altmann, *Nucleic Acids Res.*, 1997, **25**(22), 4429–4443.
39 T. H. Keller and R. Häner, *Nucleic Acids Res.*, 1993, **21**(19), 4499–4505.
40 J. Wengel, *Acc. Chem. Res.*, 1999, **32**, 301–310.
41 D. L. Bergman, L. Laaksonen and A. Laaksonen, *J. Mol. Graph. Modell.*, 1997, **15**, 301–306.
42 L. Laaksonen, *J. Mol. Graph.*, 1992, **10**, 33–34.