Measurement of Nucleobase pK_a Values in Model Mononucleotides Shows RNA−RNA Duplexes To Be More Stable than DNA−DNA Duplexes

P. Acharya, P. Cheruku, S. Chatterjee, S. Acharya, and J. Chattopadhyaya*

Contribution from the Department of Bioorganic Chemistry, Box 581, Biomedical Center, Uppsala University, S-751 23 Uppsala, Sweden

Received September 22, 2003; E-mail: jyoti@boc.uu.se

Abstract: To understand why the RNA−RNA duplexes in general has a higher thermodynamic stability over the corresponding DNA−DNA duplexes, we have measured the pK_a values of both nucleoside 3′,5′-bis-ethyl phosphates [Etp(d/rN)pEt] and nucleoside 3′-ethyl phosphates [(d/rN)pEt] (N = A, G, C, or T/U), modeling as donors and acceptors of base pairs in duplexes. While the 3′,5′-bis-phosphates, Etp(d/rN)-pEt, mimic the internucleotidic monomeric units of DNA and RNA, in which the stacking contribution is completely absent, the 3′-ethyl phosphates, (d/rN)pEt, mimic the nucleotide at the 5′-end. The pK_a values of the nucleobase in each of these model nucleoside phosphates have been determined with low pK_a error ($\sigma = \pm 0.01$ to 0.02) by ^{1}H NMR (at 500 MHz) with 20−33 different pH measurements for each compound. This study has led us to show the following: (1) All monomeric DNA nucleobases are more basic than the corresponding RNA nucleobases in their respective Etp(d/rN)pEt and (d/rN)pEt. (2) The pK_a values of the monomeric nucleotide blocks as well as ΔpK_a values between the donor and acceptor can be used to understand the relative base-pairing strength in the oligomeric duplexes in the RNA and DNA series. (3) The $\Delta G^{\circ}_{pK_a}$ of the donor and acceptor of the base pair in duplexes enables a qualitative dissection of the relative strength of the base-pairing and stacking in the RNA−RNA over the DNA−DNA duplexes. (4) It is also found that the relative contribution of base-pairing strength and nucleobase stacking in RNA−RNA over DNA−DNA is mutually compensating as the % A−T/U content increases or decreases. This interdependency of stacking and hydrogen bonding can be potentially important in the molecular design of the base-pair mimicks to expand the alphabet of the genetic code.

Introduction

The pK_a values of the nucleobase[1a,b] in guanosine 5′-phosphate and adenosine 5′-phosphate and in their 2′-deoxy counterparts have earlier been determined[2] by spectrophotometric titration, and they were found to be the same (9.33 and 3.79, respectively) in both the ribo and 2′-deoxy series. It was also foun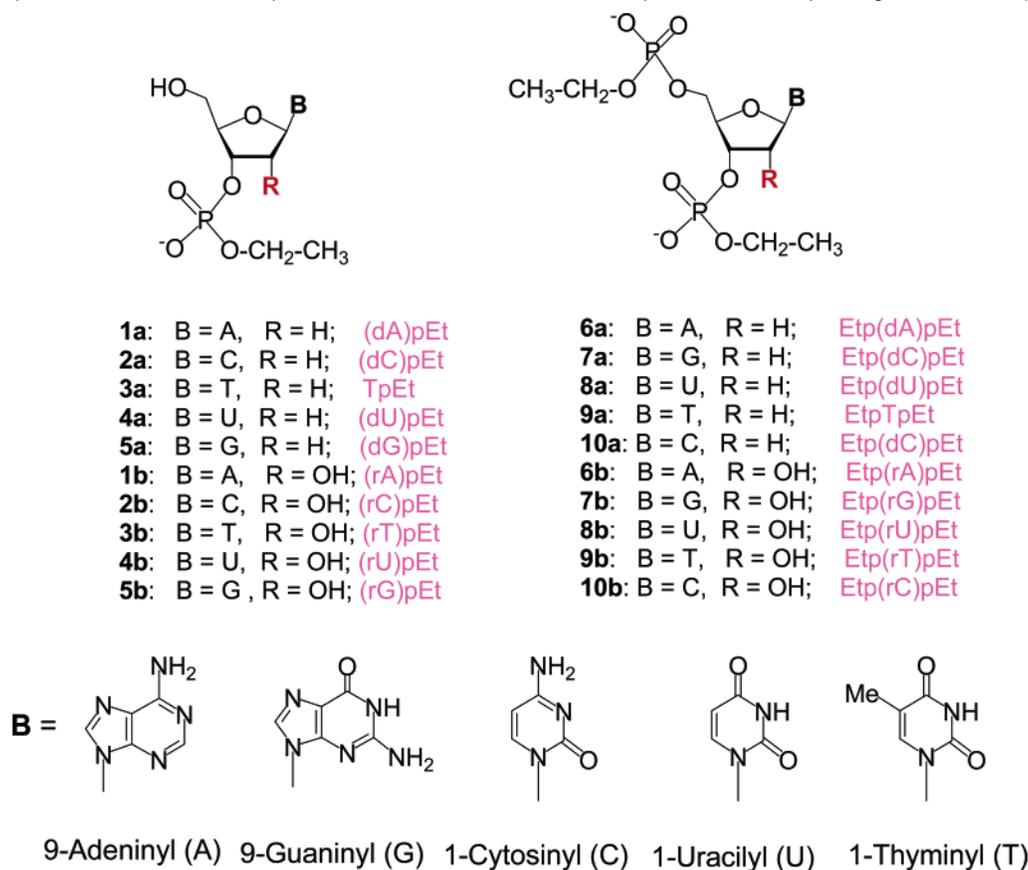d[2a] that because of the electrostatic interaction between the negatively charged 5′-phosphate and the nucleobase, the pK_a value of the nucleobases in nucleoside 5′-phosphate increases by ca. 0.5 pK_a unit compared to that of nucleoside 3′-phosphate. Clearly, pH titration by high-field NMR is a more suitable method for determination of more accurate pK_a[1d−f] of nucleobases using appropriate monomeric model building blocks of DNA and RNA (Scheme 1). We here report the pK_a values (Table 1 and Figure 1) of both nucleosides 3′-ethyl phosphates [(d/rN)pEt] and nucleosides 3′,5′-bis-ethyl phosphates [Etp(d/rN)pEt] in both 2′-deoxy (dN) and ribo (rN) series (Scheme 1), mimicking the monomeric components of DNA and RNA (see the Experimental Section for details) under uniform NMR conditions (1 mM, in the absence of any added mono- or divalent salt) at 500 MHz with high accuracy ($\sigma = \pm 0.01$ to 0.02). This shows that we indeed observe the effect of 2′-deoxy versus 2′-OH in the pentose sugar moiety on the pK_a value of the nucleobases, which we have utilized here to estimate the relative strength of the base-pairing and the stacking in RNA−RNA versus DNA−DNA duplex.

Result and Discussion

A direct comparison of the pK_a values (Table 1 and Figure 1) for all four deoxy (**6a**−**10a**) and ribo (**6b**−**10b**) pairs of nucleoside 3′,5′-bis-ethyl phosphates [Etp(d/rN)pEt; see Scheme

(1) (a) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: Berlin, 1988. (b) Bloomfield, V. A.; Crothers, D. M.; Tinoco, I. *Nucleic Acids: Structures, Properties and Functions*; University Science Books: Sausalito, CA, 1999. (c) Apart from hydrogen bonding and stacking, the other two forces that have been implicated in the overall helix stabilization are phosphate repulsions and conformational entropy; see ref 1b. We have, however, recently found that attractive electrostatic interactions among the stacked nearest-neighbors (see refs 1d−f) play a dominant role in the self-organization of the single-stranded RNA and DNA, which we believe will even be stronger in the duplexes because of the gain in enthalpy due to hydrogen bonding. (d) Acharya, S.; Acharya, P.; Földesi, A.; Chattopadhyaya, J. *J. Am. Chem. Soc.* **2002**, *124*, 13722. (e) Acharya, P.; Acharya, S.; Földesi, A.; Chattopadhyaya, J. *J. Am. Chem. Soc.* **2003**, *125*, 2094. (f) Acharya, P.; Acharya, S.; Cheruku, P.; Amirkhanov, N. V.; Földesi, A.; Chattopadhyaya, J. *J. Am. Chem. Soc.* **2003**, *125*, 9948. (g) Velikyan, I.; Acharya, S.; Trifonova, A.; Földesi, A.; Chattopadhyaya, J. *J. Am. Chem. Soc.* **2001**, *123*, 2893. (h) Acharya, S.; Acharya, P.; Chatterjee, S.; Chattopadhyaya, J., unpublished results.
(2) (a) Clauwaert, J.; Stockx, J. *Z. Naturforsch. B.* **1968**, *23*, 25. (b) Fasman, G. D., Ed. *Nucleic Acids*; Handbook of Biochemistry and Molecular Biology, Vol. 1; Chemical Rubber Co.: Cleveland, OH, 1975; pp 76−206. (c) Sober, H. A.; Harte, R. A.; Sober, E. K. *Handbook of Biochemistry. Selected Data for Molecular Biology*; Chemical Rubber Co.: Cleveland, OH, 1970; pp G3−G98.

Scheme 1. Compounds Used in the PD-Dependent ^{1}H NMR Titration To Give the pK$_a$ of the Corresponding Nucleobases (See Table 1)^a

1a: B = A, R = H; (dA)pEt	**6a**: B = A, R = H; Etp(dA)pEt
2a: B = C, R = H; (dC)pEt	**7a**: B = G, R = H; Etp(dC)pEt
3a: B = T, R = H; TpEt	**8a**: B = U, R = H; Etp(dU)pEt
4a: B = U, R = H; (dU)pEt	**9a**: B = T, R = H; EtpTpEt
5a: B = G, R = H; (dG)pEt	**10a**: B = C, R = H; Etp(dC)pEt
1b: B = A, R = OH; (rA)pEt	**6b**: B = A, R = OH; Etp(rA)pEt
2b: B = C, R = OH; (rC)pEt	**7b**: B = G, R = OH; Etp(rG)pEt
3b: B = T, R = OH; (rT)pEt	**8b**: B = U, R = OH; Etp(rU)pEt
4b: B = U, R = OH; (rU)pEt	**9b**: B = T, R = OH; Etp(rT)pEt
5b: B = G , R = OH; (rG)pEt	**10b**: B = C, R = OH; Etp(rC)pEt

B =

9-Adeninyl (A)   9-Guaninyl (G)   1-Cytosinyl (C)   1-Uracilyl (U)   1-Thyminyl (T)

^{a} All abbreviations for compounds **1a**−**10b** are shown in pink (r = ribo and d = 2′-deoxyribo).

Table 1. pK$_a$ and $\Delta G°_{pKa}$ (in kJ mol^{-1}) of Both Nucleoside 3′-Ethyl phosphate [(d/rN)PEt] (**1a**−**5a** and **1b**−**5b**; See Scheme 1) as Well as Nucleoside 3′,5′-Bisethyl phosphate [Etp(d/rN)PEt] (**6a**−**10a** and **6b**−**10b**; See Scheme 1) in Both 2′-Deoxy (dN) and Ribo (rN) Series Calculated from ^{1}H NMR Titration and Hill Plot Analysis (See Experimental Section and Supporting Information)

dNpEt	pK$_a$^{a} (nucleobase)	$\Delta G°_{pKa}$	rNpEt	pK$_a$^{a} (nucleobase)	$\Delta G°_{pKa}$
(dA)pEt (**1a**)	H8A: 3.35	19.1	(rA)pEt (**1b**)	H8A: 3.11	17.7
	H2A: 3.35	19.1		H2A: 3.10	17.7
(dC)pEt (**2a**)	H5C: 4.12	23.5	(rC)pEt (**2b**)	H5C: 3.84	21.8
	H6C: 4.11	23.5		H6C: 3.84	21.8
TpEt (**3a**)	H6T: 9.94	56.7	(rT)pEt (**3b**)	H6T: 9.65	55.0
	CH$_3$T: 9.91	56.5		CH$_3$T: 9.67	55.1
(dU)pEt (**4a**)	H5U: 9.35	53.3	(rU)pEt (**4b**)	H5U: 9.22	52.6
	H6U: 9.35	53.3		H6U: 9.20	52.5
(dG)pEt (**5a**)	H8G: 9.40	53.6	(rG)pEt (**5b**)	H8G: 9.27	52.9
Etp(dA)pEt (**6a**)	H8A: 3.82	21.8	Etp(rA)pEt (**6b**)	H8A: 3.71	21.2
	H2A: 3.83	21.8		H2A: 3.74	21.3
Etp(dG)pEt (**7a**)	H8G: 9.59	54.7	Etp(rG)pEt (**7b**)	H8G: 9.29	53.0
Etp(dU)pEt (**8a**)	H5U: 9.58	54.6	Etp(rU)pEt (**8b**)	H5U: 9.25	52.8
	H6U: 9.59	54.7		H6U: 9.27	52.9
EtpTpEt (**9a**)	H6T: 10.12	57.7	Etp(rT)pEt (**9b**)	H6T: 9.77	55.7
	CH$_3$T: 10.12	57.7		CH$_3$T: 9.78	55.8
Etp(dC)pEt (**10a**)	H5C: 4.34	24.8	Etp(rC)pEt (**10b**)	H5C: 4.24	24.0
	H6C: 4.35	24.8		H6C: 4.25	24.2

^{a} pK$_a$ was obtained from the specific marker proton shown on the left in the column. Error for pK$_a$: $\sigma = \pm 0.01$ to 0.02. Error for $\Delta G°_{pKa}$: ± 0.1 (see refs 1d−f and the Experimental Section as well as Figure 1 and Figure S3 in the Supporting Information).

1] and four deoxy (**1a**−**5a**) and ribo (**1b**−**5b**) pairs of nucleoside 3′-ethyl phosphates [(d/rN)pEt; see Scheme 1 for numbering and abbreviations] leads us to unequivocally establish that monomeric RNA nucleobases are indeed more acidic than the corresponding DNA nucleobases because the former nucleobases experience the electron-withdrawing effect of its 2′-hydroxyl group. It is likely that this pK$_a$ change of the aglycone results from the change of the inductive effect of the sugar substituents at C2′. This is consistent with two sets of observa-

tions: (i) A pairwise comparison of the chemical shifts of the aromatic protons of each aglycone of the mono-2′-deoxyribo-nucletides (**1a**−**5a**) shows that they are more shielded (see Table S3 in the Supporting Information) than the monoribonucleotides (**1b**−**5b**) because of the inductive effect of the 2′-OH in the latter: 9-adeninyl in 2′-deoxyribo-**1a** (δ_{H8} 8.34, δ_{H2} 8.26) and ribo-**1b** (δ_{H8} 8.35, δ_{H2} 8.27); 9-guaninyl in 2′-deoxyribo-**5a** (δ_{H8} 7.998) and ribo-**5b** (δ_{H8} 8.011); 1-cytosinyl in 2′-deoxyribo-**2a** (δ_{H5} 6.062, δ_{H6} 7.836) and ribo-**2b** (δ_{H5} 6.069, δ_{H6} 7.848). (ii)
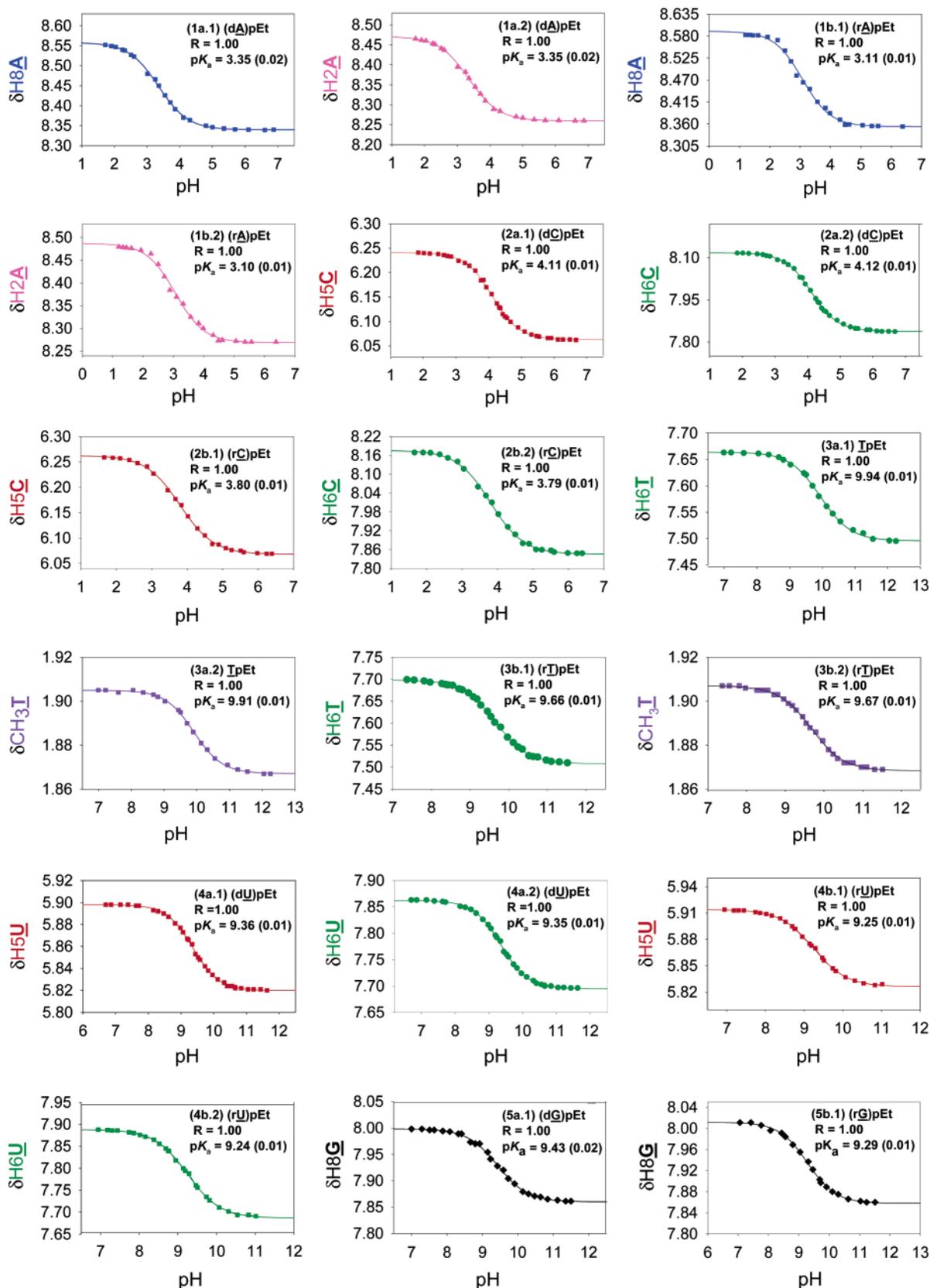
Figure 1. (Continued on next page)

The change of the electronic nature of the aglycon in ribonucleosides furthermore alters the respective pK_a values of the 2′-hydroxyl group. Thus, the pK_a value of 2′-OH is 12.10 ± 0.02 in adenosine,[1g] 12.26 ± 0.04 in 3-deazaadenosine,[1h] and 12.94 ± 0.03 for abasic 1-deoxy-D-ribofuranose.[1g] (iii) It is also noteworthy that the change of the electronic character of the 2′-substituent of the furanose moiety in ribonucleosides also alters the pK_a value of respective aglycone: pK_a of 9-adeninyl
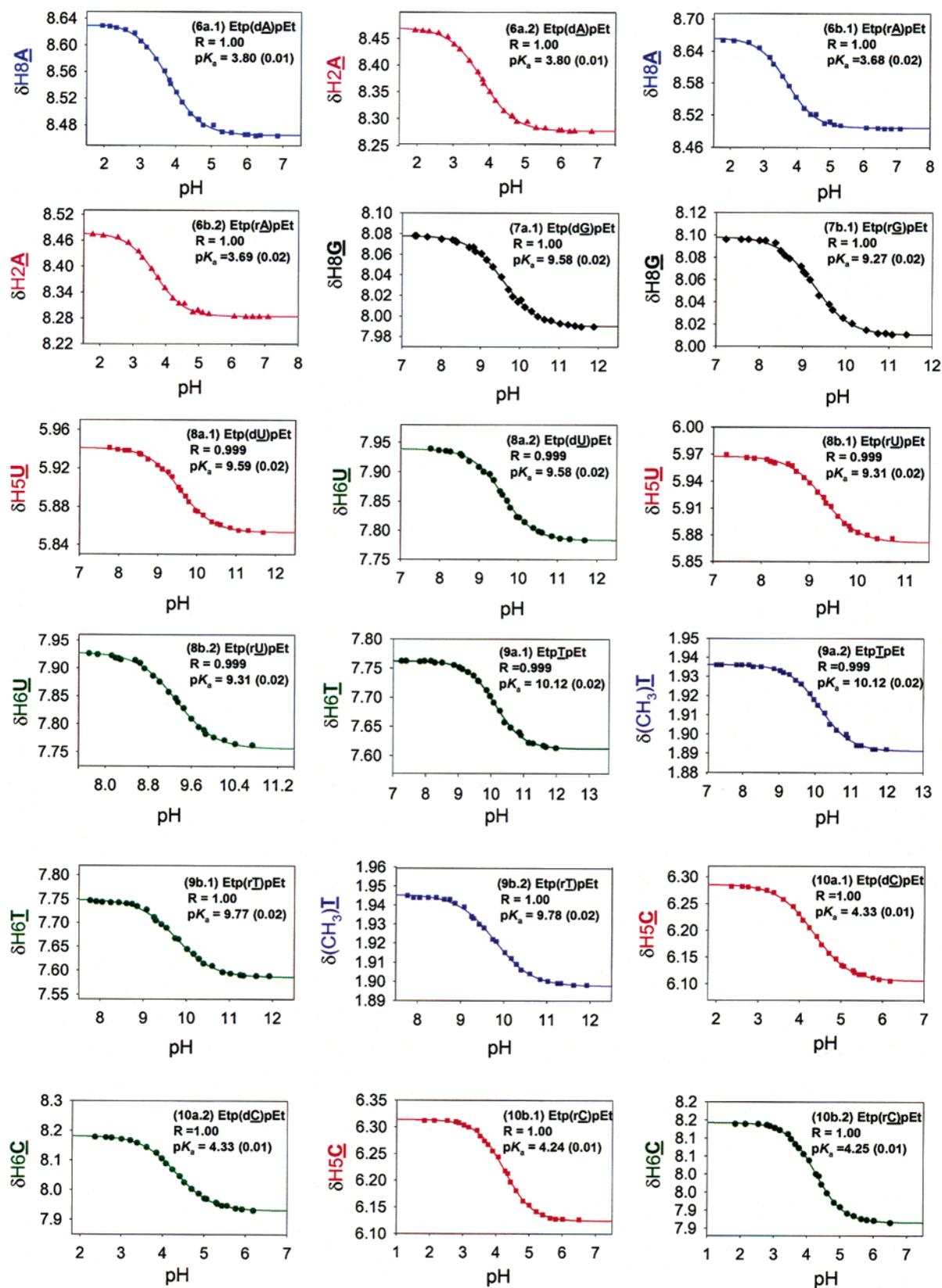
Figure 1. Panels 1a.1−10b.2 show the plot of the pH-dependent (1.6 ≤ pH ≤ 12.24) ^{1}H chemical shifts (δH) for different aromatic marker protons for compounds **1a−10a** and **1b−10b** (Scheme 1) at 298 K, showing the pK_a at the inflection point. ^{1}H NMR chemical shift variations have been measured at 20−33 different pH values in an interval of 0.2−0.3 pH units to obtain the sigmoidal curves with low pK_a error ($\sigma = \pm0.01$ to 0.02; see Experimental Section). Each graph shows chemical shift change with pH for one particular aromatic proton in a compound. The name of the compound with the corresponding panel number along with the particular aromatic proton chosen for titration, the correlation coefficient (R)[10] obtained from curve fitting, and the pK_a values obtained from the subsequent Hill plot analyses (see Figure S3 in the Supporting Information) are shown in the respective graphs [see Experimental Section for details].

Table 2. Results from Model Monomeric Donors and Acceptors

(A) ΔpK_a and $\Delta\Delta G^\circ_{pKa \text{ of each mid bp}}$ (kJ mol^{-1})^a
Values from the Model Monomeric Donors and Acceptors
Representing the H-Bonding Contribution of the Middle
Base Pairs (mid bp, i.e., Excluding the Terminal Base Pairs)^b

monomeric bp	r(C–G)	d(C–G)	r(A–U)	d(A–T)
ΔpK_a ^{c}	5.04	5.24	5.53	6.29
$\Delta\Delta G^\circ_{pKa \text{ of each mid bp}}$ ^{a}	28.9	29.9	31.6	35.9
$[\Delta\Delta pK_a]_{(deoxy)-(ribo)}$ ^{c}		0.20		0.76
$\Delta\Delta G^\circ_{pKa \text{ of each mid bp}}$ ^{a}		−1.0		−4.3

(B) Calculation of the $\Delta\Delta\Delta G^\circ_{pKa \text{ of each terminal bp}}$
(kJ mol^{-1})^d Values from the Model Monomeric
Model Donors and Acceptors To Represent the H-Bonding
Contribution of the Base-Pairing of Terminal Base Pairs^e

terminal bp	C3′/G5′	G3′/C5′	5′G/3′C	5′C/3′G	A3′/T/U5′	T/U3′/A5′	5′U/T/3′A	5′A/3′U/T
$\Delta\Delta\Delta G^\circ_{pKa \text{ of each terminal bp}}$	0.0	0.0	0.0	0.0	−3.5	−3.4	−3.5	−3.4

^{a} See ref 7 for calculations of the free energies of the of each of the middle bp. ^{b} For the middle bp residues in duplex, we have used the ΔG°_{pKa} values (Table 1) of the Etp(d/rN)pEt (**6a−10** and **6b−10a**). The net stabilization of a single r(G−C) over the d(G−C) or r(A−U) over d(A−T) $[\Delta\Delta\Delta G^\circ_{pKa \text{ of each mid bp}}$, in kJ mol$^{-1}]$ has been calculated from the subtraction of $\Delta\Delta G^\circ_{pKa-r(GC) \text{ or } r(AU)}$ from that of $\Delta\Delta G^\circ_{pKa-d(GC) \text{ or } d(AT)}$. ^{c} ΔpK_a for r(G−C): $[pK_a]_{7b} - [pK_a]_{10b}$. For d(G−C): $[pK_a]_{7a} - [pK_a]_{10a}$. For r(U−A): $[pK_a]_{8b} - [pK_a]_{6b}$. For d(T−A): $[pK_a]_{9a} - [pK_a]_{6a}$. $[\Delta\Delta pK_a]_{(deoxy)-(ribo)} = [\Delta pK_a]_{deoxy} - [\Delta pK_a]_{ribo}$. ^{d} See ref 8 for calculations of the free energies of the terminal base pair. ^{e} For the 3′- or 5′-terminal base-pairing in duplex, we have used the ΔG°_{pKa} values of the (d/rN)pEt (**1a−5a** and **1b−5a**) and Etp(d/rN)pEt (**6a−10** and **6b−10a**) (see also Tables S3 and S4 in the Supporting Information for details). We have used Etp(d/rN)pEt (**6a−10** and **6b−10a**) as the model for the 5′-phosphate because the 3′-phosphoryl group in the bis-phosphate does not have any influence on the pK_a of the constituent nucleobase.

is 3.26 ± 0.01 in adenosine, 3.45 ± 0.02 in 2′-*O*-methoxy adenosine, and 3.59 ± 0.01 in 2′-deoxyadenosine.[1h]

The difference in the pK_a modulation found for the respective aglycone in 2′-deoxynucleotides (as in **1a−10a**) by the 2′-H vis-à-vis 2′-OH in the corresponding ribonucleotides (as in **1b−10b**) suggests that an appropriate C2′ substitution with a group of defined electronegativity should also be able to change the nucleobase pK_a accordingly. The potential application of such pK_a engineering is that it can facilitate the general acid−base catalysis in a predefined site of an RNA catalyst at the physiological pH. Alternatively, the change of the electronic nature of the aglycone at a specific center can be used to lower the pK_a value of the constituent 2′-OH group within a specific

(3) (a) Larson, J. W.; McMahon, T. B. *J. Am. Chem. Soc.* **1983**, *105*, 2944. (b) Chen, J.; McAllister, M. A.; Lee, J. K. Houk, K. N. *J. Org. Chem.* **1998**, *63*, 4611 and references therein. (c) Shan, S.-O.; Herschlag, D. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14474. (d) Chen, D. L.; McLaughlin, L. W. *J. Org. Chem.* **2000**, *65*, 7468.

(4) Lesnik, E. A.; Freier, S. M. *Biochemistry* **1995**, *34*, 10807.

(5) (a) Umezawa, Y.; Nishio, M. *Nucleic Acids Res.* **2002**, *30*, 2183 and references therein. (b) Koo, H.-S.; Wu, H.-M.; Crothers, D. M. *Nature* **1986**, *320*, 501. (c) Nelson, H. C. M.; Finch, J. T.; Luisi, B. F.; Klug, A. *Nature* **1987**, *330*, 221. (d) Crothers, D. M.; Shakked, Z. In *Oxford Handbook of Nucleic Acid Structure*; Neidle, S., Ed.; Oxford University Press: Oxford, 1998; pp 455−470. (e) Haran, T. E.; Kahn, J. D.; Crothers, D. M. *J. Mol. Biol.* **1994**, *244*, 135. (f) Koskov, K. M.; Gorin, A. A.; Lu, X.-J.; Olson, W. K. *J. Am. Chem. Soc.* **2002**, *124*, 4838.

(6) (a) The equation $\Delta G^\circ_{Ka} = 2.303(RT)pK_a$ has been used (refs 6b,c) to estimate the free energy of protonation for all compounds (see Table 1). (b) Perrin, D. D.; Dempsey, B.; Serjeant, E. P. *pKa Prediction for Organic Acids and Bases*; Chapman and Hall: New York, 1981. (c) Sharp, K. A.; Honig, B. *Annu. Rev. Biophys. Chem.* **1990**, *19*, 301.

(7) For the calculation of free energy (in kJ mol^{-1}) of base-pairing between middle nucleotidyl residues. For r(G−C) base-pair (bp): $\Delta\Delta G^\circ_{pKa \text{ r(G−C)}} = [\Delta G^\circ_{pKa(7b)}] - \Delta G^\circ_{pKa(10b)}]$. For r(U−A) bp: $\Delta\Delta G^\circ_{pKa \text{ r(U−A)}} = [\Delta G^\circ_{pKa(8b)} - \Delta G^\circ_{pKa(6b)}]$. For d(G−C) bp: $\Delta\Delta G^\circ_{pKa \text{ d(G−C)}} = [\Delta G^\circ_{pKa(7a)} - \Delta G^\circ_{pKa(10a)}]$. For d(A−T) bp: $\Delta\Delta G^\circ_{pKa \text{ d(T−A)}} = [\Delta G^\circ_{pKa(9a)} - \Delta G^\circ_{pKa(6a)}]$. Thus, for G−C bp: $\Delta\Delta G^\circ_{pKa \text{ mid(G−C) bp}} = \Delta\Delta G^\circ_{pKa \text{ r(G−C)}} - \Delta\Delta G^\circ_{pKa \text{ d(G−C)}}$. For T/U−A bp: $\Delta\Delta G^\circ_{pKa \text{ mid(U/T−C) bp}} = \Delta\Delta G^\circ_{pKa \text{ r(U−A)}} - \Delta\Delta G^\circ_{pKa \text{ d(T−A)}}$.

oligorbonucleotide, which can then steer the transesterification reaction at the vicinal phosphate center in a convenient manner to cause the RNA cleavage at a defined center.

(A) Strength of the Base-Pairing Based on the pK_a Difference (ΔpK_a) of the Donor and Acceptor. The strength of a hydrogen bond (A−H···B) between a donor (A) and acceptor (B) has earlier been assessed on the basis of pK_a difference (ΔpK_a) between the two heteroatoms involved in the hydrogen bond.[3] Thus, a hydrogen bond (A−H···B $\leftrightarrows$ A···H−B) is considered[3a] to be weaker when the proton belonging to donor A is more strongly covalently bound to one of the participating heteroatoms (either to A or B), thereby indicating a much stronger A−H bond than B−H bond or vice versa. On the other hand, stronger hydrogen bonds are those where the donor and acceptor have similar pK_a values ("pK_a match")[3b] and that allow the donor and acceptor to share the proton equally, which has been evidenced from both experimental and theoretical studies in the literature.[3b] Thus, the larger the ΔpK_a is between the donor and acceptor, the weaker is the hydrogen bond and vice versa. Recently, Shan and Herschlag calculated[3c] hydrogen bond energies as a function of ΔpK_a for the homologous series of phenol−phenolate complexes in solution. Linear relationships between hydrogen bond strength (log K^{HB}) and ΔpK_a were observed in both DMSO and aqueous solvent.

(8) For the calculation (see ref 6 for equation used to convert pK_a to ΔG°_{pKa}) of free energy (in kJ mol^{-1}) of base-pairing between terminal bases. (I) $\Delta\Delta G^\circ_{pKa \text{ of terminal ribo } C3'/G5' \text{ bp}} \equiv \Delta\Delta G^\circ_{pKa \text{ of terminal ribo } 5'G/3'C \text{ bp}} \equiv {}^{OH}G_P - {}^{P}C_{OH} = \Delta G^\circ_{pKa(5b)} - \Delta G^\circ_{pKa(10b)} = 52.9−24.1 = 28.8.$ (II) $\Delta\Delta G^\circ_{pKa \text{ of terminal ribo } G3'/C5' \text{ bp}} \equiv \Delta\Delta G^\circ_{pKa \text{ of terminal ribo } 5'C/3'G \text{ bp}} \equiv {}^{P}G_{OH} - {}^{OH}C_P = \Delta G^\circ_{pKa(7b)} - \Delta G^\circ_{pKa(2b)} = 53.0−21.8 = 31.2.$ (III) $\Delta\Delta G^\circ_{pKa \text{ of terminal deoxy } C3'/G5' \text{ bp}} \equiv \Delta\Delta G^\circ_{pKa \text{ of terminal deoxy } 5'G/3'C \text{ bp}} \equiv {}^{OH}dG_P - {}^{P}dC_{OH} = \Delta G^\circ_{pKa(5a)} - \Delta G^\circ_{pKa(10a)} = 53.6−24.8 = 28.8.$ (IV) $\Delta\Delta G^\circ_{pKa \text{ of terminal deoxy } G3'/C5' \text{ bp}} \equiv \Delta\Delta G^\circ_{pKa \text{ of terminal deoxy } 5'C/3'G \text{ bp}} \equiv {}^{P}dG_{OH} - {}^{OH}dC_P = \Delta G^\circ_{pKa(7a)} - \Delta G^\circ_{pKa(2a)} = 54.7−23.5 = 31.2.$ (V) $\Delta\Delta G^\circ_{pKa \text{ of terminal ribo } A3'/U5' \text{ bp}} \equiv \Delta\Delta G^\circ_{pKa \text{ of terminal ribo } 5'U/3'A \text{ bp}} \equiv {}^{OH}U_P - {}^{P}A_{OH} = \Delta G^\circ_{pKa(4b)} - \Delta G^\circ_{pKa(6b)} = 52.6−21.3 = 31.3.$ (VI) $\Delta\Delta G^\circ_{pKa \text{ of terminal ribo } U3'/A5' \text{ bp}} \equiv \Delta\Delta G^\circ_{pKa \text{ of terminal ribo } 5'A/3'U \text{ bp}} \equiv {}^{P}U_{OH} - {}^{OH}A_P = \Delta G^\circ_{pKa(8b)} - \Delta G^\circ_{pKa(1b)} = 52.9−17.7 = 35.2.$ (VII) $\Delta\Delta G^\circ_{pKa \text{ of terminal deoxy } A3'/T5' \text{ bp}} \equiv \Delta\Delta G^\circ_{pKa \text{ of terminal deoxy } 5'T/3'A \text{ bp}} \equiv {}^{OH}T_P - {}^{P}dA_{OH} = \Delta G^\circ_{pKa(4a)} - \Delta G^\circ_{pKa(6a)} = 56.6−21.8 = 34.8.$ (VIII) $\Delta\Delta G^\circ_{pKa \text{ of terminal deoxy } T3'/A5' \text{ bp}} \equiv \Delta\Delta G^\circ_{pKa \text{ of terminal deoxy } 5'A/3'T \text{ bp}} \equiv {}^{P}T_{OH} - {}^{OH}dA_P = \Delta G^\circ_{pKa(8a)} - \Delta G^\circ_{pKa(1a)} = 57.7−19.1 = 38.6.$ So for the terminal base pairing in RR over DD: (I) − (III) = $\Delta\Delta\Delta G^\circ_{pKa \text{ of terminal } C3'/G5' \text{ bp}} \equiv \Delta\Delta\Delta G^\circ_{pKa \text{ of terminal } 5'G/3'C \text{ bp}} = 28.8 − 28.8 = 0.0.$ (II) − (IV) = $\Delta\Delta\Delta G^\circ_{pKa \text{ of terminal } G3'/C5' \text{ bp}} \equiv \Delta\Delta\Delta G^\circ_{pKa \text{ of terminal } 5'C/3'G \text{ bp}} = 31.2 − 31.2 = 0.0.$ (V) − (VII) = $\Delta\Delta\Delta G^\circ_{pKa \text{ of terminal } A3'/T5' \text{ bp}} \equiv \Delta\Delta\Delta G^\circ_{pKa \text{ of terminal } 5'U/3'A \text{ bp}} = 31.3 − 34.8 = −3.5.$ (VI) − (VIII) = $\Delta\Delta\Delta G^\circ_{pKa \text{ of terminal } 5'U/T/3'A \text{ bp}} \equiv \Delta\Delta\Delta G^\circ_{pKa \text{ of terminal } 5'A/3'U/T \text{ bp}} = 35.2 − 34.8 = −3.4.$ The final equation used (see Table S1 in the Supporting Information for the number of middle and terminal base pairs for each duplex) for calculating the net stabilization due to base-pairing (Table 3) of RNA−RNA (RR) over DNA−DNA (DD) is $[\Delta G^\circ_{pKa}]_{RR−DD}$ (in kcal mol^{-1}) = {[$\Delta\Delta G^\circ_{pKa \text{ of each mid bp}}]_{model \text{ ribo}−monomer} - [\Delta\Delta G^\circ_{pKa \text{ of each mid bp}}]_{model \text{ deoxy}−monomer}$ (Table 2A)} × (number of middle bp) + {$\Delta\Delta G^\circ_{pKa \text{ of each terminal bp}}]_{model \text{ ribo}−monomer} - [\Delta\Delta G^\circ_{pKa \text{ of each terminal bp}}]_{model \text{ deoxy}−monomer}$ (Table 2B)} × (number of terminal bp). The relation 1 kcal mol^{-1} = 4.2 kJ mol^{-1} has been used to convert all values in kJ mol^{-1} to kcal mol^{-1}. Similarly, the ΔpK_a values of all middle base-pairing, except for the two terminals, shown by $\Sigma\Delta pK_a$, which is calculated on the basis of ΔpK_a values of the model monomeric donors and acceptors (shown in Scheme 1), represent the number of G−C and A−U base-pairing in RR and the number d(G−C) and d(A−T) base-pairing in DD (see Table 2A) for each oligo duplex (except the terminal basepairs) shown in Figure 2. Thus, $\Sigma\Delta pK_a = [\Delta pK_a]_{for \text{ model A−U/T bp}}(\text{number of A−U/T bp in middle}) + [\Delta pK_a]_{for \text{ model G−C bp}}(\text{number of G−C bp in middle})$. See also Tables S3 and S4 in the Supporting Information.

(9) (a) Maltseva, T.; Chattopadhyaya, J. *Tetrahedron* **1995**, *51*, 5501. (b) Maltseva, T.; Agback, P.; Chattopadhyaya, J. *Nucleic Acids Res.* **1993**, *21*, 4246. (c) Maltseva, T.; Zarytova, V. F.; Chattopadhyaya, J. *J. Biochem. Biophys. Methods* **1995**, *30*, 163. (d) Becker, M.; Lerum, V.; Dickson, S.; Nelson, N. C.; Matsuda, E. *Biochemistry* **1999**, *38*, 5603.
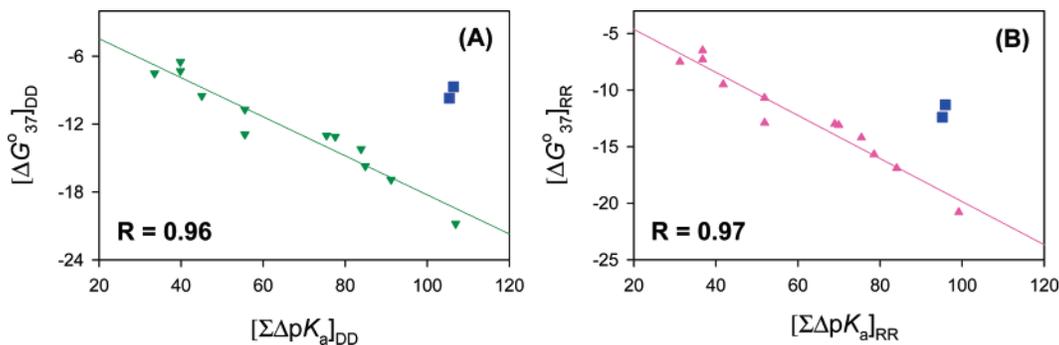
Figure 2. Panels A and B show plot of overall free energy stabilization DNA−DNA (DD) and RNA−RNA (RR) duplexes ([ΔG°_{37}]$_{DD}$ for DD duplexes and [ΔG°_{37}]$_{RR}$ for RR duplexes, in kcal mol^{-1}) as a function of sum of the pK$_a$ differences ($\Sigma\Delta$pK$_a$) between the model monomeric donors and acceptors (shown in Scheme 1), representing G−C and A−U base-pairing in RR duplexes ([$\Sigma\Delta$pK$_a$]$_{RR}$) and d(G−C) and d(A−T) base-pairing in DD duplexes ([$\Sigma\Delta$pK$_a$]$_{DD}$). See Table 2A and refs 7 and 8 for details of calculation. R = linear correlation coefficient.[10] Both [ΔG°_{37}]$_{DD}$ and [ΔG°_{37}]$_{RR}$ showed linear correlations [R = 0.96 in panel A and R = 0.97 in panel B]. See Table 3 and ref 4 for details of [ΔG°_{37}]$_{DD}$ and [ΔG°_{37}]$_{RR}$ for DD and RR duplexes (**1−14**), respectively, used in these linear regression analyses[10] except duplexes **7** and **12**, which (blue ■) did not show the correlation most probably because of the presence of AATA sequence, which has opposing conformational tendencies of adjacent TA and AT steps, thereby showing unusual melting tendencies.

The ΔpK$_a$ values for the G−C base-pairing in the two-model ribo pair (**7b**/**10b**) and deoxy pair (**7a**/**10a**) are 5.04 and 5.24, respectively (Table 2A). Similarly, the ΔpK$_a$ for the A−U/T base-pairing in the two-model ribo pair (**6b**/**8b**) and deoxy pair (**6a**/**8a**) are 5.53 and 6.29, respectively. Comparison of these ΔpK$_a$ values shows that the ribonucleotide base-pairing is stronger than the corresponding 2′-deoxyribonucleotide base-pairing. The fact that the ΔpK$_a$ is considerably less for the G−C base-pairing in both the ribo (ΔΔpK$_a$ = 0.49) and the deoxy (ΔΔpK$_a$ = 1.05) series than those for the A−U/T base-pairing shows that this is consistent, as expected, with the fact that the former is stronger than the latter.

(B) Validity of Use of pK$_a$ Differences (ΔpK$_a$) among the Monomer Blocks To Understand the Base-Pairing Contributions in the Free Energy of Oligo DNA−DNA and RNA−RNA Duplex Stability. Both the stacking and hydrogen bonding are two of the most essential components[1a−c] in the stabilization of the double-stranded DNA or RNA helix, which contribute to the free energy [ΔG°] of the RNA−RNA and DNA−DNA duplex formation. Both the computer simulation and the dangling base studies have earlier shown that the stacking interactions between the neighboring nucleobases stabilize the self-assembly process of double-stranded DNA or RNA helix[1a,b] perhaps more strongly (ca 0.4−3.6 kcal mol^{-1}) than the H-bonding-promoted stabilization (0.5−2 kcal mol^{-1} per H-bond[1]). Although it is well-known that RNA−RNA (RR) duplexes are thermodynamically more stable than DNA−DNA (DD) duplexes and the stability of RR over DD increases as the content of the A−T/U base pair (bp) decreases,[4] it is not clear why the former, in general, is thermodynamically more stable than the latter.

Since we have now independently estimated the hydrogen-bonding strength (ΔpK$_a$) between the model monomeric donors and the acceptors, as in Scheme 1 (in which the stacking is completely absent), representing each ribo and 2′-deoxy G−C and A−T/U base pairs in aqueous solution, we argued that the ΔG°_{37} of the RR and DD duplex formation and the ΔpK$_a$ contribution should give a high degree of correlation. Since the strength of the total hydrogen-bonding contribution, calculated from the sum of ΔpK$_a$ values ($\Sigma\Delta$pK$_a$) of the total number of middle base pairs,[8] is a contributing component to the total stabilization expressed in ΔG°_{37} of the RR and DD duplex formation, we have plotted ΔG°_{37} as a function of $\Sigma\Delta$pK$_a$. It

gives a straight line with high correlation coefficient (R, based on linear regression analysis[10]) for both DNA−DNA (R = 0.96, Figure 2A) and RNA−RNA (R = 0.97, Figure 2B) duplexes. This shows that although the potential for hydrogen bonding between the monomer model systems in the aqueous solution is relatively weak, the relative magnitude of the hydrogen-bonding contribution (without considering stacking interaction), as determined from the simple pK$_a$ determination, to ΔG°_{37} of the duplex melting is well correlated. This is particularly interesting in view of the fact that the ΔG°_{37} of the RR and DD duplexes and $\Sigma\Delta$pK$_a$ originate from two completely independent experiments, the former from the oligomer melting and the latter from the pK$_a$ measurements of the model monomeric compounds. Most importantly, the above correlation suggested that a subtraction of the free energy of the total base-pairing ([ΔG°_{bp}]$_{RR−DD}$; see eq 1), as calculated from the donors and the acceptors in the model monomeric blocks (see refs 7 and 8 as well as Supporting Information for details), from the ΔG°_{37} of helix melting of the RR and DD duplexes may give us a good independent experimental measure of the relative contribution of the free energy of the stacking:

$$[\Delta G^\circ_{bp}]_{RR-DD} \text{ (Table 3)} =$$
$$\{[\Delta\Delta G^\circ_{pKa \text{ of each mid bp}}]_{\text{model ribo-monomer}} -$$
$$[\Delta\Delta G^\circ_{pKa \text{ of each mid bp}}]_{\text{model deoxy-monomer}} \text{ (Table 2A)}\} \times$$
$$(\text{number of middle bp}) +$$
$$\{[\Delta\Delta G^\circ_{pKa \text{ of each terminal bp}}]_{\text{model ribo-monomer}} -$$
$$[\Delta\Delta G^\circ_{pKa \text{ of each terminal bp}}]_{\text{model deoxy-monomer}} \text{ (Table 2B)}\} \times$$
$$(\text{number of terminal bp})... \text{ (1)}$$

(C) Dissection of Contributions from Stacking Vis-à-Vis Base-Pairing in the DNA−DNA and RNA−RNA Duplexes. Thus, the subtraction of ΔG°_{pKa} of appropriate donor and acceptor (Table 1) gave $\Delta\Delta G^\circ_{pKa}$ of A−T/U and G−C base-pairing in both DNA and RNA series in Tables 2 and 3 (see also ref 7 and 8 for details of calculations). A plot of the free energy gain ([ΔG°_{bp}]$_{RR−DD}$) in the base-pairing of RR duplex over the DD duplex as a function of % A−T/U content (△, Figure 3) shows a linear correlation (R = 0.84 by linear

---

(10) *SigmaPlot*, version 8.0: SPSS Science Software GmbH, Schimmel-buschstrasse 25, Potfach 4107, 40688 Erkrath, Germany. For the downloading of the manual, see http://www.spss.com/spssbi/sigmaplot/.

Table 3. Thermodynamic Analyses of the Helix Stability[a] [ΔG°_{37}]$_{RR-DD}$, Free Energy for Base-Pairing[b] [ΔG°_{bp}]$_{RR-DD}$, and Free Energy for Stacking[c] [$\Delta G^{\circ}_{stacking}$]$_{RR-DD}$ in RNA–RNA (RR) over the DNA–DNA (DD) Duplexes (**1**–**14**)[d] and Their A–T/U Vis-à-Vis G–C Base-Pair Content

| sequence | [ΔG°_{37}]$_{RR-DD}$ | | | T/U–A bp | G–C bp | [ΔG°_{bp}]$_{RR-DD}$ | [$\Delta G^{\circ}_{stacking}$]$_{RR-DD}$ |
	RR	DD	RR–DD				
1	−18.6	−10.7	−7.9	4	8	−5.6	−2.3
2	−12.2	−7.3	−4.9	4	5	−4.8	−0.1
3	−11.9	−6.5	−5.4	4	5	−5.8	−0.4
4	−15.7	−9.5	−6.2	3	7	−4.3	−1.9
5	−21.9	−15.6	−6.2	6	11	−8.8	2.6
6	−21.3	−16.9	−4.4	7	11	−9.9	5.5
7	−12.4	−9.7	−2.8	11	4	−12.7	9.9
8	−12.3	−7.5	−4.8	2	6	−3.0	−1.8
9	−18.9	−12.9	−6.0	3	9	−4.7	−1.3
10	−24.1	−20.8	−3.3	7	14	−8.3	5.0
11	−15.1	−13.0	−2.1	9	6	−10.6	8.5
12	−11.3	−8.7	−2.6	12	3	−12.2	9.6
13	−13.8	−13.1	−0.7	9	6	−10.2	9.5
14	−13.1	−14.2	1.1	10	6	−11.2	12.3

[a] The data for helix stability [ΔG°_{37}]$_{RR-DD}$ (in kcal mol^{-1}) of the above-mentioned DNA–DNA (DD) and RNA–RNA (RR) sequences have been performed on the basis of T_m analyses and taken from ref 4. [b] Free-energy for base-pairing of RR over DD [ΔG°_{bp}]$_{RR-DD}$ (in kcal mol^{-1}) have been calculated from pK_a studies as shown in Table 2 and refs 7 and 8. [c] Free-energy for stacking of RR over DD [$\Delta G^{\circ}_{stacking}$]$_{RR-DD}$ (in kcal mol^{-1}) have been calculated from [ΔG°_{37}]$_{RR-DD}$ − [ΔG°_{bp}]$_{RR-DD}$. [d] The duplexes **1**–**14** used in this study are taken from ref 4: (**1**) 5′-TCCCTCCTCTCC-3′/3′-AGGGAGGAGAGG-5′; (**2**) 5′-CCTTCCCTT-3′/3′-GGAAGGGAA-5′; (**3**) 5′-TTCCCTTCC-3′/3′-AAGGGAAGG-5′; (**4**) 5′-GCTCTCTGGC-3′/3′-CGAGAGACCG-5′; (**5**) 5′-CTCGTAC CTTCCGGTCC-3′/3′-GAGCATG GAAGGCCAGG-5′; (**6**) 5′-CTCGTACCTTTCCGGTCC-3′/3′GAGCATG GAAAGGCCAGG-5′; (**7**) 5′-TAGTTATCTCTATCT-3′/3′-ATCAATA-GAGATAGA-5′; (**8**) 5′GCACAGCC-3′/3′-CGTGTCGG-5′; (**9**) 5′-GAGCTC-CCAGGC-3′/3′-CTCGAGGGTCCG-5′; (**10**) 5′-GCCGAGGTCCATG TCG-TACGC-3′/3′-CGGCTCCAGGTACAGCATGCG-5′; (**11**) 5′-TGTAC-GTCACAACTA-3′/3′ACATGCAGTGTTGAT-5′;(**12**)5′-TATACAAGTTATCTA-3′/3′-ATATGTTCAATAGAT-5′; (**13**) 5′CGCCTATGCA AAAAC-3′/3′-GCTGATACGTTT TTG-5′; (**14**) 5′-CGCAAAAAAAAAAACGC-3′/3′G-CGTTTTTTT TTTGCG-5′.
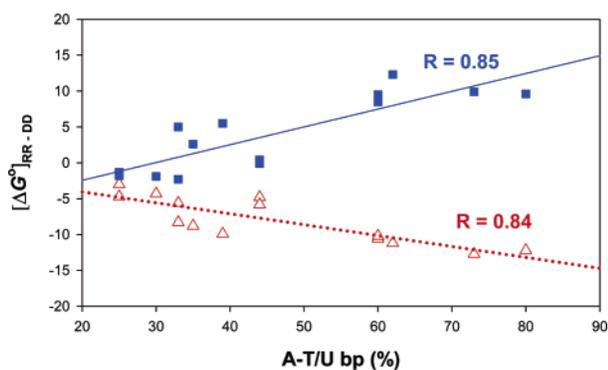

Figure 3. Plot of free energy stabilization of the base-pairing ([ΔG°_{bp}]$_{RR-DD}$) [$\triangle$, $R = 0.84$] as well as that for stacking ([$\Delta G^{\circ}_{stacking}$]$_{RR-DD}$) [$\blacksquare$, $R = 0.85$] in RNA–RNA (RR) over the DNA–DNA (DD) duplex as a function of % A–T/U bp content in the duplexes. R = correlation coefficient. Note that the free energy for stacking of RR over DD [$\Delta G^{\circ}_{stacking}$]$_{RR-DD}$ has been calculated [in $\blacksquare$] from [ΔG°_{37}]$_{RR-DD}$ − [ΔG°_{bp}]$_{RR-DD}$ (see Supporting Information). See Table 3 and refs 7 and 8 for details of calculations. The linear regression analyses[10] for both these plots without duplexes **7** and **12** give $R = 0.76$ for $\triangle$ and $R = 0.84$ for $\blacksquare$.

regression analysis as in ref 10). This means that as the % A–T content increases, the base-pairing contribution for the overall stability of DD duplex over the RR duplex decreases, given all other external factors remaining the same. The reason for this is that as the % A–T base-pair content increases, the 5-methyl group increasingly destabilizes (by [$\Delta\Delta pK_a$]$_{(A.T)-(A.U)}$ = 0.76 pK_a unit, which is equivalent to 4.3 kJ mol^{-1}, Table 2A) the

base-pairing in DD duplex over the % A–U base-pair content in the corresponding RR duplex. For the sake of simplicity,[1c] if one considers that the hydrogen bonding and stacking are the two main components in the stability of a helix, then a subtraction of the free-energy contribution to the base-pairing, [ΔG°_{bp}]$_{RR-DD}$, from the free energy of the helix stability ([ΔG°_{37}]$_{RR-DD}$)[4] should give us an approximate idea of the contribution of stabilization achieved through the stacking (i.e., [ΔG°_{37}]$_{RR-DD}$ − [ΔG°_{bp}]$_{RR-DD}$ = [$\Delta G^{\circ}_{stacking}$]$_{RR-DD}$). Thus, a plot of [$\Delta G^{\circ}_{stacking}$]$_{RR-DD}$ as a function of % A–T/U bp content (■, Figure 3) in the duplexes shows a linear correlation ($R = 0.82$), however, with the reverse slope with respect to that of base-pairing. A comparison of these two linear plots (Figure 3) with opposite slope shows that with the increase of % A–T bp the stability of DNA–DNA duplex weakens over the corresponding RNA–RNA duplexes ([ΔG°_{bp}]$_{RR-DD}$) while the strength of stacking ([$\Delta G^{\circ}_{stacking}$]$_{RR-DD}$) of A–T rich DNA–DNA sequence increases in comparison with the A–U rich sequence in RNA–RNA duplexes. It is likely that this increased stacking contribution from T compared to U, in DNA–DNA over RNA–RNA duplex, comes from the favorable electrostatic CH/π interaction[5a] between the 5-methyl group of T with the nearest-neighbor A in the AT-rich sequence. This is consistent with the recent crystal structure analysis[5a] of various A/T-rich oligo-DNAs, which shows that the structure of 5′-ApTpApT-3′ is stabilized by the favorable interaction of the 5-methyl group of T with the π ring of the 9-adeninyl moiety preceding it in the same strand. This interaction is duplicated in the opposite complementary strand, thereby giving a "twin A/T–Me interaction".[5a] Thus, a successive A/T–Me stacking has been suggested to be responsible for making the A tracts robust and straight.[5a] It is also known that the successive occurrence of the N/T–Me and the A/T–Me motifs is responsible for the deformability of DNA.[5b–f]

Conclusions

(1) The nucleobases of the monomeric DNA are uniformly more basic than the corresponding RNA counterparts.

(2) The strength of the base-pairing based on the pK_a difference (ΔpK_a) of the monomeric donor and acceptor can be used to understand the relative base-pairing strength of larger oligomeric DNA–DNA and RNA–RNA duplexes.

(3) The use of pK_a differences (ΔpK_a) among the monomer blocks, modeling the A–T/U and G–C base pairs, allows us to understand the base-pairing contributions in the free energy of DNA–DNA and RNA–RNA duplex stability, which is evident from a high correlation coefficient based on linear regression ($R = 0.96$ and 0.97, respectively) found between ΔG°_{37} of the helix stability and the sum of the ΔpK_a values ($\sum\Delta pK_a$) of donor/acceptor in the base-pair formation.

(4) The high correlation of ΔG°_{37} of the helix stability and the sum of ΔpK_a values in a duplex ($\sum\Delta pK_a$) showed that a simple subtraction of the base-pairing contribution of RNA–RNA over DNA–DNA ([ΔG°_{bp}]$_{RR-DD}$) from the free energy of the total helix stability ([ΔG°_{37}]$_{RR-DD}$) gives us the relative stacking contribution ([$\Delta G^{\circ}_{stacking}$]$_{RR-DD}$) in a qualitative manner.

(5) A comparison of these two linear plots ([ΔG°_{bp}]$_{RR-DD}$ versus [$\Delta G^{\circ}_{stacking}$]$_{RR-DD}$ as a function of % A–T/U bp content) with opposite slope shows that with the increasing content of

A−T base pairs the stability of DNA−DNA duplex weakens over the corresponding RNA−RNA duplexes ($[\Delta G^\circ_{bp}]_{RR-DD}$), while the strength of stacking ($[\Delta G^\circ_{stacking}]_{RR-DD}$) of A−T rich DNA−DNA sequence increases in comparison with A−U rich sequence in RNA−RNA duplexes. This increased stacking contribution from T compared to U, in DNA−DNA over RNA−RNA duplex, comes from favorable electrostatic CH/π interaction between the 5-methyl group of T with the nearest-neighbor A in the AT rich sequence.

(6) On the basis of NOESY/ROESY experiments,[9a] base-pair exchange kinetics,[9b,c] and analysis of energy of activation[9d] for the base-paired imino proton exchange with the bulk water, it has been shown that the grooves of the fully matched DNA duplexes are relatively less hydrated than those in the mismatched or single-stranded counterparts. That the core of the double helix is dehydrated has also recently been validated by a comparison of the relative hydrolysis rate[9d] of the DNA-tethered acridinium ester in that the acridinium ester hydrolysis rate was relatively slower in the matched duplex (because of poorer water availability) compared to the mismatched and the single-stranded DNA. These studies lead us to suggest that the actual strength of the base-pairing inside the duplexes is stronger than that found from the present consideration of pK_a in the model monomeric donors and acceptors. This means that the relative strength of stacking within the DNA duplex might be more reduced in the matched duplexes by more efficient hydrogen bonding than found in the present work.

Experimental Section

(A) pH-Dependent ¹H NMR Measurement. All NMR experiments were performed using Bruker DRX-500 and DRX-600 spectrometers. The NMR samples for compounds of all four 2′-deoxy (**6a**−**10a**) and ribo (**6b**−**10b**) pairs of nucleosides 3′,5′-bis-ethyl phosphates, Etp(d/rN)pEt, and four 2′-deoxy (**1a**−**5a**) and ribo (**1b**−**5b**) pairs of nucleosides 3′-ethyl phosphates (Scheme 1) were prepared in D$_2$O solution (concentration of 1 mM in order to rule out any chemical shift change owing to self-association) with δ_{DSS} 0.015 ppm as the 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 of 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$_2$O) of pH 4, 7, and 10. The pD of the sample has been adjusted by simple addition of microliter

volumes of NaOD solutions (0.5, 0.1, and 0.01 M). The assignments (see Figure S2 in the Supporting Information) for all compounds have been performed on the basis of selective homonuclear (¹H) and heteronuclear (³¹P) decoupling experiments. All ¹H spectra have been recorded using 128K data points and 64 scans.

(B) pH Titration of Aromatic Protons. The pH titration studies [over the range 1.8 < pH < 12.2, with an interval of pH 0.2−0.3, Figure 1] were carried out for compounds of all four 2′-deoxy (**6a**−**10a**) and ribo (**6b**−**10b**) pairs of nucleosides 3′,5′-bis-ethyl phosphates, Etp(d/rN)pEt, and four 2′-deoxy (**1a**−**5a**) and ribo (**1b**−**5b**) pairs of nucleosides 3′-ethyl phosphates (Scheme 1, Table 1). All pH titration studies consist of ∼20−33 data points (see Figure 1). The corresponding Hill plots for all compounds are given in Figure S3 in the Supporting Information, and the pK_a values shown in Table 1 have been calculated from Hill plot analyses (see section C for details).

(C) pK_a Determination. The pH-dependent [over the range 1.8 < pH < 12.2, with an interval of pH 0.2−0.3] ¹H chemical shifts (δ, with error of ±0.001 ppm) for all compounds (for 2′-deoxy series **1a**−**5a** and **6a**−**10a** as well as ribo series **1b**−**5b** and **6b**−**10b**) show a sigmoidal behavior [Figure 1]. The pK_a determination is based on the Hill plot analysis using the equation pH = log$((1 - \alpha)/\alpha)$ + 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 neutral (N) or the deprotonated (D) states at a given pH ($\Delta_P = \delta_N - \delta_{obs}$ for protonation and $\Delta_D = \delta_D - \delta_{obs}$ for deprotonation, where δ_{obs} is the experimental chemical shift at a particular pH), divided by the total change in chemical shift between the neutral (N) and protonated (P) or deprotonated (D) states (Δ_T). So the Henderson−Hasselbach type equation can then be written as pH = log$((\Delta_T - \Delta_{P/D})/\Delta_{P/D})$ + pK_a. The pK_a is calculated from the linear regression analysis of the Hill plot [Figure S3 in the Supporting Information].

Supporting Information Available: Figures and tables showing ¹H NMR, ³¹P NMR, and Hill plot analysis of **1a**−**10a** and **1b**−**10b** and tables listing the total number of base-pairing and $\sum\Delta$pK_a for **1**−**14**. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0386546