The pK_a's of 2′-Hydroxyl Group in Nucleosides and Nucleotides

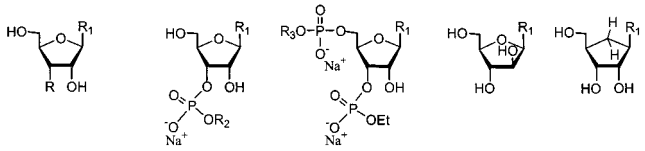
Irina Velikyan, Sandipta Acharya, Anna Trifonova, Andras Földesi, and Jyoti Chattopadhyaya*

Department of Bioorganic Chemistry
Box 581, Biomedical Centre
University of Uppsala, S-751 23 Uppsala, Sweden

Received October 10, 2000

One of the most important features which distinguishes DNA from RNA is the presence of the 2′-hydroxyl group in the latter.[1a,b] The 2′-OH group has major structural implications in that it is involved in recognition, processing, and catalytic properties of RNA,[1c−e] such as the transesterification reactions involved in the Group I and Group II splicing reactions,[1c,e] self-cleavage in lariat-RNA,[1e] RNA catalysis in ribozyme[1d] and in ribonuclease[1d] action. To understand why a specific phosphate function in a large RNA molecule becomes chemically active toward stereospecific transesterification reaction, it is important to address how the reactivity of the 2′-OH group is modulated by the nature of the RNA folding (with or without metal ion as cofactor). This is indeed a very complex issue to tackle because the pK_a of 2′-OH in an RNA molecule is almost impossible to measure because of its inherent lability under alkaline conditions,[1f] hence the experimental pK_a can only be obtained for small model molecules. There are only a few reports thus far on the ionization process of the ribo 2′-OH.[2a−h] The main disadvantage of these studies is that they give inconsistent values for the same compound mainly because of employment of different techniques.[2a−h] For example, the pK_a of 12.35 was obtained for adenosine by thermometric titration,[2a] whereas electrometric titration[2e] reported a value of 12.5. A potentiometric titration of phosphorylated ribose[2d] gives a pK_a of 2′-OH to be 13.9, whereas quantum chemical calculations[2i] produced a value of 14.9. Clearly, the availability of accurate pK_a of 2′-OH for a wide set of ribonucleos(t)ides, measured under uniform experimental conditions, will be of considerable value to model and understand the chemical reactivity of large biologically functional RNA molecules, in general.

We here report the pK_a values of 2′-OH group in ribonucleosides (**3**, **10**, **14**, **16**),[3a] their 3′,5′-bis-alkyl phosphodiester derivatives[3b] (**5**, **11**, **12**, **17**), 3′-monophosphates (**6**, **13**, **15**, **18**),[3a] adenosine 3′-ethyl phosphate (**4**),[3b] 3′-deoxyadenosine (**8**)[3c] as well as those of *ara*-adenosine (**7**),[3a] aristeromycin (**9**),[3a] and the abasic counterparts (**1** and **2**)[3d] by using simple pH-dependent proton chemical shift measurements under *identical conditions* (see Experimental Section in the Supporting Information), showing a pK_a variation of 2′-OH of up to 1.5 units (12.17−13.59) with maximum standard error of ±0.07 pK_a unit (Table 1). The 3′,5′-bis-alkyl phosphodiesters (**5**, **11**, **12**, **17**)[3b] have been used, as a model for the native dinucleoside(3′→5′)phosphate, for pK_a measurement of the 2′-OH group because of their high alkaline stability up to pD 14, mainly owing to poorer leaving-group character of the ethyl group. It has emerged that the pK_a's for 2′-hydroxyl of the ribonucleosides and their phosphate derivatives simply change because of different abilities of various sugar substituents to stabilize the 2′-oxyanion.



1: 1-d-rf (R = OH; R_1 = H) **4**: ApEt (R_1 = A; R_2 = Et) **2**: EtprfpEt (R_1 = H; R_3 = Et) **7**: *ara*-Ado (R_1 = A)
3: Ado (R = OH; R_1 = A) **6**: AMP (R_1 = A; R_2 = H) **5**: EtpApEt (R_1 = A; R_3 = Et) **9**: Aristeromycin (R_1 = A)
8: 3′-dAdo (R = H; R_1 = A) **13**: GMP (R_1 = G; R_2 = H) **11**: MepGpEt (R_1 = G; R_3 = Me)
10: Guo (R = OH; R_1 = G) **15**: UMP (R_1 = U; R_2 = H) **12**: EtpGpEt (R_1 = G; R_3 = Et)
14: Urd (R = OH; R_1 = U) **18**: CMP (R_1 = C; R_2 = H) **17**: MepCpEt (R_1 = C; R_3 = Me)
16: Cyd (R = OH; R_1 = C)

1-d-rf = 1-deoxy-D-ribofuranose; EtprfpEt = 1-deoxy-D-ribofuranose 3,5- ethyl-bisphosphodiester;
A = adenin-9-yl (**3-9**); G = guanin-9-yl (**10-13**);U = uracil-1-yl (**14,15**); C = cytosine-1-yl (**16-18**)

The pH-dependent proton chemical shift (H1′, H2′, and H3′) titration shifts for **1**−**18** were analyzed by nonlinear least-squares curve fitting[4a] to the Henderson−Hasselbalch equation: [eq 1, footnote of Table 1], and have simple sigmoidal shapes between the deprotonated and the neutral states (see Supporting Information for experimentals and for pH-dependent chemical shift titration curves; Figure 1S). The ^{13}C NMR experiments conducted for ^{13}C-labeled adenosine, cytidine, and uridine (Figures 3S, 4S in the Supporting Information) have shown that the pK_a values obtained by proton- and ^{13}C chemical shift titrations are the same within the error (Table 1S), which is also evident from the pH-dependent correlation plot of δ^1H1′ versus δ^{13}C1′, -2′, or -3′, giving correlation coefficient of 0.999 (Figure 5S in the Supporting Information). The pK_a values calculated from the eq 2[4b] (Table 1, footnote; Experimental Section and Figure 2S in Supporting Information) nicely match the ones obtained by the curve-fitting procedure[4a] of the experimental pH-dependent chemical shifts with the exception of those for **2**, **13**, **15**, **18** which show incomplete ionization at pH 13.6. The pK_a values of these compounds have been estimated from the Hill slope[4c] values (see Experimental Section in the Supporting Information) to be not less than 13.55, 13.46, 13.55, and 13.59, respectively, for **2**, **13**, **15**, **18**.

In general, any effect within the molecule that results in electron withdrawal from the 2′-oxyanion has a stabilizing effect because any delocalization of the charge reduces the energy of the system, resulting in increased acidity of the 2′-OH group through either H-bonding,[2a] through-space field effect,[5a] through-bond inductive

(1) (a) Saenger, W. In *Principles of Nucleic Acid Structure*; Springer-Verlag: Berlin, 1988. (b) Lehninger, A. L.; Nelson, D. L.; Cox, M. M. *Principles of Biochemistry;* Worth Publishers: New York, 1993. (c) Cech, T. R. *Annu. Rev. Biochem*. **1990**, *59*, 543 and refrrences therein. (d) Fersht, A. R. *Enzyme Structure and Mechanism*; W. H. Freeman & Co: San Francisco, 1985; pp 426−431. (e) Rousse, B.; Puri, N.; Viswanadham, G.; Agback, P.; Glemarec, C.; Sandström, A.; Sund, C.; Chattopadhyaya, J. *Tetrahedron* **1994**, *50*, 1777−1810. (f) Kochetkov, N. K., Budovskii, E. I. *Organic Chemistry of Nucleic Acids*; Plenum Press: London and New York, 1972.

(2) (a) Izatt, R. M.; Hansen, L. D.; Rytting, J. H.; Christensen, J. J. *J. Am. Chem. Soc*. **1965**, *87*, 2760−2761. (b) Izatt, R. M.; Rytting, J. H.; Hansen, L. D.; Christensen, J. J. *J. Am. Chem. Soc*. **1966**, *88*, 2641−2645. (c) Christensen, J. J.; Rytting, J. H.;.Izatt, R. M. *J. Phys. Chem.* **1967**, *71*, 2700−2705. (d) Usher, D. A.; Richardson, D. I., Jr.; Oakenfull, D. G. *J. Am. Chem. Soc*. **1970**, *92*, 4699−4712. (e) Levene, P. A.; Bass L. W.; Simms H. S. *J. Biol. Chem.* **1926,** *70*, 229. (f) Järvinen, P.; Oivanen, M.; Lönnberg, H. *J. Org. Chem*. **1991**, *56*, 5396−5401. (g) Weinstein, L. B.; Earnshaw, D. J.; Cosstick, R.; Cech, T. R. *J. Am. Chem. Soc*. **1996**, *118*, 10341−10350. (h) Li, Yi.; Breaker, R. R. *J. Am. Chem. Soc*. **1999**, *121*, 5364−5372. (i) Lyne, P. D.; Karplus, M. *J. Am. Chem. Soc*. **2000**, *122*, 166−167.

(3) (a) Adenosine (**3**), AMP (**6**), *ara*A (**7**), aristeromycin (**9**), guanosine (**10**), uridine (**14**), UMP (**15**), cytidine (**16**), CMP (**18**) are commercial products of Sigma. They were used as received. (b) Polak, M.; Plavec, J.; Trifonova, A.; Földesi, A.; Chattopadhyaya, J. *J. Chem. Soc., Perkin Trans. 1* **1999**, 2835−2843. (c) Bazin, H.; Chttopadhyaya, J. *Synthesis* **1985**, 1108−1111. (d) Plavec, J.; Tong, W.; Chttopadhyaya, J. *J. Am. Chem. Soc.* **1993**, *115*, 9734.

(4) (a) Press: W. H.; Teukolsry, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed.; Cambridge University Press: New York, 1993. (b) Albert, A.; Serjeant, E. P. *The Determination of Ionization Constants*; Chapman and Hall: London, 1971; pp 44−52. (c) Wyman, J.; Gill, S. J. *Binding and Linkage. Functional Chemistry of Biological Macromolecules*; University Science Books: Mill Valley, CA, 1990.

(5) (a) Reynolds, W. F. *Prog. Phys. Org. Chem*. **1983**, *14*, 165−203 and references therein. (b) Epshtein, L. M. *Russ. Chem. Rev.* **1979**, *48*, 854−867. (c) Thibaudeau, C.; Plavec, J.; Chattopadhyaya, J. *J. Org. Chem*. **1996**, *61*, 266−286.

10.1021/ja0036312 CCC: $20.00
Published on Web 03/01/2001

Table 1. The pK_a Values with Standard Error at 25 °C for Sugar 2′-Hydroxyl Dissociation in Various Nucleosides, Their 3′-Monophosphates, Nucleotides, Their Abasic Counterparts as Well as 3′-Deoxyadenosine, *ara*-Adenosine and Aristeromycin.

	pK_a for 2'-OH from δH1'		pK_a for 2'-OH from δH2'		pK_a for 2'-OH from δH3'		
cmpd	titration curves (eq 1)[b]	calculated, using eq 2	titration curves (eq 1)[b]	calculated, using eq 2	titration curves (eq 1)[b]	calculated, using eq 2	Overall average pK_a
1	13.02 ± 0.03	13.12 ± 0.01	12.94 ± 0.03	13.05 ± 0.02	13.02 ± 0.03	13.12 ± 0.01	13.05 ± 0.03
2[a]	13.55 ± 0.11		13.61 ± 0.12		13.50 ± 0.10		13.55 ± 0.03
3	12.14 ± 0.04	12.15 ± 0.04	12.19 ± 0.05	12.14 ± 0.03	12.21 ± 0.05	12.18 ± 0.03	12.17 ± 0.01
4	12.55 ± 0.01	12.59 ± 0.01	12.88 ± 0.06	12.86 ± 0.02	12.58 ± 0.02	12.61 ± 0.01	12.68 ± 0.06
5	12.63 ± 0.03	12.70 ± 0.02	12.76 ± 0.06	12.72 ± 0.04	12.58 ± 0.02	12.66 ± 0.02	12.68 ± 0.03
6	13.22 ± 0.03	13.50 ± 0.02	13.25 ± 0.04	13.46 ± 0.02	13.25 ± 0.04	13.59 ± 0.01	13.38 ± 0.06
7			12.55 ± 0.07	12.31 ± 0.03	12.41 ± 0.04	12.31 ± 0.02	12.40 ± 0.06
8	12.53 ± 0.03	12.80 ± 0.02	12.66 ± 0.05	12.64 ± 0.04	12.72 ± 0.06	12.90 ± 0.01	12.71 ± 0.05
9	13.01 ± 0.05	12.99 ± 0.01	13.11 ± 0.08	13.05 ± 0.01	13.00 ± 0.04	12.99 ± 0.01	13.03 ± 0.02
10	12.47 ± 0.02	12.54 ± 0.02	12.54 ± 0.02	12.54 ± 0.04	12.54 ± 0.04	12.52 ± 0.01	12.53 ± 0.01
11	12.87 ± 0.03	13.06 ± 0.01	12.85 ± 0.05	12.88 ± 0.04	12.84 ± 0.02	12.93 ± 0.01	12.91 ± 0.03
12	12.85 ± 0.03	12.95 ± 0.05	12.83 ± 0.08		12.84 ± 0.02	13.02 ± 0.01	12.90 ± 0.04
13[a]	13.43 ± 0.06		13.40 ± 0.08		13.55 ± 0.12		13.46 ± 0.05
14	12.52 ± 0.02	12.66 ± 0.01	12.55 ± 0.02	12.66 ± 0.02	12.62 ± 0.01	12.73 ± 0.02	12.62 ± 0.03
15[a]	13.47 ± 0.06		13.62 ± 0.11		13.57 ± 0.09		13.55 ± 0.04
16	12.40 ± 0.02	12.46 ± 0.02	12.49 ± 0.02	12.58 ± 0.03	12.53 ± 0.02	12.63 ± 0.02	12.52 ± 0.03
17	12.95 ± 0.04	13.12 ± 0.01	13.04 ± 0.04	13.25 ± 0.01	12.95 ± 0.04	13.16 ± 0.01	13.08 ± 0.05
18[a]	13.44 ± 0.10		13.67 ± 0.20		13.65 ± 0.18		13.59 ± 0.07

[a] The pK_a values for **2**, **13**, **15**, and **18** were estimated to be not less than 13.55, 13.46, 13.55, 13.59, respectively. [b] $pH = pK_a + \log [A^-]/[AH] = pK_a + (1\text{-}\alpha)/\alpha$ ..... equation 1. [c] $pK_a = pH + \log(\delta_h - \delta_{obs})/(\delta_{obs} - \delta_l)$ (2) ..... equation 2.

effect,[5a] solvation,[5b] or stereoelectronic anomeric and gauche[5c] effects. Reduction of the pK_a in **3** compared to **8** by 0.54 units is a result of H-bonding[2a] between 2′-OH and 3′-OH with the 2′-OH acting as a donor.[6] The comparison of the pK_a's (Table 1) of 2′-OH in AMP (**6**) (13.38), adenosine 3′-ethyl phosphate (**4**) (12.68), and adenosine (**3**) (12.17) to that of 3′-deoxyadenosine (**8**) (12.71) reveals relative stabilization of the 2′-oxyanion by the 3′-substituent: 3′-OH > 3′−H ≈ 3′-OPO_2EtO^- > 3′-OPO_3^{2-}. Interestingly, the pK_a value for 2′-OH in AMP (**6**) increases by 0.61 units compared to those of **4** and **8**, thereby showing that the 3′-phosphate dianion destabilizes the 2′-oxyanion

The comparison of pK_a for 2′-OH in 1-deoxy-D-ribofuranose (**1**) (13.05) with those of nucleosides (**3**, **10**, **14**, **16**) (12.7−12.62), and 1-deoxy-D-ribofuranose 3,5-*O*-bis-ethyl phosphate (**2**) (13.55) with those of nucleotides (**5**, **11**, **12**, **17**) (12.68−13.08) clearly shows that the C1′-aglycons, depending on their chemical character, have varied influence on the stabilization of the 2′-oxyanion. The 2′-OH in adenosine (**3**) and its 3′,5′-bis-ethyl phosphate derivatives (**5**) is the most acidic (Table 1) because adenin-9-yl is a better stabilizer for the 2′-oxyanion (by 0.08−0.45 pK_a units) compared to any other aglycons in the nucleoside, 3′-phosphomonoester, 3′-phosphodiester, or 3′,5′-bis-phosphodiester series. In addition to the effect of aglycon, the acidity of the 2′-OH is further modulated by the nature of the 3′-substituent, that is −OH versus $-PO_3^{-2}$, with virtually no effect for 3′-PO_3-$(Et)^{-1}$.

It is also likely that the pK_a of 2′-OH of different ribonucleotide units in an RNA molecule is steered by the variation of the local microenvironment and its hydrophobic character, depending upon the nature of RNA-folding, as observed for the aglycons in at least two-folded RNAs[7a] and for a DNA triplex.[7b] It is also possible that the different reactivities of phosphodiester functions, as observed in the pre-mRNA processing reaction (splicing)[1c,e] or in RNA catalysis (ribozyme),[1d] are steered by different acidities/basicities of the 2′-OH, dictating the site-specific 2′-OH-assisted transesterification reactions, in general. Finally, the absence of the ring oxygen in aristeromycin (**9**) (13.03) results in an increase of pK_a of 2′-OH by 0.86 units compared to that in adenosine (**3**) (12.17) and in the abasic sugar (**1**) (13.05). This means that the substitution of pentose sugar by a cyclopentane moiety in an RNA molecule will make the general acid−base-catalyzed 2′-OH-assisted transesterification reactions slower in the latter compared to the former. This again shows the importance of the pentose—vis-à-vis the cyclopentane-based RNA-world.

Acknowledgment. We thank the Swedish Board for Technical Development (TFR), Swedish Natural Science Research (NFR) Council, Swedish Board for Technical Development (NUTEK) for generous financial support. Thanks are due to the Wallenbergstifelsen and University of Uppsala for funds for the purchase of 500 and 600 MHz Bruker DRX NMR spectrometers. I.V. and S.A. have made equal contributions in the determination of pK_a's, while A.T. and A.F. have prepared the compounds for the pK_a measurements.

Supporting Information Available: 1. Table 1S: pK_a values of 2′-OH in ^{13}C-labeled adenosine, uridine, and cytidine obtained from the ^{13}C chemical shift of C1′, C2′, and C3′. 2. Figures 1S and 3S: proton and carbon chemical shift titration curves, respectively for **1**−**18** and ^{13}C-labeled adenosine, cytidine, and uridine. 3. Figures 2S and 4S: graphical determination of δ_h, respectively for **1**, **3**−**12**, **14**, **16**, **17,** and ^{13}C-labeled adenosine, cytidine, and uridine. 4. Figure 5S. Correlation plots between ^{13}C- and ^{1}H chemical shift for ^{13}C-labeled adenosine, cytidine, and uridine. 7. Experimental Section (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0036312

(6) Bernet, B.; Vasella, A. *Helv. Chim. Acta* **2000**, *83*, 995−1021.

(7) (a) Legault, P.; Pardi, A. *J. Am. Chem. Soc.* **1997**, *119*, 6621−6628. (b) Singleton, S. F.; Dervan, P. B. *Biochemistry* **1992**, *31*, 10995−11003.