Non-identical electronic characters of the internucleotidic phosphates in RNA modulate the chemical reactivity of the phosphodiester bonds†‡

Jharna Barman,§[a] Sandipta Acharya,§[a] Chuanzheng Zhou,[a] Subhrangsu Chatterjee,[a] Åke Engström[b] and Jyoti Chattopadhyaya*[a]

Received 28th November 2005, Accepted 20th December 2005
First published as an Advance Article on the web 2nd February 2006
DOI: 10.1039/b516733g

We here show that the electronic properties and the chemical reactivities of the internucleotidic phosphates in the heptameric ssRNAs are dissimilar in a sequence-specific manner because of their non-identical microenvironments, in contrast with the corresponding isosequential ssDNAs. This has been evidenced by monitoring the δ H8(**G**) shifts upon pH-dependent ionization (pK_{a1}) of the central 9-guaninyl (**G**) to the 9-guanylate ion ($\mathbf{G}^-$), and its electrostatic effect on each of the internucleotidic phosphate anions, as measured from the resultant δ ^{31}P shifts (pK_{a2}) in the isosequential heptameric ssRNAs *vis-à-vis* ssDNAs: [d/r(5′-Cp$_1$Ap_2Q$^1p_3$**G**p_4Q^{2}p$_5$Ap$_6$C-3′): Q^1 = Q^2 = A (**5a/5b**) or C (**8a/8b**), Q^1 = A, Q^2 = C (**6a/6b**), Q^1 = C, Q^2 = A (**7a/7b**)]. These oligos with single ionizable **G** in the centre are chosen because of the fact that the pseudoaromatic character of **G** can be easily modulated in a pH-dependent manner by its transformation to $\mathbf{G}^-$ (the 2′-OH to 2-O$^-$ ionization effect is not detectable below pH 11.6 as evident from the N$^{1\text{-Me}}$-**G** analog), thereby modulating/titrating the nature of the electrostatic interactions of **G** to $\mathbf{G}^-$ with the phosphates, which therefore constitute simple models to interrogate how the variable pseudoaromatic characters of nucleobases under different sequence context (*J. Am. Chem. Soc.*, 2004, **126**, 8674–8681) can actually influence the reactivity of the internucleotide phosphates as a result of modulation of sequence context-specific electrostatic interactions. In order to better understand the impact of the electrostatic effect of the **G** to $\mathbf{G}^-$ on the tunability of the electronic character of internucleotidic phosphates in the heptameric ssRNAs **5b**, **6b**, **7b** and **8b**, we have also performed their alkaline hydrolysis at pH 12.5 at 20 °C, and have identified the preferences of the cleavage sites at various phosphates, which are p_2, p_3 and p_4 (Fig. 3). The results of these alkaline hydrolysis studies have been compared with the hydrolysis of analogous N$^{1\text{-Me}}$-**G** heptameric ssRNA sequences **5c**, **7c** and **8c** under identical conditions in order to establish the role of the electrostatic effect of the 9-guanylate ion (and the 2′-OH to 2-O$^-$ ionization) on the internucleotidic phosphate. It turned out that the relative alkaline hydrolysis rate at those particular phosphates (p_2, p_3 and p_4) in the N$^{1\text{-Me}}$-**G** heptamers was reduced from 16–78% compared to those in the native counterparts [Fig. 4, and ESI 2 (Fig. S11)]. Thus, these physico-chemical studies have shown that those p_2, p_3 and p_4 phosphates in the native heptameric RNAs, which show pK_{a2} as well as more deshielding (owing to weaker ^{31}P screening) in the alkaline pH compared to those at the neutral pH, are more prone to the alkaline hydrolysis because of their relatively enhanced electrophilic character resulting from weaker ^{31}P screening. This screening effect originates as a result of the systematic charge repulsion effect between the electron cloud in the outermost orbitals of phosphorus and the central guanylate ion, leading to delocalization of the phosphorus pπ charge into its dπ orbitals. It is thus likely that, just as in the non-enzymatic hydrolysis, the enzymatic hydrolysis of a specific phosphate in RNA by general base-catalysis in RNA-cleaving proteins (RNase A, RNA phosphodiesterase or nuclease) can potentially be electrostatically influenced by tuning the transient charge on the nucleobase in the steric proximity or as a result of specific sequence context owing to nearest-neighbor interactions.

[a]*Department of Bioorganic Chemistry, Box 581, Biomedical Center, Uppsala University, S-751 23 Uppsala, Sweden. E-mail: jyoti@boc.uu.se*
[b]*Department of Medical Biochemistry and Microbiology, Box 582, Biomedical Center, Uppsala University, S-751 23 Uppsala, Sweden*

† Electronic supplementary information (ESI) available: detailed experimental section for pH titration by NMR, Hill plot analysis for pK_a determination, titration, stacked NMR plots for pH-dependent ^{1}H and ^{31}P shifts. Explanatory notes for pK_{a1} and pK_{a2}, and mechanism of deshielding of ^{31}P shifts as pH became alkaline, HPLC analysis of the alkaline hydrolysis products, rate constants and mass spectral analysis.. See DOI: 10.1039/b516733g

‡ In honor and celebration of the 75th Birthday of Professor Colin Bernard Reese, F.R.S.

§ The first authorship of this article is equally shared by Sandipta Acharya and Jharna Barman.

Introduction

Recently, the base-promoted RNA transesterification reaction and cleavage has been studied quite in some detail in many laboratories.[1–5] These studies on the susceptibility of cleavage of RNA phosphodiesters under the conditions of general base-catalysis have shed light on how the process of base-catalysis by RNA-cleaving proteins (such as RNase A[2a] or RNA phosphodiesterase[2b] or nuclease[2]) works in biology. The process of general base-catalysis, commonly used as one of the catalytic strategies by RNA-cleaving proteins, utilizes histidine-12 for deprotonation of the 2′-OH group for nucleophilic attack to the vicinal phosphodieser. It is likely that these studies will lead us to a better understanding of the action of catalytic nucleic acids and thus to the rational design of artificial RNases, which could be used in gene therapy, as well as to understanding more about the central role that RNA possibly played during the origins of life. Monomethyl and monoisopropyl esters of adenosine 2′- and 3′-monophosphates,[1a,b] dimers[1c] as well as various short 2′-*O*-methylated[1d–f,3j] or 2′-deoxy[1g,3b] oligonucleotide sequences containing only **one** reactive ribonucleotide unit in chimeric DNA/RNA oligonucleotides have been chosen for this purpose. Several factors[3] have so far emerged as prerequisites in this general base-catalysed hydrolysis of RNA phosphodiesters: (i) the nucleophilicity of the 2′-hydroxy group and its pK_a, (ii) the electrophilicity of the reacting phosphate, (iii) the in-line conformation of the attacking 2′-oxyanion with the developing 5′-oxyanion, (iv) the readiness with which the 5′-oxyanion leaving group departs (for example, upon binding to a metal ion), as well as (v) the intramolecular environment[1,3f–l] (stacking, hydrogen bonding and the nucleobase composition) around the transesterification site, which is also believed to modulate the structure of oligonucleotide sequence that facilitates or retards the transesterification reaction. It has also been proposed[3f] that the general base-catalysis and cleavage of RNA phosphodiesters can be effectively accelerated when the attacking nucleophile is in a 'near attack conformation' for the reaction to proceed.[3f] For example, a well-stacked rigid structure (giving an A-type conformation) would perhaps retard the base-promoted cleavage of a specific RNA phosphodiester[3g] compared with a disordered structure. The reaction rate also depends on structural factors such as whether or not the cleaving site is within a single or double strand or in a hairpin region.[1g,3j] The cleavage rate of the sequence within a hairpin was shown[1g,3j] to be different depending on its position in the stem or loop of the hairpin.

On the other hand, for an efficient transesterification to take place in the enzymatic hydrolysis of RNA phosphodiester bonds in the self-splicing of pre-mRNA[4] in *Tetrahymena thermophila*,[4,5] processing of tRNA precursors by the RNA moiety of RNase P,[5] group II intron ribozyme[4,5] hammerhead[4,5] and hairpin ribozymes,[4,5] genomic and antigenomic hepatitis delta virus (HDV) ribozymes[4,5] and Varkud satellite (VS) ribozymes[4,5] a precisely folded structure is required: precise substrate recognition, much like their protein counterparts, is achieved by intricate structure formation. It is now a well-known fact[5] that both large and small ribozymes possess an 'internal guide sequence' (IGS) that after Watson–Crick base-pairing with the target RNA juxtaposes the reacting and catalytic groups. This process of substrate recognition and biocatalysis is often aided by folding of the enzyme–substrate complex[5] using flexible domains as well as tertiary interactions to form the catalytic core such that the reacting groups come close to each other, and with the help of metal ions like Mg^{2+} or nucleic acid bases with environmentally perturbed pK_a values[6] result in cleavage or ligation.[5]

It is known that the electronic environment around the phosphate in RNA plays a vital role in the RNA–protein interaction: for example, the crystal structure[7a] of the glutaminyl tRNA synthetase (GlnRS)-tRNA^{Gln} complex shows that hydrogen bonding between the amino group of a guanine and the 5′-phosphate of an adenine in tRNA contributes to the recognition of GlnRS by tRNA^{Gln}. The RNA binding domain in sex-lethal proteins[7b,c] recognizes and binds uridine-rich sequences. A particular turn in this binding is stabilized by a protein-RNA interaction as well as by three hydrogen bonds formed between 2′-OH and the phosphates within the RNA. Binding of hairpin loop IV of U2 small nuclear RNA to the splicesomal protein U2B′ involves[7d] hydrogen bonds between the amino groups and phosphates in the RNA as well as salt bridges between the amino acid residues of the protein and the phosphates of RNA.

However, it is not understood if some specific scissile phosphodiesters in RNA have variable charges in comparison with the other phosphates owing to some sequence-specific unique scaffolds or folding motifs. This could be the result of nearest-neighbor stacking interactions, non-canonical basepairing as well as by complexation with some specific metal ion or protein. Any of these interactions can alter structural motifs of RNA, which can potentially alter the local microenvironment of a specific internucleotidic phosphate, and making it electronically and chemically non-uniform by charge rearrangements. It is also not understood if the variable phosphate charges can potentially tune the chemical reactivity in the biologically ubiquitous transesterification reactions. Clearly, any further information in this complex problem has a deep consequence in our further understanding of the chemical basis of the ubiquitous RNA transesterification reactions in general.

In this work, we have interrogated this complex problem – as to whether or not there is any modulation of the chemical reactivity of the phosphates by the electronic nature of a specific sequence context – by generating a single charge at the 9-guaninyl center of the model heptameric DNA and RNA sequences [d/r(5′-Cp$_1$Ap_2Q$^1p_3$**G**p_4Q^{2}p$_5$Ap$_6$C-3′): Q^1 = Q^2 = A (**5a/5b**) or C (**8a/8b**), Q^1 = A, Q^2 = C (**6a/6b**), Q^1 = C, Q^2 = A (**7a/7b**)]. Owing to the gradual changes of the electrostatic interactions of the internucleotidic phosphate anions with the central pseudoaromatic 9-guaninyl group (**G**) as it is ionized to N^1-guanylate ion (**G**$^-$) from pH 7 to 12.5, this model study has allowed us to electrostatically titrate the electronic/chemical nature of each of the phosphates, depending upon their respective microenvironments, with an intramolecular ionization center (**G**$\rightarrow$ **G**$^-$) as a reference point. We considered that the easiest way to study this titration of the electrostatic interaction of internucleotidic phosphates and **G**/**G**$^-$ and 2′-OH/2-O$^-$ ionizations is to monitor simultaneously the pH-dependent ^{31}P chemical shift changes of all the phosphates as well as of δ H8(**G**) in a systematic manner in isosequential ssDNAs, *vis-à-vis* ssRNAs as a result of 9-guanylate formation at the center. This electrostatic titration was then compared with the alkaline degradation studies of our ssRNAs in order to examine if there is any selectivity or modulation in the spontaneous alkaline cleavage

by transesterification due to the presence of a charged center ($\mathbf{G}^-$) in the steric proximity. Furthermore, **G** in **5b**, **7b** and **8b** was replaced with $N^{1\text{-Me}}$-**G** in the heptameric ssRNA sequences **5c**, **7c** and **8c** in order to quench the formation of the N^1-guanylate ion ($\mathbf{G}^-$), and the results of their alkaline degradation studies were then compared with those of the native heptameric RNAs (**5b**, **7b** and **8b**). This helped us to verify how the presence of $N^{1\text{-Me}}$-**G** in place of **G** can influence the preferential cleavage of internucleotidic phosphates (taking into consideration the 2′-OH to 2-O^- ionization). This methylation study also allowed us to address, at least partly, how the electrostatic effect of the pseudoaromatic characters of various nucleobases (as a result of sequence context variation) can potentially dictate the reactivity of the internucleotidic phosphates in the proximity.

Here, we provide the first unequivocal evidence in these model heptameric ssRNA systems, showing that the electronic character of the internucleotidic phosphates in **5a/b**–**8a/b** (Fig. 1) in ssRNAs is indeed non-identical, whereas they are more or less uniform in the corresponding ssDNAs. We also show that there are indeed some interesting preferences in the alkaline hydrolytic cleavage of some specific internucleotidic phosphodiester linkages in the native ssRNAs as a result of electrostatic interactions between those phosphates and the guanylate ion ($\mathbf{G}^-$), which could be remarkably reduced by substitution of the central **G** with $N^{1\text{-Me}}$-**G**. Thus, the alteration of the central **G** to $\mathbf{G}^-$ and **G** to $N^{1\text{-Me}}$-**G**, as simple model systems, allows us to interrogate and compare the electronic and hydrolytic properties of the phosphates by engineering the electrostatic effects within a given

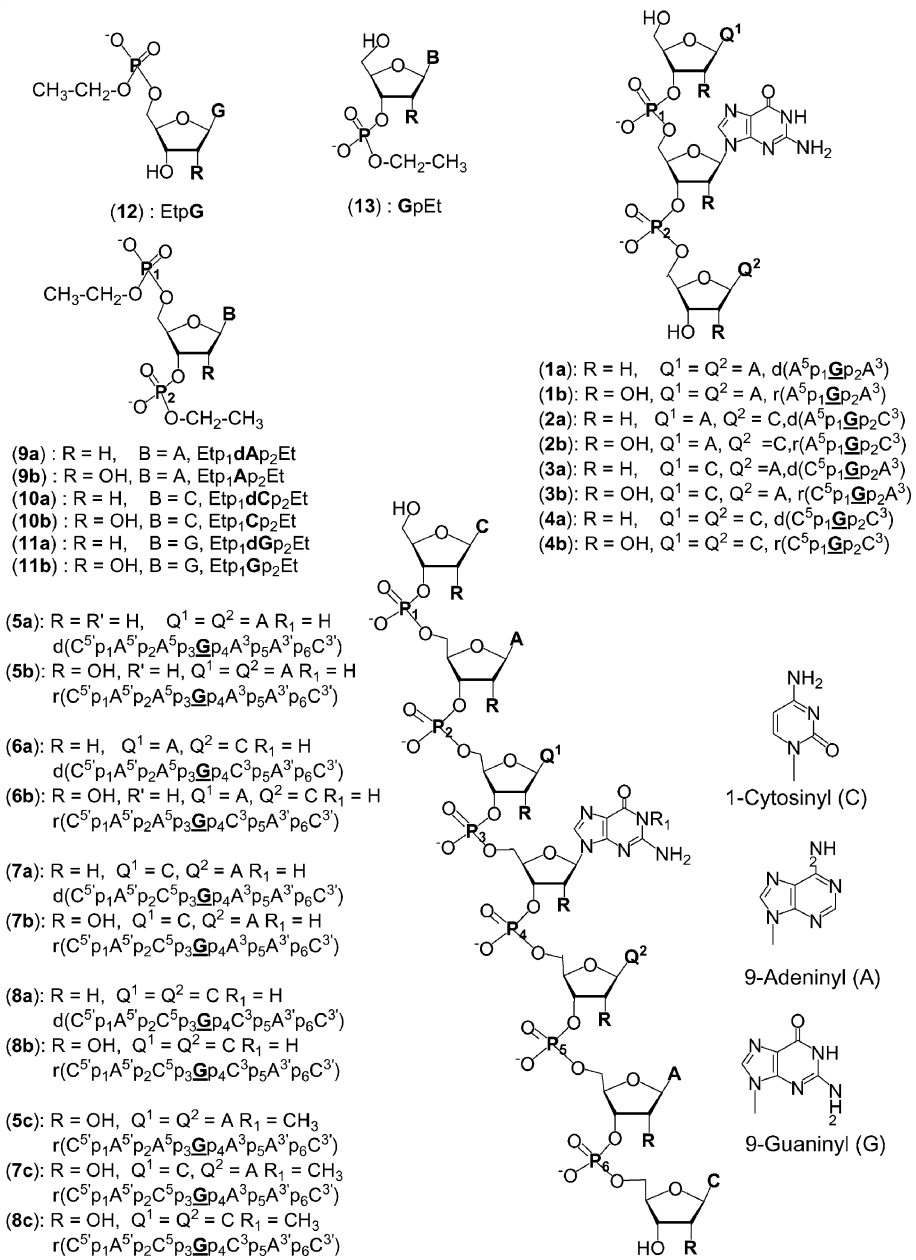

Fig. 1 Model compounds used in the present work.

set of analogous RNAs (only partly isosequential) with different sequence contexts in which both the hydrogen bonding and dipole-moment pseudoaromatic characteristics have been modulated.

Results and discussion

The eight ssDNA/ssRNA heptameric sequences used for the pH-dependent ^{31}P chemical shift study (pH 6.6–12.5) had a central **G** moiety with either 5′-purine(A)-**G**-purine(A)-3′, 5′-purine(A)-**G**-pyrimidine(C)-3′, 5′-pyrimidine(C)-**G**-purine(A)-3′ or 5′-pyrimidine(C)-**G**-pyrimidine(C)-3′ as the immediate neighboring nucleobases in the 5′- and 3′-ends respectively. Trimeric ssDNA/ssRNA sequences [d/r(Ap$_1$**G**p$_2$A) (**1a**/**1b**), d/r(Ap$_1$**G**p$_2$C) (**2a**/**2b**), d/r(Cp$_1$**G**p$_2$A) (**3a**/**3b**), d/r(Cp$_1$**G**p$_2$C) (**4a**/**4b**), Fig. 1], constituting the central trinucleotidyl core of the corresponding heptamer, were used as the internal reference compounds. In the isosequential heptameric ssDNAs (**5a**–**8a**) and ssRNAs (**5b**–**8b**) the central trimeric sequences (Q^1**G**Q^2) are extended both at the 3′- and 5′-ends by the AC residues.

The internucleotidic phosphodiesters (pK_a 1.5)[9] in the ssDNAs and ssRNAs (**1**–**8**) are fully negatively charged under our measurement conditions (pH 6.6–12.5). The ^{31}P resonances for each of the phosphates are shifted downfield[11] [ESI 1 (Note 2 and Table S6)†] due to the formation of **G**$^-$ from pH 6.6 to 12.5 for **1**–**8**, and show sigmoidal behavior giving an inflection point typical of a titration curve. An example of the pH-dependent titration profile is shown for the heptameric ssRNA **8b** (Fig. 2), and similar titration profiles for the other ssDNAs and ssRNAs are shown in ESI 1 (Fig. S1–S3). The pK_a values[8a] of **G**, obtained from the titration curves from each of the δ ^{31}P resonances (pK_{a2}), are determined by different methods including non-linear curve-fitting [ESI 1 (Fig. S1–S3)] and Hill plot analysis [ESI 1 (Experimental section D and Fig. S6 and S7)]. The pK_a of **G** obtained from the pH-dependent ^{31}P chemical shifts [Table 1 and ESI 1 (Table S1 and Fig. S1–S3)] of various ^{31}P resonances in the oligo-RNAs is a result of the systematic change (titration profile) of the electrostatic interaction of each of the negatively charged phosphates with the central pseudoaromatic 9-guaninyl group as it is gradually transformed into the 9-guanylate ion as the pH becomes alkaline. Interestingly, it can be seen that the pK_a of **G** obtained from the pH-dependent ^{31}P chemical shifts (pK_{a2}) are variable in a sequence-dependent manner, and are indeed different from those of the directly measured pK_a from its own pH-dependent δ H8(**G**) shifts (pK_{a1}).[8a]

The reason for our observation of variable pK_{a2} values from various internucleotidic phosphodiesters is as follows: the electrostatic potential energy $E = Q_1 \times Q_2/4\pi\varepsilon_0 r$, where Q is the electric charge for $Q_1 = \mathbf{G}^-$ and $Q_2 = PO_2^{-1}$, ε_0 = the permittivity factor (depending upon the microenvironment around each of the phosphates) and r = the distance between the charge generation site **G**$^-$ and the phosphate. Thus, depending upon the local phosphate charge variation, *i.e.* if all the phosphates are not electronically identical in a given ssDNA or ssRNA sequence, then Q_2 can be different for different phosphates depending upon each of their local microenvironments, which can potentially be dictated by any or a combination of the following factors: (i) the specific electronic character of a nucleobase in a given sequence context (through which the Q_1 will be modulated) due to the nearest-neighbor relationship (stacking/destacking equilibrium), (ii) the interaction with a metal ion cofactor, (iii) the partial charge generation – cationic or anionic, (iv) the non-covalent interaction through distant neighboring group participation, (v) the folding pattern, and (vi) the varying hydration capabilities around each of the internucleotidic phosphates, where hydration is due to conformational changes. Thus, this variable nature of Q_1 and Q_2 makes the E for each of the internucleotidic phosphates variable, which can be observed from the differently perturbed pK_{a2} values.[8a]

(I) Accuracy of the pH-dependent NMR titration studies

The pK_{a1} values for the ionization at N^1 center of **G** [obtained from δH8(**G**) marker protons of the 9-guaninyl residue] used for comparison in our present work were obtained in our earlier work[8a] by the pH-dependent ^{1}H chemical shifts measured by both 500 and 600 MHz NMR,[8a] followed by the Hill plot analysis. The pK_{a2} values reported here [Table 1 and ESI 1 (Table S1 and Fig. S1–S3)†] are obtained by the pH-dependent chemical shift of the ^{31}P markers (p_1–p_6 markers in the case of heptameric ssDNA and ssRNA **5**–**8**; and the p_1 and p_2 markers for trimeric ssDNA and ssRNA **1**–**4**) in the pH range 6.6–12.5. The ^{31}P chemical shifts have three sources of error [see ESI 1 (Experimental section B)]: (i) the error owing to the digital resolution (the maximum error is 0.001 ppm); (ii) the error from the line broadening of the phosphorus (the maximum error is 0.06 ppm); and (iii) the error due to the salt effect[10] as the pH is changed in a stepwise manner by adding NaOD from pH 6.6 [the maximum error is 0.061 ppm upon the addition of a maximum of 16 mM NaOD for reaching pH 12.5; ESI 1 (Experimental section B)]. Taking into account all of the above sources of error in our observed ^{31}P chemical shifts and applying the salt effect correction [ESI 1 (Experimental section B)] over the pH values 6.6–12.5, the pK_a of guanine from ^{31}P markers was estimated only when the total change in the salt-effect-corrected ^{31}P chemical shifts was above 0.10 ppm [*i.e.* $\Delta\delta\,^{31}P_{(pH'''6.6-pH12.5)} >$ 0.10 ppm].

Exceptions are ^{31}P resonances of some ssRNA sequences (p_6 in **6b**, **7b** and **8b**) where $\Delta\delta\,^{31}$P was >0.10 ppm over the pH range 6.6–12.5, but no pK_a value was calculated from these resonances since the pH *vs.* $\delta\,^{31}$P did not show satisfactory sigmoidal behavior. The change in $\delta\,^{31}$P at higher alkaline pH values in these resonances not showing pK_{a2} was a result of the effect of the 2′-OH ionization on the ^{31}P chemical shifts. The effect of 2′-OH→2′-O$^-$ on the evolution of the pH-dependent $\delta\,^{31}$P of a specific internucleotidic phosphate was found to be variable depending upon the sequence context [see Fig. 2, and ESI 1 (Fig. S1)†]. This effect was found to be prevailing in most of the ^{31}P resonances in the pH range 11.9–12.5 [see ESI 1 (Experimental section D) for a detailed discussion of the different pH values from which $\delta\,^{31}$P was affected by the 2′-OH ionization for different ^{31}P markers]. As a result, although pH titrations for all of the heptameric ssRNAs were performed at pH 12.5 for **5b**, **6b** and **8b** and pH 12.75 for **7b** all of the data points have not been used for the determination of pK_{a2} [Fig. 2 and ESI 1 (Fig. S1) show the data points not considered for pK_{a2} determination in (▲)]. The pH titration of compound **8c**, an N$^{1\text{-Me}}$-**G** analogue of **8b**, was done in the pH range 8.0–12.5 as a control study and it was found that the ^{31}P chemical shifts of the phosphate markers in **8c** were influenced by the vicinal 2′-OH ionization but not before pH 11.6 (Fig. 2).

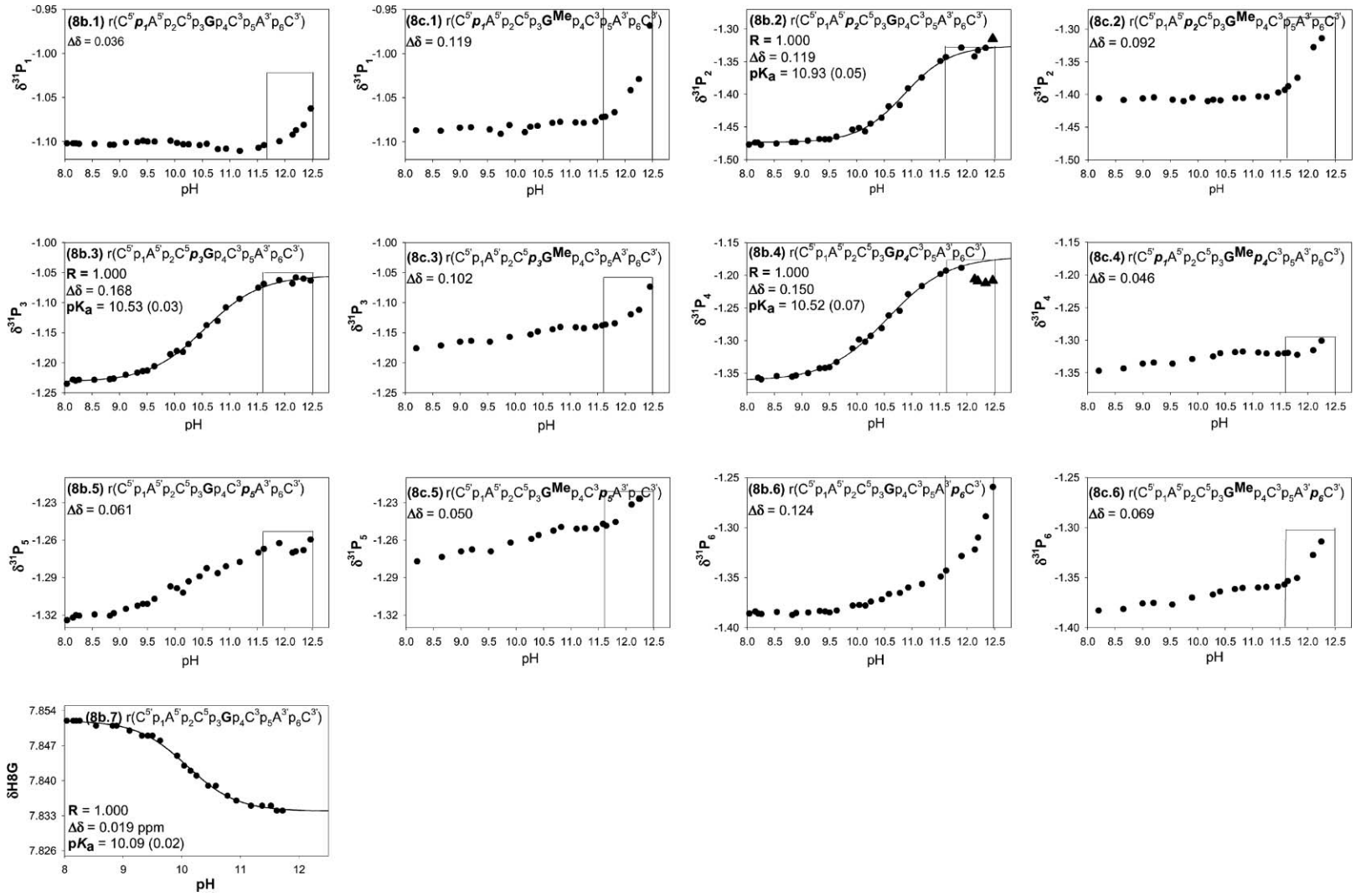

Fig. 2 Plots of pH-dependent (6.6–12.5) ^{31}P chemical shifts (δ^{31}P) for different ^{31}P markers (p_1–p_6) of ($C^{5'}p_1A^{5'}p_2C^{5}p_3\underline{G}p_4C^{3}p_5A^{3'}p_6C^{3'}$) (**8b**) (panels **8b**.**1**–**8b**.**6**) and ($C^{5'}p_1A^{5'}p_2C^{5}p_3\underline{G}^{Me}p_4C^{3}p_5A^{3'}p_6C^{3'}$) (**8c**) (panels **8c**.**1**–**8c**.**6**), as well as ^{1}H chemical shifts for the H8(G) marker of **8b** (panel **8b**.**7**). The pH was corrected for the deuterium effect, and δ^{31}P was corrected for the salt effect (see Experimental). The pH-dependent δ^{31}P/^{1}H plots show the pK_a at the inflection point. Note that in the case of markers p_2 and p_4 in **8b**, data points for δ^{31}P have been taken up to pH values 12.34 and 11.9 respectively for the pK_a calculation as higher pH values involved the effect of the 2′-OH ionization on δ^{31}P. *Data points with pH values of 12.34–12.5 in the case of p_2 and pH values of 11.9–12.5 in the case of p_4 are shown in the graph with the symbol ▲ but were not considered while doing the curve-fitting analysis because of the contribution from the 2′- OH ionization.* The correlation coefficient (R) obtained from each non-linear curve fitting (in panels **8b**.**1**–**8b**.**4**) through the experimental NMR titration points, the average pK_a values obtained for each of the ^{31}P and ^{1}H markers, and the total change in ^{31}P and ^{1}H chemical shift over the pH range (6.6–12.5) ($\Delta\delta$) has been shown in the graph. *The data points for the phosphate markers in the box in each panel over the pH range 11.6–12.5 are shown to emphasize the relative rate of formation of the 2′-oxyanion in **8b*** versus ***8c***.

Table 1 Comparison of pK_a values of **G** obtained from the H8(G) marker (pK_{a1}) and the phosphorus markers (pK_{a2}) in ssDNA (**5a**–**8a**) and ssRNA (**5b**–**8b**)

Compound name			pK_a[a]	pK_a[b]	pK_a[c]	pK_a[d]	pK_a[e]	average pK_a	pK_a from δH8(**G**)
d(Cp$_1$Ap$_2$Ap$_3$**G**p$_4$Ap$_5$Ap$_6$C)	**5a**	p_2	—	—	—	—	—	—	11.06 ± 0.01
		p_3	11.07 ± 0.03	—	11.04 ± 0.03	—	10.95 ± 0.03	11.02 ± 0.06	
		p_4	11.01 ± 0.03	—	10.99 ± 0.02	—	10.84 ± 0.05	10.95 ± 0.09	
r(Cp$_1$Ap$_2$Ap$_3$**G**p$_4$Ap$_5$Ap$_6$C)	**5b**	p_2[f]	10.46 ± 0.03	10.42 ± 0.04	—	10.41 ± 0.10	—	10.43 ± 0.03	10.58 ± 0.01
		p_3[f]	10.56 ± 0.04	10.53 ± 0.06	—	10.49 ± 0.10	—	10.53 ± 0.04	
		p_4	10.94 ± 0.04	10.89 ± 0.04	10.96 ± 0.06	10.77 ± 0.02	10.83 ± 0.02	10.88 ± 0.08	
d(Cp$_1$Ap$_2$Ap$_3$**G**p$_4$Cp$_5$Ap$_6$C)	**6a**	p_2	—	—	—	—	—	—	10.74 ± 0.02
		p_3	10.68 ± 0.03	—	10.68 ± 0.03	—	10.68 ± 0.02	10.68 ± 0.00	
		p_4	10.64 ± 0.03	—	10.64 ± 0.03	—	10.64 ± 0.03	10.64 ± 0.00	
r(Cp$_1$Ap$_2$Ap$_3$**G**p$_4$Cp$_5$Ap$_6$C)	**6b**	p_2	10.46 ± 0.03	10.44 ± 0.02	10.44 ± 0.02	10.29 ± 0.04	10.46 ± 0.04	10.42 ± 0.07	[g]
		p_3	10.40 ± 0.03	10.39 ± 0.03	—	10.36 ± 0.05	—	10.38 ± 0.02	
		p_4	10.72 ± 0.05	10.67 ± 0.04	10.65 ± 0.04	10.50 ± 0.05	10.61 ± 0.09	10.63 ± 0.08	
d(Cp$_1$Ap$_2$Cp$_3$**G**p$_4$Ap$_5$Ap$_6$C)	**7a**	p_2	—	—	—	—	—	—	10.79 ± 0.04
		p_3	10.64 ± 0.06	—	10.71 ± 0.04	—	10.59 ± 0.06	10.65 ± 0.06	
		p_4	10.69 ± 0.03	10.69 ± 0.03	—	10.64 ± 0.02	—	10.67 ± 0.03	
r(Cp$_1$Ap$_2$Cp$_3$**G**p$_4$Ap$_5$Ap$_6$C)	**7b**	p_2	10.31 ± 0.04	10.30 ± 0.04	—	10.14 ± 0.06	—	10.25 ± 0.10	[g]
		p_3	10.15 ± 0.03	10.15 ± 0.03	—	10.06 ± 0.06	—	10.12 ± 0.05	
		p_4	11.00 ± 0.06	10.90 ± 0.07	—	10.75 ± 0.08	—	10.88 ± 0.13	
d(Cp$_1$Ap$_2$Cp$_3$**G**p$_4$Cp$_5$Ap$_6$C)	**8a**	p_2	10.26 ± 0.04	10.32 ± 0.03		10.36 ± 0.11		10.31 ± 0.05	10.39 ± 0.01
		p_3	10.49 ± 0.03	10.43 ± 0.04		10.32 ± 0.06		10.41 ± 0.09	
		p_4	—	—	—	—	—	—	
r(Cp$_1$Ap$_2$Cp$_3$**G**p$_4$Cp$_5$Ap$_6$C)	**8b**	p_2	10.86 ± 0.04	10.94 ± 0.04	10.92 ± 0.04	10.98 ± 0.02	10.97 ± 0.02	10.93 ± 0.05	10.09 ± 0.02
		p_3	10.55 ± 0.02	10.55 ± 0.03	—	10.49 ± 0.07	—	10.53 ± 0.03	
		p_4	10.55 ± 0.04	10.56 ± 0.05	—	10.44 ± 0.07	—	10.52 ± 0.07	

All pK_a values and their corresponding errors are calculated from non-linear curve-fitting analysis and Hill plot analyses (see Experimental section D for details).[a] pK_a is obtained from non-linear curve fitting of pH *vs.* experimental δ^{31}P plot. [b] pK_a is obtained from non-linear curve fitting of pH *vs.* experimental δ^{31}P including the calculated δ^{31}P at the deprotonated state using δ^{31}P values between pH 10 and 11.6 in the calculation of the chemical shift of the deprotonated state. [c] pK_a is obtained from non-linear curve fitting of pH *vs.* experimental δ^{31}P including the calculated δ^{31}P at the deprotonated state using δ^{31}P values between pH 10 and 11.9 in the calculation of the δ^{31}P of the deprotonated state. [d] pK_a is obtained from the Hill plot analysis using δ^{31}P values between pH 10 and 11.6 in the calculation of the chemical shift at the deprotonated state as well as the Hill plot analysis. [e] pK_a is obtained from the Hill plot analysis using δ^{31}P at the deprotonated state using δ^{31}P values between pH 10 and 11.9 in the calculation of the δ^{31}P of the deprotonated state as well as the Hill plot analysis. [f] Could not be assigned exactly whether or not it is p_2 or p_3 due to the unavailability of ^{1}H–^{31}P cross peaks (see Fig. S12 in the ESI for details). [g] No titration profile found (see ref. 8*d*).

The corresponding error in pK_a determination is estimated to be between ±0.01 and ±0.09 in the case of trimeric ssDNAs and ssRNAs **1**–**4** and between ±0.00 and ±0.13 for the heptameric ssDNAs and ssRNAs **5**–**8**. All individual errors of the respective pK_a values are shown in Table 1 and ESI 1 (Table S1)†. The pH measurements were performed twice inside the NMR tube, both before and after each NMR titration point [30–40 pH points within the pH range 6.6–12.5 for each compound shown in Table 1 and ESI 1 (Table S1)], and the pH readings were found to vary by only ±0.025, hence no buffer was used for our study. The details of the pH measurement and calibration procedure can be found in ESI 1.

(II) Non-identical electronic characters of the internucleotidic phosphodiesters in the heptameric RNAs, but not in the isosequential DNAs

Table 1 and ESI 1 (Table S1), respectively, summarize the pK_{a1} values of **G** as obtained from its own δH8 (from our previous work[8a]) as well as the pK_{a2} values obtained from the pH-dependent ^{31}P chemical shift data for different phosphate markers in the trimeric [**1**(**a**/**b**)–**4**(**a**/**b**)] and heptameric [**5**(**a**/**b**)–**8**(**a**/**b**)] RNAs (with the exception of ^{31}P markers p_1, p_5 and p_6, from which pK_a values of G could not be estimated). By comparing the pK_as of 9-guaninyl, obtained from the pH-dependent ^{31}P chemical shifts (pK_{a2}) within the different phosphate markers as well as that obtained from its δH8(**G**) (pK_{a1}), the following key observations can be made.

(1) The maximum difference between pK_{a1} and any of the pK_{a2}s within a particular sequence is 0.07 pK_a units in **1a** among the trimers (**1**–**4**) and 0.14 pK_a units in **7a** among the ssDNA heptamers (**5a**–**8a**), which are very close to the error limit of ±0.13 pK_a units [see Table 1 and ESI 1 (Table S1 and Experimental section D)†]. The maximum difference is 0.02 pK_a units between the pK_{a2} values obtained from the ^{31}P markers p_1 and p_2 in **2b** and **3a** among the trimeric ssDNAs/ssRNAs (**1**–**4**). Similarly, among the heptameric ssDNAs (**5a**–**8a**), the maximum difference is 0.10 pK_a units between the pK_{a2} values obtained from the ^{31}P markers p_2 and p_3 in **8a**.

(2) The pK_{a1} and pK_{a2} values differ from each other significantly in that the maximum difference between pK_{a1} and pK_{a2} (from the ^{31}P marker p_2) is 0.84 pK_a units in the ssRNA sequence **8b**. On the other hand, the minimum difference between pK_{a1} and pK_{a2} (from the ^{31}P marker p_4) is found to be 0.05 pK_a units in heptameric ssRNA **5b** (Table 1). The pK_{a2} values obtained from each of the marker phosphates (p_2–p_4) in the heptameric ssRNAs (**5b**–**8b**) also differ from each other (Table 1) depending upon the sequence context. In the two ssRNA sequences **6b** and **7b**, unfortunately we do not have the pK_{a1} of guanine from δH8(**G**) to compare with the pK_{a2} obtained from the ^{31}P marker because of the repulsive electrostatic effect between the 5′-phosphate of **G** and its imidazole moiety.[8] The pK_{a2}s obtained from the ^{31}P markers within a sequence, however, differ from each other as observed from 10.38 (from the δ ^{31}P marker p_3) to 10.63 (from the δ ^{31}P marker p_4) in **6b** as well as from 10.12 (from the δ ^{31}P marker p_3) to 10.88 (from the δ ^{31}P marker p_4) in **7b**. Thus, among heptameric ssRNAs (**5b**–**8b**) the maximum difference between the pK_{a2} values obtained from the ^{31}P markers p_3 and p_4 in **7b** is 0.76 pK_a units and the minimum difference is 0.25 pK_a units between the pK_{a2} values obtained from the ^{31}P markers p_3 and p_4 in **6b**.

(3) The internucleotidic phosphates in the ssDNAs and ssRNAs are fully ionized at the studied pH range of 6.6–12.5. Hence, the observed downfield shifts of all ^{31}P resonances from neutral to alkaline pH is a result of through-space repulsive electrostatic interaction[11] of the phosphate anion and the **G**$^-$ [see ESI 1 (Notes 1 and 2)†]. The observed downfield shift of the ^{31}P resonances [Fig. 2, and ESI 1 (Fig. S1–S3)] reflects the weaker screening of the ^{31}P nucleus due to the delocalization of charge into its dπ orbitals[11] [ESI 1 (Note 2)] as **G** becomes **G**$^-$ over the pH range of 6.6–12.5. This is very similar to the earlier observed downfield ^{31}P shifts in various types of phosphates,[11a–h] phosphonates[11a,b] and aminophosphonates,[11b] as they are ionized with an increase of pH[11] [see also Section III(3) on the alkaline hydrolysis of the heptameric RNAs].

(4) The effect of the formation of **G**$^-$ is felt by the internucleotidic phosphates only in close proximity at both the 3′- and 5′-ends of the ionization site in the heptameric ssDNA and ssRNA (mainly the p_2, p_3 and p_4 markers show the apparent pK_{a2} of guanine, Fig. 3). Thus, the generation of **G**$^-$ at the center of the ssDNA or ssRNA chain allows us to probe only the relative electronic character of the neighboring internucleotide phosphates (*i.e.* of p_2, p_3 and p_4). It can be seen from the various pK_{a2}s obtained from the ^{31}P markers (Table 1) that the largest increases in pK_{a2} are seen when **G** is sandwiched between two C residues, as in **8b**.

(5) As stated above, the pK_{a2} of **G** obtained from the trimeric and heptameric ssDNAs/ssRNAs is a result of the interaction of **G**$^-$ with the internucleotidic phosphate anion. The relationship $E = Q_1 \times Q_2/4\pi\varepsilon_0 r$ suggests that as the distance r increases the electrostatic potential energy decreases, which means that the pK_{a2} readout should decrease with respect to pK_{a1} as the distance increases, in an expected manner, provided that all of the internucleotidyl phosphates experience an identical microenvironment. It is also implicit that when a particular phosphate experiences a more hydrophobic local environment [ESI 1 (Note 3)†] it should have a lower dielectric contribution $4\pi\varepsilon_0$ compared with the other phosphates, which means that it should show a higher pK_{a2}. This is consistent with the fact that the pK_a of the carboxylic acid group in salicylic acid in DMSO (6.6) is much higher compared with that in water (2.9).[12] Note that this dielectric argument also applies to the nucleobases, which are in a more hydrophobic microenvironment[12] [ESI 1 (Note 3)] compared to the internucleotidic phosphates, and hence they should show a larger pK_{a1} than the pK_{a2} from the phosphates. Hence, our observations of those pK_{a2} values which are lower than or similar to the pK_{a1} (as in p_4 of **5a**; p_2 and p_3 of **5b**; p_3 and p_4 of **7a**) are no surprise, they are indeed well expected (Table 1). In contradistinction, what is interesting and unexpected are our observations of unique pK_{a2}s from some phosphates (p_4 of **5b**; and p_2, p_3 and p_4 of **8b**, Table 1), which show a larger pK_{a2} than the pK_{a1}. The larger pK_{a2} values of these phosphates compared with the pK_{a1} can be clearly attributed to their charge rearrangements[2a,12a] due

to the electrostatic modulation of their electronic character by a hydrophobic local microenvironment [ESI 1 (Note 3)], which has a lower dielectric contribution compared with that of the aqueous environment. Thus, those specific phosphates which are in a more hydrophobic environment than the other phosphates are likely to show a relatively higher pK_a than the average pK_a normally measured for the internucleotide phosphates (1.5 [9a,c] to 2.1 [9b]). This means that those phosphates in the hydrophobic pocket will be relatively less ionized and hence likely to be more electrophilic; clearly, they will also have a poorer ability to repel nucleophiles such as the 2′-oxyanion or the hydroxide. In fact, they are the ones which are more vulnerable than the others to a potential transesterification reaction by the 2′-oxyanion or the hydroxide ion. Interestingly, we have observed a very similar trend: that is, those internucleotidic phosphates which show a pK_{a2} are more prone to cleavage in the alkaline hydrolysis of the heptameric RNAs [see also Section III(3) on the alkaline hydrolysis of the heptameric RNAs].

We are tempted to postulate that such a hydrophobic pocket around the scissile phosphate may be created in the catalytic core of both large and small ribozymes when an 'internal guide sequence' (IGS) Watson–Crick base-pairs with the target RNA to form the catalytic core such that the reacting and catalytic groups come close to each other and also with the help of metal ions such as Mg^{2+} or nucleic acid bases with environmentally perturbed pK_a values. This increases the pK_a of the scissile phosphate slightly, which will make it partially protonated at the physiological pH in preference to the others, thereby steering the electrophilicity of the internucleotidyl scissile phosphate to be distinctly higher than the others in a large polymeric RNA. Consequently, those protonated phosphates will have a poorer ability to repel nucleophiles such as the 2′-oxyanion or the hydroxide, causing specific transesterification to take place.

(6) It is noteworthy that, although the nucleobases in both ssDNA and ssRNA do become destacked[8,13,14] as the 9-guanylate ion is formed in the alkaline pH, the sugar-phosphate backbone does not change appreciably [ESI 1 (Note 4)†], as became evident from the monitoring endocyclic $^3J_{HH}$ of the constituent sugar rings [ESI 1 (Table S2 and S3)]. The present data therefore suggest that the role of the conformational effect due to sugar-phosphate backbone preorganization is minimal for an O–P–O bond angle change [see ESI 1 (Note 2)] as well as for the change of charge densities around each phosphate marker across the RNA chain compared with that of the contribution due to the change in the local microenvironment.

(III) Comparative alkaline hydrolysis of the heptameric native RNAs and their N^1-methylated G counterparts

In order to correlate the results from the above electrostatic titration studies of the internucleotidic phosphodiesters by the central $\mathbf{G}^-$ and the 2′-oxyanion in the heptameric RNAs **5b**, **6b**, **7b** and **8b** with their respective chemical reactivities, we have carried out their alkaline hydrolytic cleavage at pH 12.5 at 20 °C, and compared the results with those from the cleavage of the analogous N^1-**G**-methylated (N$^{1\text{-Me}}$-**G**) heptameric **5c**, **7c** and **8c** under identical conditions. Their time-dependent alkaline hydrolysis was monitored by RP-Hplc analysis [ESI 2 (Fig. S12A); ESI 3 (Fig. S12B); and ESI 5 (Fig. S15A) to ESI 11], and their pseudo first-order rate constants were closely comparable [Fig. 3 and Fig. 4, and ESI 2 (Fig. S10 and S11)†]. The types of hydrolysis products formed after 15, 30 and 60 min were identical, which meant that those that were formed up to 60 min of the digestion

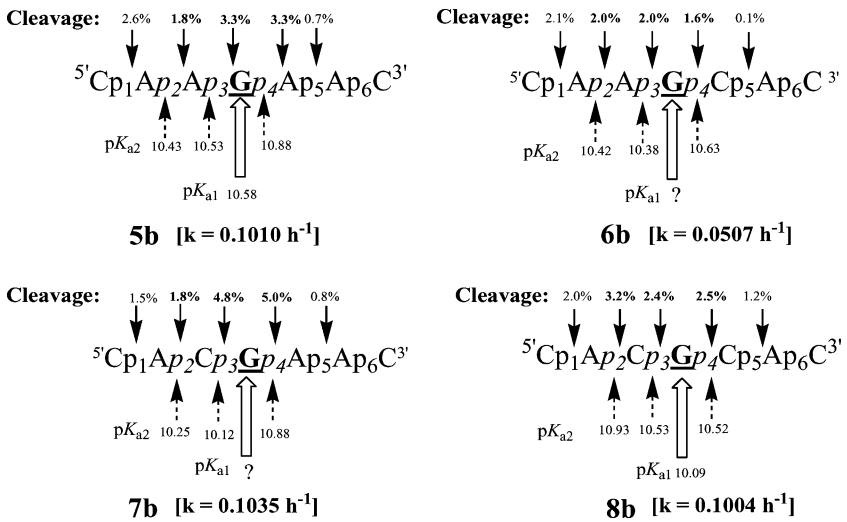

Fig. 3 Alkaline hydrolysis (in 0.03 N aqueous NaOH) of the internucleotidic phosphates at pH 12.5 at 20 °C for heptameric ssRNAs (**5b**, **6b**, **7b** and **8b**, see Fig. 1) for 1 h, and quenching by aqueous acetic acid (0.03 N). Each cleavage product after 1 h of digestion was separated as a pure component by RP-Hplc and SMART™ RP-Hplc analyses. Each of the pure isolated fragments was subsequently characterized by high resolution MALDI TOF mass measurements. Cleavage sites and the percentage cleavage are shown by solid arrows. Broken arrows show the pK_{a2} of those phosphates which are also shown in italics, and the blank arrows show the pK_{a1}. The rate constants (k) shown in bold are for the disappearance of the parent heptamer during the alkaline hydrolysis [see ESI 2 (Fig. S10–S15)]. Electrostatic interaction in the phosphate markers is propagated at the 3′-end (p_4–p_6) and at the 5′-end (p_1–p_3) as a result of $\mathbf{G}^-$ formation. The phosphate markers in italics (p_2, p_3 and p_4) show the pK_{a2} (Table 1) and are also deshielded due to the 9-guaninyl ionizations shown in the titration profile.

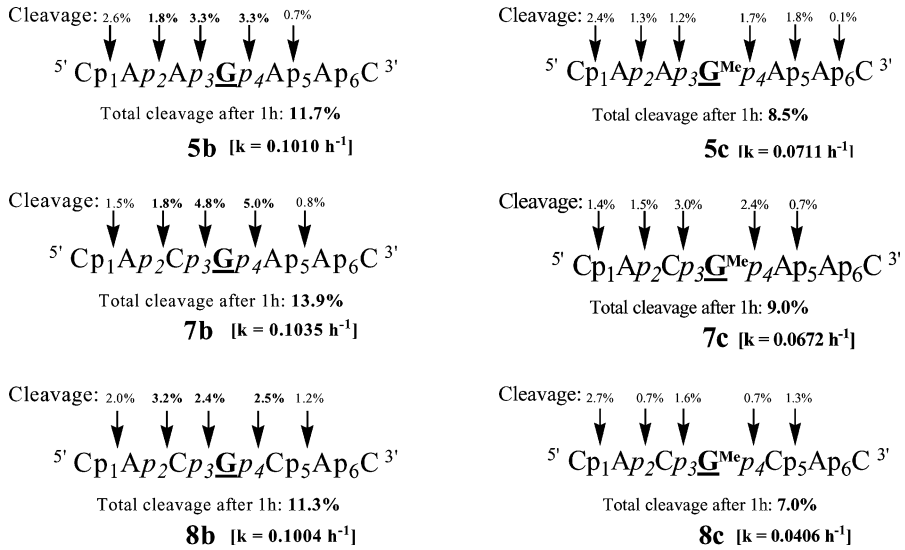

Fig. 4 Comparison of the alkaline hydrolysis at 1 h (in 0.03 N aqueous NaOH, pH 12.5, 20 °C) for native (**5b**, **7b**, **8b**) and $N^{1\text{-Me}}$-**G** (**G**^{Me})-containing heptameric ssRNAs (**5c**, **7c**, **8c**). See ESI 2 for the rate plots.

were the *primary alkaline hydrolysis products*. These primary cleavage products were subsequently found to produce secondary hydrolysis products as the alkaline hydrolysis progressed, and since this resulted in too many pathways hence the products could not be used for further time-dependent quantitation of the *primary alkaline hydrolysis products* {RP-Hplc results were monitored and compared up to 48 h for the native RNAs, and up to 27 h for $N^{1\text{-Me}}$-**G**-containing heptameric ssRNAs [see also ESI 2 (Fig. S12A); ESI 3 (Fig. S12B); and ESI 5 (Fig. S15A) to ESI 11]}. Each of the *primary alkaline hydrolysis products* (after 60 min of the alkaline digestion at 20 °C) formed from each of the heptameric ssRNAs **5b**, **6b**, **7b**, **8b** and the $N^{1\text{-Me}}$-**G**-containing **5c**, **7c**, **8c** was separated *as a pure entity* either by single-step RP-Hplc or by a two-step RP-Hplc/SMART™ RP-Hplc and quantitated {Fig. 3 and Fig. 4; ESI 2–ESI 11; Experimental sections (E)–(G) in this article; and ESI 1 [Experimental sections (E)–(G)]}. The high resolution MALDI TOF mass spectra in the negative ion mode are shown in ESI 4 (Fig. S13), and characterization and structural assignment of each of the pure cleavage products are shown in ESI 5 [Table S9 (A)–(G)]. Our key observations are given in the following list, (1) to (9).

(1) It has been found that the ssRNAs are degraded at different rates depending upon the sequence context: the total cleavage after one hour at 20 °C is 11.7% for **5b**, 7.8% for **6b**, 13.9% for **7b** and 11.3% for **8b** [time-dependent hydrolysis in the RP-Hplc elution profiles is shown in ESI 2 (Fig. S12A); ESI 3 (Fig. S12B); and ESI 5 (Fig. S15A) to ESI 11].

(2) In the case of the $N^{1\text{-Me}}$-**G** heptameric ssRNAs (**5c**, **7c** and **8c**), it has been observed that the **total** alkaline degradation occurs at a slower rate compared with that of the corresponding native sequences. For example, after one hour at 20 °C: (i) degradation of the $N^{1\text{-Me}}$-**G** heptamer **5c** is 8.5%, which is 27% less than the native **5b** (11.7%); (ii) degradation of the $N^{1\text{-Me}}$-**G** heptamer **7c** is 9.0%, which is 35% less than the native **7b** (13.9%); and (iii) degradation of the $N^{1\text{-Me}}$-**G** heptamer **8c** is 7.0%, which is 38% less than the native **8b** (11.3%) (Fig. 4).

(3) In all four native heptameric ssRNAs (**5b**, **6b**, **7b** and **8b**), alkaline hydrolysis is preferred to give the initial products at those internucleotide phosphates (p_2, p_3 and p_4, Fig. 3) which show both pK_{a2} and a weaker screening of the ^{31}P nucleus in alkaline pH compared to neutral conditions. Some other relatively minor cleavages were also observed at other internucleotide phosphates, which do not show a pK_{a2}, such as p_1 of Cp_1A (in all ssRNAs) as well as at p_5 of the Cp_5A block in **6b** and **8b**, and in that of the Ap_5A block in **5b** and **7b** (note, no cleavage is observed at p_6 of any of our native ssRNAs). The rates for the cleavage of the Cp_1A fragments[1d–f,3i–l] are quite comparable to or lower than the cleavage rates of the phosphodiester bonds (p_2, p_3 and p_4) which show pK_{a2} (Fig. 3). In contrast, the reactivities of the Cp_1A fragments are always higher than the cleavage rates of that of internucleotidic p_5, which does not show any pK_{a2}.

(4) The **preferential** cleavages found at the internucleotidic phosphates p_2, p_3 and p_4 (which also show pK_{a2}) in the native heptameric ssRNAs (**5b**, **7b** and **8b**), are 16–78% reduced in the case of the $N^{1\text{-Me}}$-**G**-containing RNAs (**5c**, **7c** and **8c**) because of the disappearance of the electrostatic effect of $\mathbf{G}^-$ (Fig. 4).

(5) In Fig. 5, the relative $\Delta\delta^{31}P$ shifts of the N^1-methylated analog **8c** signify the sole electrostatic interactions (modulated by the microenvironment) between the 2′-O^- and the vicinal 3′-phosphate, whereas for **8b** it is the triple interaction amongst three actors, **G**/$\mathbf{G}^-$, 2′-O^- and phosphate. Thus, the comparison of the alkaline cleavage rates at pH 12.5 (Fig. 4) and the relative $\Delta\delta^{31}P$ shifts between pH 11.6 and 12.5 (Fig. 5) in the native heptameric RNA **8b** with that of N^1-methylated analog **8c** show the following:

(i) The variable $\Delta\delta^{31}P$ shifts suggest that the pK_a values of all 2′-hydroxys are not uniform in **8c**, which is in contrast with the very similar pK_a values[3c,15] of the internucleotidic 2′-OH obtained for eight different diribonucleoside (3′→5′) phosphates (12.71 for ApG, 12.81 for ApA, 13.13 for GpG, 13.11 for GpA, 13.17 for CpG, 13.28 for CpA,

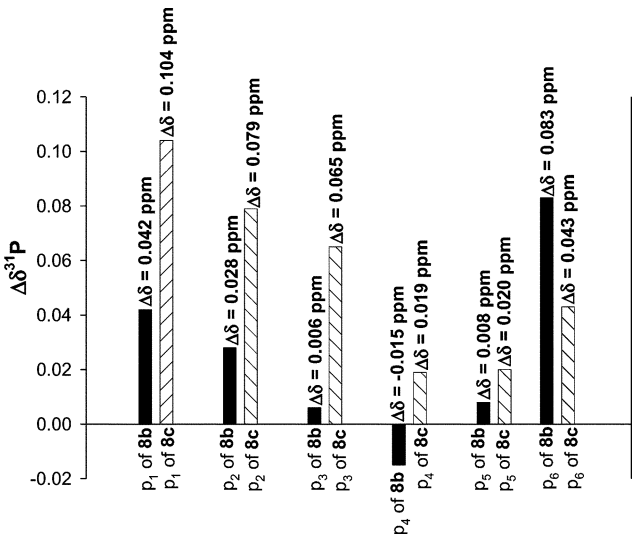

Fig. 5 Bar plots representing comparative $\Delta\delta^{31}$P shifts between pH values 11.6–12.5 for each of the internucleotidic p_1–p_6 phosphates of ($C^{5'}p_1A^{5'}p_2C^{5}p_3\underline{G}p_4C^{3}p_5A^{3'}p_6C^{3'}$) (**8b**) and ($C^{5'}p_1A^{5'}p_2C^{5}p_3\underline{G^{Me}}p_4C^{3}$-$p_5A^{3'}p_6C^{3'}$) (**8c**). The values for the $\Delta\delta^{31}$P shifts for each of the internucleotidic phosphate markers of **8b** and **8c** are shown on the top of each bar plot along the *y*-axis. Relative $\Delta\delta^{31}$P shifts suggest that all 2′-OH groups (except for p_6) are more ionized in **8c** compared with **8b**.

13.16 for UpG and 13.10 for UpA). The relative $\Delta\delta^{31}$P shifts of each of the internucleotidic phosphates suggests that the vicinal 2′-OH acidity in **8c** increases in the following order: $p_5 \approx p_4 < p_3 < p_2 < p_1$.

(ii) The relative cleavage rates at the internucleotidic phosphates cannot be correlated with the relative population of the corresponding vicinal 2′-oxyanion because of contributions from other factors (see Introduction). The percentile hydrolytic cleavage values at the internucleotidic p_2, p_3 and p_4 phosphates in **8c** are relatively small compared with those in **8b**, despite the fact that the vicinal 2′-oxyanion population is considerably higher in **8c** (as is evident from the $\Delta\delta^{31}$P shifts in Fig. 5). This suggests that the relatively high electrophilic character of the phosphates in **8b** (contributed to by both $\mathbf{G}^-$ and 2-O^- in the proximity) is perhaps more important for its higher rate of the alkaline cleavage reaction than its 2′-oxyanion population (compared with that of **8c**), keeping in view that all other cleavage requirements in **8b** and **8c** are perhaps very similar because of their closely similar sequence context.

(iii) It is evident from the relatively larger $\Delta\delta^{31}$P shifts for p_1, p_2, p_3, p_4 and p_5 in **8c** compared with those in **8b** that the 2′-OH groups in **8c** are relatively more acidic compared with **8b**, with the one exception of p_6. The reason that 2′-OH in **8b** is relatively poorly ionized compared with **8c** is because of the high energy penalty of the electrostatic repulsion between $\mathbf{G}^-$ and the 2′-oxyanion.

(6) **The 5′-terminal Cp_1A blocks in the ssRNAs are uniquely cleaved at a higher rate than the others.** Fig. 3 shows that the relative hydrolytic cleavage rates at the 5′-terminal Cp_1A moiety are comparable to or lower than the respective cleavage rates at the p_2, p_3 and p_4 internucleotidic phosphodiesters which show pK_{a2}. In contradistinction, an unusually high hydrolysis is found to take place at the internucleotidic phosphates of the 5′-terminal Cp_1A blocks in general (2.6% in **5b**, 2.1% in **6b**, 1.5% in **7b**, and 2.0% in **8b**), compared with the intrastrand phosphodiesters of the Cp_5A block of **6b** (0.1%) and **8b** (1.2%) or of the Ap_5A block of **5b** (0.7%) and **7b** (0.8%), which do not show any pK_{a2}.

Thus, the systematically higher cleavage rates of the Cp_1A fragments, outside the group of phosphates with pK_{a2}, are consistent (see below) with earlier observations[1d–f,3i–l] that the 5′-r(CpA)-3′ or 5′-r(UpA)-3′ blocks undergo cleavage that far exceeds other internucleotide linkages in the chimeric DNA/RNA or 2′-*O*-methylated-RNA/RNA oligonucleotides as well as in many natural RNA polymers.

(7) **The relationship between the internucleotidic cleavage rate and pK_{a2}.** The importance of the sequence context in RNA can be further understood by comparison of the cleavage rates at p_2, which shows pK_{a2} as the pH becomes alkaline: we observe (Fig. 3) that the Ap_2C fragment in **8b** (3.2%) cleaves at a much faster rate than that in **7b** (1.8%), whereas the relative cleavage rates of the Ap_2A fragment in **5b** (1.8%) and **6b** (*ca.* 2.0%) are comparable. The Ap_2C block cleaves more easily in **7b** (1.8%) or **8b** (3.2%), while the Ap_6C block, which does not show pK_{a2}, is completely resistant to any hydrolysis in all of our ssRNAs. This highly preferential cleavage of the Ap_2C block over the Ap_6C block in **7b** or **8b** is remarkable, assuming that the 2′-OH of the adeninyl nucleotide moieties in Ap_2C and Ap_6C have comparable pK_a values. Surprisingly, despite the fact that the internucleotide phosphate linkage of the Ap_2C block is capable of sampling fewer in-line cleavage configurations because of its more conformationally constrained intrastrand location than that of the internucleotidic phosphate of 3′-terminal Ap_6C (particularly, its α and β torsions can act as a free rotor), it clearly was not the decisive factor favoring the cleavage of the former over the latter. These suggest that the chemical character of the internucleotidic p_2 in **7b** or **8b**, which shows both pK_{a2} as well as weaker screening of their ^{31}P nucleus in the alkaline pH compared with that of neutral conditions, are very special, perhaps due to its enhanced electrophilic character compared with that of p_6 (see below for further evidence of enhanced electrophilic character of those phosphates with pK_{a2}).

On the other hand, a comparison of the chemical character/reactivity of the p_2 [the Ap_2C block in **7b** (1.8%) and **8b** (3.2%) and the Ap_2A block in **5b** (1.8%) or **6b** (2.0%)], also shows that there is most probably a complex set of stereoelectronic and conformational/dynamic factors[1d–f,3h–l] that are responsible in tandem for dictating the cleavage (see the Introduction), depending upon the sequence context and the resulting folding of the RNA molecule.

(8) **The electrostatic charge repulsion between the phosphate and the guanylate enhances the electrophilic character of the phosphate.** A comparison of the cleavage rates at p_2, p_3 and p_4 in all four ssRNAs shows (Fig. 3) that the cleavage at p_3 and p_4 far exceed those of the cleavage rate at p_2 in –Ap_2A$p_3\underline{G}p_4$A– in **5b** and –Ap_2C$p_3\underline{G}p_4$A– in **7b**, whereas the cleavage at p_2 is certainly preferred over p_3 and p_4 in –Ap_2C$p_3\underline{G}p_4$C– in **8b**. In contradistinction, the relative

cleavage rates at p_2, p_3 and p_4 in –Ap_2Ap_3Gp_4C– in **6b** are comparable. This shows the importance of the sequence context in a *competitive hydrolysis experiment with many reactive internucleotide phosphates* as in our ssRNAs, in which all internucleotide phosphates compete freely for stabilizing the stereoelectronic and conformational/dynamic hyperspace[3] as prerequisites for the general base-catalysed hydrolysis of RNA phosphodiesters.

Thus, our observations of high alkaline hydrolysis rates at the internucleotidic phosphates of p_2, p_3 and p_4 (which show pK_{a2}) compared with those which do not show pK_{a2} [such as those of the Ap_5A blocks in **5b**/**7b** and the Ap_6C block (in which no cleavage is observed in any of the four ssRNAs), along with the known exception[1d–f,3i–l] of high cleavage rates of CpA dimers as in the Cp_1A and Cp_5A blocks] clearly suggest that the enhancement of the electrophilic character of the phosphorus of those specific internucleotide phosphates with pK_{a2} can be responsible for a highly preferential hydrolysis rate of the intrastrand p_2, p_3 and p_4 (see below for the mechanism).

Earlier alkaline hydrolysis studies[1d–g,3b–j] were performed with chimeric DNA/RNA or 2′-*O*-methylated-RNA/RNA oligonucleotides, in which only **one** scissile RNA phosphodiester was incorporated. These studies showed that the alkaline transesterification reactivity of 5′-*Up*A-3′ and 5′-*Cp*A bonds (the ribo-phosphodiester bond is shown in italic) were definitely much larger than those of the 5′-*Ap*A-3′, 5′-*Gp*A-3′, 5′-*Gp*G-3′ or 5′-*Ap*G-3′ moieties. When we, however, compare these preferences for the cleavage of the 5′-pyrimidine-purine-3′ phosphodiester bonds in the chimeric oligonucleotides with all the *six freely competing ribo-phosphodiester bonds* in our heptameric ssRNAs, we clearly note that the 5′-purine-purine-3′ phosphodiester bonds are **not** more resistant to cleavage compared with those of 5′-pyrimidine-purine-3′ phosphodiester bonds, as earlier elucidated for **one** scissile RNA phosphodiester bond using chimers. In fact, with the exceptions of Ap_5A (0.7% cleavage) in **5b** and Ap_5A (0.8% cleavage) in **7b**, all other 5′-purine-purine-3′ phosphodiester bonds are either more reactive [Ap_3G (3.3%), Gp_4A (3.3%) in **5b** and Gp_4A (5.0%) in **7b**] or are very comparable in reactivity [Ap_2A (2.0%) and Ap_3G (2.0%) in **6b**] with that of the 5′-pyrimidine-purine-3′, Cp_1A, phosphodiester bond [(2.6%) in **5b**, (2.1%) in **6b**, and (1.5%) in **7b**]. In the case of **6b** and **8b**, the 5′-purine-purine-3′ phosphodiester bonds are clearly more reactive than that of the 5′-pyrimidine-purine-3′, Cp_5A phosphodiester bond.

(9) **A general mechanism explaining the enhancement of the electrophilic character of some specific internucleotidic phosphates.** It is well known that the chemical shift is dictated by the screening of a nucleus,[9b] which in turn is directly correlated to the diamagnetic shielding of the neighboring electrons (*i.e.* a function of the electron density). This would normally mean that the phosphate ionization to an anion should shield the phosphate to a higher field as it happens for proton chemical shift change as the electron density around it increases. However, it is well known, in general, for various types of phosphates,[11a–h] phosphonates[11a,b] and aminophosphonates[11b] that the chemical shift goes downfield in the alkaline pH compared with those under neutral conditions. This is also true with some of the phosphates of the heptameric RNAs [Fig. 2, and ESI 1 (Fig. S1 and Table S6)], which also show pK_{a2}: as the pH increases, the ^{31}P chemical shift of those adjoining phosphates goes downfield [see ESI 1 (Table S5)], showing a typical titration curve for the internucleotidic p_2, p_3 and p_4 (Fig. 2) as a result of the systematic charge repulsion effect between the electron cloud in the outermost orbitals of phosphorus and the central 9-guanylate ion. As a result, the excess negative charge accumulated around the phosphorus nucleus is delocalized into its own dπ orbitals through pπ–dπ orbital overlap, causing the O–P–O bond angle to decrease as a manifestation of the p-orbital expansion (an increased p-orbital unbalancing term).[11a] Thus, as the guanylate ion is generated at pH 12.5, the electrostatic repulsion between the phosphorus electrons and the 9-guanylate ion results in a weaker ^{31}P screening due to delocalization of the charge into the phosphorus dπ orbitals. This weaker screening of the ^{31}P nucleus is mostly felt at p_2, p_3 and p_4, as demonstrated here by the downfield shift (deshielding) of δ^{31}P [Fig. 2, and ESI 1 (Fig. S1 and Table S6)] as the population of the guanylate ion increases in the pH-dependent electrostatic titration studies from pH 7 to 12.5 (Fig. 2 and Fig. 3). The chemical manifestation of this *weaker screening* of the ^{31}P nucleus (*i.e.* the systematic downfield shift of δ^{31}P from pH 7 to 12.5 as a result of deshielding of the ^{31}P nucleus) upon generation of the 9-guanylate ion is an enhancement of the electrophilic character of those phosphates showing pK_{a2} and those more electrophilic phosphates which are p_2, p_3 and p_4 (Fig. 3). Hence, their alkaline transesterification rates are higher than others.

We considered that the best possible way to establish this enhanced electrophilicity in the internucleotide phosphates of p_2, p_3 and p_4 is to compare the alkaline cleavage rates of identical dinucleotide blocks constituting p_1, p_5 and p_6. For this comparison we ruled out the relative cleavage rates of the Cp_1A and Cp_5A blocks because they have unusual cleavage rates[1d–f,3i–l] as discussed above, but we can safely compare the rest. Thus, a comparison of the relative rates of cleavage at Ap_2A (1.8%) *versus* Ap_5A (0.7%) in **5b**, Ap_2C (1.8%) *versus* Ap_6C (no cleavage) in **7b**, and Ap_2C (3.2%) *versus* Ap_6C (no cleavage) in **8b** show that the cleavage rate is much higher at p_2 than at p_5 or p_6 (Fig. 3).

Conclusions and implications

(1) It is found that, unlike the trimeric ssDNAs/ssRNAs and heptameric ssDNAs, the electronic nature of some of the phosphates in the heptameric ssRNAs is dissimilar (non-equivalent) in a sequence-specific manner, as probed by titrating the intramolecular electrostatic interactions of the **G** to **G**$^-$ transition (pK_{a1}) with each of the internucleotidic phosphates (pK_{a2}) across the ssRNA chain by the pH-dependent ^{31}P chemical shift study. This is most probably due to the specific folded nature of each of these heptameric ssRNAs, which makes the chemical environment around each of their phosphates sufficiently different to exhibit a sequence-dependent chemical reactivity toward alkaline hydrolysis.

(2) The specific molecular microenvironment, due to a specific sequence context in the ssRNA, has been known[1d–f,3h–l] to accelerate the alkaline cleavage rate of a scissile phosphate. Molecular dynamics simulation has also shown that a specific base sequence can also induce a higher stability of the penta-coordinated phosphorane structure, mimicking that of the transition state of the scissile phosphodiester bond.[1e,f] On the other hand, the present work showed that, owing to a specific sequence context in the ssRNA, the charge densities around each of the internucleotidic phosphates can vary due to the electrostatic influence of an electron-rich center that is in close steric proximity. This dictates the electrophilic character that is available for the different internucleotidic phosphates toward the alkaline hydrolytic cleavage reaction. In the heptameric RNAs this variable electrophilic character of the phosphates results from a systematic electrostatic charge repulsion effect between the electron cloud in the outermost orbitals of the phosphorus and the central guanylate ion as the pH becomes gradually more alkaline, leading to subsequent delocalization of the phosphorus pπ charge into its dπ orbitals. The net effect of this delocalization of the pπ charge into the dπ orbitals is a weaker ^{31}P screening, which is evidenced by systematic deshielding of the δ^{31}P shift as the pH becomes more alkaline, resulting in a neat titration curve with an inflection point. Hence, those adjoining phosphates that are in stereochemical proximity around an electron-rich center show pK_{a2} and a resulting higher electrophilicity; they are also, therefore, relatively more susceptible to the spontaneous alkaline hydrolysis by transesterification than those which are further away.

It is known that some nucleobases in large, folded RNA show a perturbed[6] pK_a because of their orientation in a more hydrophobic microenvironment. These perturbed nucleobases normally show more basic pK_a [2a,12a] than those of the mononucleotide or any other unfolded nucleobase(s) within a given RNA sequence. Thus this change of electron density can also enhance the electrophilicity of the phosphodiester bonds within the steric proximity, simply because of the enhanced electrostatic repulsion mechanism between them and the neighboring phosphates. Thus, this present study constitutes a simple model for understanding how the modulation of a specific sequence context, owing to its microenvironment, can dictate the reactivity of the interenucleotidic phosphates within the steric proximity.

(3) It may be possible to modulate the chemical reactivity of the internucleotidic phosphate(s) by complexing a specific nucleobase (mainly at N^1 of G or at N^3 of U) with a metal ion, such as Hg(II), which will enhance its charge density and thereby influence the electrophilicity of the neighboring phosphate(s) as a result of weaker screening of the ^{31}P nucleus. The specific base-catalysis by RNA-cleaving proteins, such as RNase A or RNA phosphodiesterase or nuclease, of a specific phosphate hydrolysis in RNA (often utilizing histidine residues for deprotonation), just as in the present non-enzymatic hydrolysis, will most probably be electrostatically influenced in a similar manner. Therefore, the present set of sequence-specific alkaline hydrolysis experiments serves as a good model to understanding the influence of electrostatics in the modulation of chemical reactivity of RNA in general.

(4) The modulation of phosphate charge density by electrostatic means (inter- or intra-molecularly) may also enhance the pK_a of the scissile phosphate slightly (because of its steric location in a more hydrophobic environment compared with the others). This will make it partially protonated *at the physiological pH* more preferentially than the others, and as a result the electrophilicity of the internucleotidyl scissile phosphate is very likely to be distinctly different from the others in a large polymeric RNA. Consequently, those protonated phosphates will have a poorer ability to repel nucleophiles such the 2′-oxyanion or the hydroxide, causing specific transesterification to take place in a preferential manner.

(5) The percentile cleavage values at the internucleotidic p_2, p_3 and p_4 phosphates are less in **8c** compared with **8b**, despite the fact that the vicinal 2′-oxyanion population is considerably higher in **8c** (as evident from the $\Delta\delta^{31}$P shifts in the pH values between 11.6 and 12.5 – Fig. 5). This suggests that the relatively high electrophilic character of the phosphates in **8b** (because of **G**$^-$ in the proximity) is perhaps more important for its higher rate of the alkaline cleavage reaction than its 2′-oxyanion population, bearing in mind that the 'in-line' cleavage conformations in **8b** and **8c** are perhaps very similar.

(6) The earlier observations on the relative rates of the alkali-promoted cleavage reaction of the RNA strand[1a–f,3b,h–l] *versus* the RNA sequence context goes hand-in-hand with our present observation of the *sequence-dependent electrostatic modulation* of the phosphate charge densities as well as the propensity to spontaneous hydrolysis of various internucleotide phosphates under the alkaline conditions.

(7) The comparative alkaline hydrolysis in all four heptameric ssRNAs (**5b**–**8b**) clearly demonstrates the complexity of the physico-chemical behavior of a particular sequence context (and the role of the resulting local structure) in dictating the preferences in the alkaline hydrolytic cleavage reaction amongst the six competing internucleotidic ribo-phosphates (compared with the hydrolysis with those in the chimeras with **one** scissile RNA phosphodiester). This might in turn help us to interpret the mechanism employed by ribonucleases and by RNA-cleaving ribozymes and to use the knowledge for the design of appropriate enzyme mimics.

(8) Incorrect splicing is known to have an important biological role in the development of various diseases. It is likely that if a specific fold or scaffold building in a large pre-mRNA stereochemically places a particular phosphate in a hydrophobic microenvironment, it becomes more electrophilic than the rest. By doing so it assumes a greater transesterification potential, causing a wrong splicing to take place. Thus, by modulating the folding pattern or varying the hydrophobic pockets by engineering the appropriate interactions, one can perhaps re-tune the phosphate reactivity in such a way that the correct splicing is re-instated.

Experimental

(A) pH-Dependent ^{31}P NMR measurement

The pH-dependent ^{31}P chemical shifts for the trimeric ssDNA and ssRNA compounds **1**–**4** as well as the heptameric ssDNA and

ssRNA compounds **5**–**8** were performed using Bruker DRX-500 and DRX-600 spectrometers at 298 K in D_2O solution (1 mM) using $O{=}P(OMe)_3$ ($\delta = 0.0$ ppm) as an external standard for ^{31}P chemical shifts. All oligos were characterized using ^{31}P-decoupled ^{1}H COSY, ^{1}H TOCSY, ^{31}P–^{1}H correlation spectroscopy and ^{1}H NOESY. The pH calibration and measurement procedure is given in ESI 1.†

(B) Accuracy of δ^{31}P chemical shifts in compounds 1–8

The error in $\delta^{31}P$ can come from the following three sources: (1) the error from the digital resolution of the NMR spectrometer at 600 MHz, (2) the error from the line broadening of the ^{31}P resonance, and (3) the error from the salt effect.

(C) pH Titration using ^{31}P markers in compounds 1–8

These have been discussed in the text and are also discussed in ESI 1.†

(D) pK_a Determination

The pH-dependent $\delta^{31}P$ plots (over the range of pH 6.6–12.5, with an interval of pH 0.2–0.3) for **1**–**8** show sigmoidal behavior [see Fig. 2, and ESI 1 (Fig. S1–S3)†] for the deprotonation of the 9-guaninyl moiety in the trimeric and heptameric ssDNAs (**1a**–**8a**) and ssRNAs (**1b**–**8b**). In some ^{31}P resonances of the ssRNA sequences (p_2 and p_4 of **6b** and p_4 of **7b**) the pH-dependent $\delta^{31}P$ plots do not reach a distinct plateau at the end of the deprotonation of 9-guaninyl. In other ^{31}P resonances of heptameric ssRNA sequences a plateau is reached for the deprotonation of 9-guaninyl, but the $\delta^{31}P$ shift at even higher pH values shows deviation from the plateau. This observation is a result of the influence on the phosphorus chemical shifts of the ionization of the vicinal 2′-OH group in the ssRNAs, which starts before complete deprotonation of 9-guaninyl in cases where a plateau is not reached and after complete deprotonation in cases where the plateau is reached [see ESI 1 (Experimental section D) for a detailed discussion on the influence of the ionization of the vicinal 2′-OH group on the $\delta^{31}P$ values of the phosphate markers in **5b**, **6b**, **7b** and **8b**].

(E) Alkaline hydrolysis of the heptameric ssRNAs

To the lyophilized solid oligonucleotides (5 od at $\lambda_{260\,nm}$) in an Eppendorf tube, aqueous sodium hydroxide (100 μL, 0.03 N, pH 12.5) was added. It was allowed to stand at room temperature (20 °C) until the last time point was taken at 48 h. Aliquots of 10 μL, each containing 0.5 od, were removed at suitable time intervals and immediately quenched with aqueous acetic acid (10 μL, 0.03N) to *ca.* pH 7. This was then stored at −20 °C until analysis by Hplc. Each peak in the Hplc elution profile of the alkaline hydrolysis of compounds **5b**, **6b**, **7b**, **8b**, **5c**, **7c**, and **8c** for 1 h were characterized by MALDI TOF and the elution profiles at zero and 1 h hydrolysis were used for quantification of the primary products formed after 1 h of alkaline digestion. All elution profiles at different time points ($\frac{1}{2}$, 1, 2, 3, 4, 8, 15, 27 and 48 h) were, however, used in determining the degradation profile and the degradation rate constants for the heptamers. For a detailed discussion on the Hplc separation of the hydrolysis products at different time intervals, a procedure for the separation of each primary product as a pure component after 1 h of alkaline hydrolysis, and a procedure for the MALDI TOF analysis, see the Experimental section in ESI 1.†

(F) Calculation of rate constants of the alkaline hydrolysis

The pseudo first-order rate constants were determined by plotting the natural log of the fraction of the heptamer remaining uncleaved, ln (% heptamer left), at various incubation times, considering the heptamer peak as 100% area at zero time. All of the four ssRNA heptamers were contaminated by a small amount of both non-nucleot(s)idic impurities (at $R_T \approx 6.6$ and $\approx 10.5'$) as well as by nucleotidic impurities which have been defined for quantitation purposes [ESI 1 (Experimental section F)†]. Hence, depending upon the retention times [ESI 1 (Experimental section F)], all peaks formed upon degradation were recalculated after subtracting the impurities (taking the parent heptamer peak as 100% before the addition of alkali).

The percentage areas of each of the degradation product peaks at 1 h of alkaline digestion in the RP-Hplc elution profile of each ssRNA were corrected according to the purity defined for each parent heptameric ssRNA [see ESI 1 (Experimental section)]. Some of the single peaks in the first round of RP-Hplc were found to contain more than one hydrolysis product (nucleotide fragments) when analyzed by MALDI TOF mass spectrometry. They were further analyzed using the SMART™ RP-Hplc micro purification system. Those separated pure components/peaks were characterized again by MALDI TOF spectrometry and were subsequently used for extinction coefficient correction to determine the actual contribution of the different components in the percentage cleavage shown in Fig. 3 and Fig. 4 of the main text [see also ESI 2 (Fig. S12A); ESI 3 (Fig. S12B), and ESI 4 (Fig. S13)]. For detailed calculation procedures see ESI 1 (Experimental section J) and ESI 5 (Fig. S14).

Acknowledgements

Generous financial support from the Swedish Natural Science Research Council (Vetenskapsrådet), the Swedish Strategic Research Foundation (Stiftelsen för Strategisk Forskning), theEuropean Union Biotechnology Program (LSHB-CT-2004-005276, Contract No. 005276) and Philip Morris USA Inc, the Wallenberg Consortium North (to Å. E.) is gratefully acknowledged.

References and notes

1 (*a*) D. M. Brown, D. I. Magrath and A. H. Neilson, *Nature*, 1956, **177**, 1124; (*b*) M. Oivanen, R. Schnell, W. Pfleiderer and H. Lönnberg, *J. Org. Chem.*, 1991, **56**, 3623; (*c*) P. Jarvinen, M. Oivanen and H. Lonnberg, *J. Org. Chem.*, 1991, **56**, 5396; (*d*) U. Kaukinen, S. Lyytikäinen, S. Mikkola and H. Lönnberg, *Nucleic Acids Res.*, 2002, **30**, 468; (*e*) U. Kaukinen, T. Venäläinen, H. Lönnberg and M. Peräkylä, *Org. Biomol. Chem.*, 2003, **1**, 2439; (*f*) U. Kaukinen, H. Lönnberg and M. Peräkylä, *Org. Biomol. Chem.*, 2004, **2**, 66; (*g*) K. Kawamura, *Bull. Chem. Soc. Jpn.*, 2003, **76**, 153.

2 (*a*) A. R. Fersht, *Structure and Mechanism in Protein Sciences*, W. H. Freeman & Co, San Francisco, 1999; (*b*) F. M Richards and H. W. Wyckoff, in *The Enzymes*, P. D. Boyer, ed., Academic Press, New York, 3rd edn., 1971, vol. iv, p. 647.

3 (*a*) P. D. Lyne and M. Karplus, *J. Am. Chem. Soc.*, 2000, **122**, 166; (*b*) Yi. Li and R. R. Breaker, *J. Am. Chem. Soc.*, 1999, **121**, 5364; (*c*) S. Acharya, A Földesi and J. Chattopadhyaya, *J. Org. Chem.*, 2003, **68**, 1906; (*d*) for a review on RNA hydrolysis, see: D. M. Perreault and

E. V. Anslyn, *Angew. Chem., Int. Ed. Engl.*, 1997, **36**, 432; (*e*) for a review on RNA hydrolysis under acidic and basic conditions and a detailed kinetic analysis, see: M. Oivanen, S. Kuusela and H. Lönnberg, *Chem. Rev.*, 1998, **98**, 961; (*f*) G. A. Soukup and R. R. Breaker, *RNA*, 1999, **5**, 1308; (*g*) B. G. Lane and G. C. Butler, *Biochim. Biophys. Acta*, 1959, **33**, 281; (*h*) R. Kierzek, *Nucleic Acids Res.*, 1992, **20**, 5079; (*i*) R. Kierzek, *Methods Enzymol.*, 2001, **541**, 657; (*j*) I. Zagorowska, S. Mikkola and H. Lonnberg, *Helv. Chim. Acta*, 1999, **82**, 2105; (*k*) K. Williams, P Ciafré and G. P. Tocchini-valentine, *EMBO J.*, 1995, **14**, 4551; (*l*) A. Bibillo, M. Figlerowicz and R. Kierzek, *Nucleic Acids Res.*, 1999, **27**, 3931.

4 T. R. Cech and B. L. Golden, in *The RNA World*, R. F. Gesteland, T. R. Cech and J. F. Atkins, eds., Cold Spring Harbor Laboratory Press, New York, 2nd edn., 1999, p. 321 and p. 525.

5 For reviews, see: J. R. Williamson and M. J. Fedor, *Nat. Rev. Mol. Cell Biol.*, 2005, **6**, 399; Y. Takagi, Y. Ikeda and K. Taira, *Top. Curr. Chem.*, 2004, **232**, 213; P. C. Bevilacqua, T. S. Brown, S. Nakano and R. Yajima, *Biopolymers*, 2004, **73**, 90; D. Lilley, *Trends Biochem. Sci.*, 2003, **9**, 495; V. J. DeRose, *Chem. Biol.*, 2002, **9**, 961; R. Schroeder, R. Grossberger, A. Pichler and C. Waldsich, *Curr. Opin. Struct. Biol.*, 2002, **12**, 296; M. J. Fedor, *Curr. Opin. Struct. Biol.*, 2002, **12**, 289; J. A. Doudna and T. R. Cech, *Nature*, 2002, **418**, 222.

6 I. H. Shih and M. D. Been, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 1489; S. Nakano, D. Chadalavada and P. Bevilacqua, *Science*, 2000, **287**, 1493; A. T. Perrotta, I. Shih and M. D. Been, *Science*, 1998, **286**, 123; A. R. A. Lupták, K. Ferré-D'Amaré, K. W. Zhou, J. A. Zilm and Doudna, *J. Am. Chem. Soc.*, 2001, **123**, 8447; P. Legault and A. Pardi, *J. Am. Chem. Soc.*, 1997, **119**, 6621; S. Ravindranathan, S. E. Butcher and J. Feigon, *Biochemistry*, 2000, **39**, 16 026; Z. Cai and I. Tinoco, Jr., *Biochemistry*, 1996, **35**, 6026; A. C. Drohat and J. T. Stivers, *J. Am. Chem. Soc.*, 2000, **122**, 1840; A. C. Drohat and J. T. Stivers, *Biochemistry*, 2000, **39**, 11 865.

7 (*a*) M. A. Rould, J. J. Perona, D. Soll and T. A. Steitz, *Science*, 1989, **246**, 1135; (*b*) A. A. Antson, *Curr. Opin. Struct. Biol.*, 2000, **10**, 87; (*c*) N. Handa, O. Nureki, K. Kuriomoto, I. Kim, H. Sakamoto, Y. Shimura, Y. Muto and S. Yokoyama, *Nature*, 1999, **398**, 579; (*d*) S. R. Price, P. R. Evans and K. Nagai, *Nature*, 1998, **394**, 645.

8 (*a*) S. Acharya, J. Barman, P. Cheruku, S. Chatterjee, P. Acharya, J. Isaksson and J. Chattopadhyaya, *J. Am. Chem. Soc.*, 2004, **126**, 8674; (*b*) P. Acharya, S. Acharya, P. Cheruku, N. V. Amirkhanov, A. Földesi and J. Chattopadhyaya, *J. Am. Chem. Soc.*, 2003, **125**, 9948; (*c*) S. Acharya, P. Acharya, A. Földesi and J. Chattopadhyaya, *J. Am. Chem. Soc.*, 2002, **124**, 13 722; (*d*) P. Acharya, S. Acharya, A. Földesi and J. Chattopadhyaya, *J. Am. Chem. Soc.*, 2003, **125**, 2094.

9 (*a*) W. D. Kumler and J. J. Eiler, *J. Am. Chem. Soc.*, 1943, **65**, 2355; (*b*) P. J. Cozzone and O. Jardetzky, *Biochemistry*, 1976, **15**, 4853; (*c*) S. Chamberlin, E. J. Merino and K. M. Weeks, *Proc. Natl. Acad. Sci. U. S. A*, .2002, **99**, 14 688.

10 Y. H. Kao, C. A. Fitch, S. Bhattacharya, C. J. Sarkisian, J. T. J. Lecomte and B. E. García-Moreno, *Biophys. J.*, 2000, **79**, 1637.

11 (*a*) D. G. Gorenstein, in *Phosphorous-31 NMR: Principles and Applications*, D. G. Gorenstein, ed., Academic Press, Orlando, FL, 1984, pp. 7–36; (*b*) K. Moedritzer, *Inorg. Chem.*, 1967, **6**, 936; (*c*) S. Pietri, M. Miollan, S. Martel, F. L. Moigne, B. Blaive and M. Culcasi, *J. Biol. Chem.*, 2000, **275**, 19 505; (*d*) R. A. Y. Jones and A. R. Katritzky, *J. Inorg. Nucl. Chem.*, 1960, **15**, 193; (*e*) M. M. Crutchfield, C. F. Callis, R. R. Irani and G. C. Roth, *Inorg. Chem.*, 1962, **1**, 831; (*f*) M. Cohn and T. R. Jr. Hughes, *J. Biol. Chem.*, 1960, **237**, 3250; (*g*) M. Blumenstein and M. A. Raftery, *Biochemistry*, 1972, **11**, 1643; (*h*) R. B. Moon and J. H. Richards, *J. Biol. Chem.*, 1973, **248**, 7276; (*i*) M. A. Porubcan, W. M. Westler, I. B. Ibañez and J. L. Markley, *Biochemistry*, 1979, **18**, 4108; (*j*) A. C. van der Drift, H. C. Beck, W. H. Dekker, A. G. Hulst and E. R. J. Wils, *Biochemistry*, 1985, **24**, 6894.

12 (*a*) S.-O. Shan and D. Herschlag, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 14 474; (*b*) K. Langsetmo, J. A. Fuchs and C. Woodward, *Biochemistry*, 1991, **30**, 7603; (*c*) H. A. Chen, M. Pfuhl, M. S. B. McAlister and P. C. Driscoll, *Biochemistry*, 2000, **39**, 6814; (*d*) W. Schaller and A. D. Robertson, *Biochemistry*, 1995, **34**, 4714; (*e*) K. Bartik, C. Redfield and C. M. Dobson, *Biophys. J.*, 1994, **66**, 1180; (*f*) J. Song, M. Laskowski, M. A. Qasim and J. L. Markley, *Biochemistry*, 2003, **42**, 2847; (*g*) W. R. Forsyth, M. K. Gilson, J. Antosiewicz, O. R. Jaren and A. D. Robertson, *Biochemistry*, 1998, **37**, 8643; (*h*) M. D. Joshi, A. Hedberg and L. P. McIntosh, *Protein Sci.*, 1997, **6**, 2667; (*i*) W. R. Baker and A. Kintanar, *Arch. Biochem. Biophys.*, 1996, **327**, 189; (*j*) V. Dillet, H. J. Dyson and D. Bashford, *Biochemistry*, 1998, **37**, 10 298; (*k*) R. Consonni, I. Arosio, B. Belloni, F. Fogolari, P. Fusi, E. Shehi and L. Zetta, *Biochemistry*, 2003, **42**, 1421.

13 H. P. M. de Leeuw, C. A. G. Haasnoot and C. Altona, *Isr. J. Chem.*, 1980, **20**, 108; C. Thibaudeau and J. Chattopadhyaya, *Stereoelectronic Effects in Nucleosides and Nucleotides and their Structural Implications*, Department of Bioorganic Chemistry, Uppsala University Press,Sweden, 1999 (ISBN 91-506-1351-0; jyoti@boc.uu.se); W. Saenger, *Principles of Nucleic Acid Structure*, Springer Verlag, Berlin, 1988.

14 C. R. Cantor, M. M. Warshaw and H Shapiro, *Biopolymers*, 1970, **9**, 1059; *Handbook of Biochemistry and Molecular Biology, Volume 1: Nucleic Acids*, G. D. Fasman, ed., CRC Press, FL, USA, 3rd edn., 1975, p. 589; M. J. Cavaluzzi and P. N. Borer, 'Revised UV extinction coefficients for nucleoside-5′-monophosphates and unpaired DNA and RNA', *Nucleic Acids Res.*, 2004, **32**(1), e13.

15 The pK_a values of internucleotidic 2′-OH in ApG, ApA, CpA, CpG, UpG and UpA, reported from this lab (see ref. 3*c*) have been recalculated because of an error in defining the pH values. Since we have used titration with a known strength of H_2SO_4 all along from pH 11.1 to 14 without any readout from the electrode, there is no need to deduct 0.4 pH units from corresponding pD values to get the pH value. Hence, we performed this recalculation, and thus the pK_a values of the internucleotidic 2′-OH groups for the following dimers are 0.4 pK_a units higher than we have reported in past. The newly calculated values are as follows: ApG (12.71), ApA (12.81), CpA (13.28), CpG (13.17), UpG (13.16) and UpA (13.10).