Electronic Supplementary Information 1

Non-identical Electronic Characters of the Internucleotidic Phosphates in RNA Modulate the Chemical Reactivity of the Phosphodiester Bonds

Jharna Barman[1], Sandipta Acharya[1], Chuanzheng Zhou[1], Subhrangsu Chatterjee[1], Åke Engström[2], and Jyoti Chattopadhyaya[1]*

[1]Department of Bioorganic Chemistry, Box 581, Biomedical Center, Uppsala University, S-751 23 Uppsala, Sweden

[2]Department of Medical Biochemistry and Microbiology, Box 582, Biomedical Center, Uppsala University, S-751 23 Uppsala, Sweden

jyoti@boc.uu.se

Table of Content for all Electronic Supplementary Information:

Electronic Supplementary Information 1

Experimental Section A – J Details of experimental procedures for pK_a and rate determination as well as procedures for calculation of pK_a and rate constants. **p. S7 – 20**

Notes 1 – 4 Explanatory notes. **p. S20 – 24**

Figure S1 Panels **1a.1** – **8a.6** show the plots of pH-dependent (6.6 – 12.75) ^{31}P chemical shifts (δ^{31}P) for different phosphate markers of the trimeric (**1a – 4a**) and heptameric (**5a – 8a**) ssDNA as well as trimeric (**1b – 4b**) and heptameric (**5b – 7b**) ssRNA. The δ^{31}P are corrected for salt effect by subtracting the average δ^{31}P of the 3' and 5' phosphorus (p_1 and p_2 respectively) of Etp_1Ap_2Et (**9b**) from the δ^{31}P of different marker phosphates of the oligomers (**1 – 8**). **p. S25 - 30**

Figure S2. Panels **1a.3** – **8b.9** show the plots of pH-dependent (6.6 – 12.5) ^{31}P chemical shifts (δ^{31}P) for different phosphate markers of the trimeric (**1a – 4a**) and heptameric (**5a – 8a**) ssDNA as well as trimeric (**1b – 4b**) and heptameric (**5b – 8b**) ssRNA using experimental ^{31}P shifts as well as theoretically determined δ^{31}P at alkaline pH in the plot. For theoretical derivation of δ^{31}P at alkaline pH data points in the pH values 10 – 11.6 have only been used. The δ^{31}P are corrected for salt effect by subtracting the average δ^{31}P of the 3' and 5' phosphorus (p_1 and p_2 respectively) of Etp_1Ap_2Et (**9b**) from the δ^{31}P of different marker phosphates of the oligomers (**1 – 8**). **p. S31 – 34**

Figure S3. Panels **5a.7 – 8b.11** show the plots of pH-dependent (6.6 – 12.5) ^{31}P chemical shifts (δ^{31}P) for different phosphate markers of heptameric (**5a – 8a**) ssDNA as well as heptameric (**5b – 8b**) ssRNA using experimental ^{31}P shifts as well as theoretically determined δ^{31}P at alkaline pH in the plot. For theoretical derivation of δ^{31}P at alkaline pH data points in the pH values 10 – 11.9 have been used. The δ^{31}P are corrected for salt effect by subtracting the average δ^{31}P of the 3' and 5' phosphorus (p_1 and p_2 respectively) of Etp_1Ap_2Et (**9b**) from the δ^{31}P of different marker phosphates of the oligomers (**1 – 8**). **p. S35 - 36**

Figure S4. Panels **1a.5** – **8b.14** show the Graphical determination of ^{31}P chemical shift (δ_h) at the plateau at high pH for for ssDNA **1a – 4a** and **5a – 8a** as well as ssRNA **1b – 4b** and **5b – 8b**. The data points between pH 10 – 11.6 have been used in the calculation. The δ^{31}P are corrected for salt effect by subtracting the average δ^{31}P of the 3' and 5' phosphorus (p_1 and p_2 respectively) of Etp_1Ap_2Et (**9b**) from the δ^{31}P of different marker phosphates of the oligomers (**1 – 8**). **p. S37 - 39**

Figure S5. Panels **5a.9** – **8b.14** show the Graphical determination of ^{31}P chemical shift (δ_h) at the plateau at high pH for for heptameric ssDNA **5a – 8a** as well as heptameric ssRNA **5b – 8b**. The data points between pH 10 – 11.9 have been used in the calculation. The δ^{31}P are corrected for salt effect by subtracting the average δ^{31}P of the 3' and 5' phosphorus (p_1 and p_2 respectively) of Etp_1Ap_2Et (**9b**) from the δ^{31}P of different marker phosphates of the oligomers (**1 – 8**).**p. S40**

Figure S6. Panels **1a.7** – **8b.17** show the Hill plots for ssDNA **1a – 4a** and **5a – 8a** as well as ssRNA **1b – 4b** and **5b – 8b**. Data points between pH 10 – 11.6 have been used in the calculation. The δ^{31}P are

corrected for salt effect by subtracting the average δ^{31}P of the 3' and 5' phosphorus (p_1 and p_2 respectively) of Etp_1Ap_2Et (**9b**) from the δ^{31}P of different marker phosphates of the oligomers (**1 – 8**).**p. S41 - 43**

Figure S7. Panels **5a.11** – **8b.18** show the Hill plots for heptameric ssDNA **5a – 8a** as well as heptameric ssRNA **5b – 8b**. Data points between pH 10 – 11.9 have been used in the calculation. The δ^{31}P are corrected for salt effect by subtracting the average δ^{31}P of the 3' and 5' phosphorus (p_1 and p_2 respectively) of Etp_1Ap_2Et (**9b**) from the δ^{31}P of different marker phosphates of the oligomers (**1 – 8**). **p. S44 – S45**

Figure S8. Panels **(A) – (P)** show the stack plots of pH-dependent ^{31}P chemical shifts (in D_2O) of the ^{31}P phosphate markers for isosequential ssDNA (**1a – 4a** and **5a – 8a**) and ssRNA (**1b – 4b** and **5b – 8b**) at 298 K. **p. S46 – S62**

Figure S9. Panels **(A) – (H)** show the stack plots of pH-dependent ^{1}H chemical shifts (in D_2O) of the aromatic and anomeric markers protons for isosequential ssDNA (**5a – 8a**) and ssRNA (**5b – 8b**) at 298 K at neutral and alkaline pHs showing that all heptameric ssDNAs and the alkali-sensitive ssRNAs are better than 97% stable at the highest alkaline pH (12.5 to 12.75). **p. S63 – S69**

Table S1: Comparison of pK_a of **G** obtained from H8G marker (pK_{a1}) and phosphorus markers (pK_{a2}) in ssDNA (**1a – 4a**) and ssRNA (**1b – 4b**) **p. S71**

Table S2: The endocyclic $\Sigma^3J_{H1'}$ (in Hz) from ^{1}H NMR at 298 K for the respective pentofuranose moiety in each nucleotidic unit in **5a** – **8a** and **5b – 8b** in both the neutral (N) and the deprotonated (D) states. **p. S72**

Table S3: Estimation of the percentage of North-type sugar population using the $\Sigma^3J_{H1'}$ (for heptameric ssDNA (**5a** – **8a**)) and using $^3J_{1'2'}$ (for heptameric ssRNA (**5b** – **8b**))from ^{1}H NMR at 298 K for the respective pentofuranose moiety in each nucleotydic units. **p. S73**

Table S4: The ^{31}P chemical shifts (δ^{31}P) in ppm for the monomeric bisphosphate model compounds (**9** – **11**) at neutral pH (~7.25) and alkaline pH (~12.5). **p. S74**

Table S5: The line broadening values for ^{31}P chemical shifts ($\delta^{31}P$) in ppm for the monomeric bisphosphate model compounds (**9** – **11**) as well as heptameric ssDNAs/ssRNAs (**5** – **8**) at neutral pH (~7.25) and alkaline pH (~12.5) pH. **p. S75**

Table S6: ^{31}P chemical shifts [δ^{31}_{P}, in ppm] at the neutral (N) and the deprotonated (D) states at 298 K for ssDNA trimers **1a** – **4a**, ssDNA heptamers **5a** – **8a** as well as ssRNA trimers **1b** – **4b**, ssRNA heptamers **5b** – **8b.** **p. S76**

Table S7: Oligomerization shift estimated from ^{1}H chemical shift at the neutral (N) state at 298K for aromatic protons ssRNA (**5b – 8b**) using appropriate monomeric reference compounds **p. S77**

Table S8: Oligomerization shift estimated from ^{1}H chemical shift at the deprotonated (D) state at 298K for aromatic protons ssRNA (**5b – 8b**) using appropriate monomeric reference compounds **p. S78**

Electronic Supplementary Information 2

Figure S10. Panel **A** shows the time dependent degradation of all four native heptameric ssRNAs and panel **B** shows the pseudo first-order rate constants (k) for the degradation of all four native heptameric ssRNAs. **p. S2**

Figure S11. Panel **A** shows pair-wise comparison of the alkaline degradation of the native heptameric ssRNAs and the $N^{1\text{-}Me}$-G containing heptameric ssRNAs at various time intervals and panel **B** shows pair-wise comparison of the pseudo first-order rate constants (k) for the degradation. **p. S3**

Figure S12A. Panels (**a1**) – (**a2j**), (**b1**) – (**b2j**), (**c1**) – (**c2j**) and (**d1**) – (**d2$_l$**) show the RP-Hplc and SMART™ RP-Hplc profiles of all native heptameric ssRNAs at neutral state (0.0 h) and at 1h of alkaline digestion. **p. S4 – S24**

Electronic Supplementary Information 3

Figure S12B. Panels (**e1**) – (**e2j**), (**f1**) – (**f2i**) and (**g1**) – (**g2j**) show the RP-Hplc and SMART™ RP-Hplc profiles of $N^{1\text{-}Me}$-G containing heptameric ssRNAs at neutral state (0.0 h) and at 1h of alkaline digestion. **p. S2 – S15**

Electronic Supplementary Information 4

Figure S13. Panels (**a1**) – (**a8**), (**b1**) – (**b10**), (**c1**) – (**c8**), (**d1**) – (**d12**), (**e1**) – (**e11**), (**f1**) – (**f9**) and (**g1**) – (**g8**) show the Maldi Tof negative ion mode spectrum of the peaks separated by RP-Hplc and SMART™ RP-Hplc after 1h of alkali digestion for all heptameric ssRNAs. **p. S2 – S70**

Electronic Supplementary Information 5

Figure S14. An example of the calculation considering the extinction co-efficients for heptamer**7b. p. S2**

Table S9: Tables **(A) – (G)** show Maldi Tof negative ion mode mass analysis of the nucleotide fragments from the peaks separated by RP-Hplc and SMART™ RP-Hplc after 1h alkali digestion of the heptameric ssRNAs. **p. S3 – S9**

Figure S15A. Panels (**a1**) – (**a8**) show the RP-Hplc and SMART™ RP-Hplc profiles at ½, 2, 3, 4, 8, 15, 27, 48 h of alkali digestion of native heptamer **5′-r(CAAGAAC)-3′** (**5b**). **p. S10 – S17**

Electronic Supplementary Information 6

Figure S15B. Panels (**b1**) – (**b8**) show the RP-Hplc and SMART™ RP-Hplc profiles at ½, 2, 3, 4, 8, 15, 27, 48 h of alkali digestion of native heptamer **5′-r(CAAGCAC)-3′** (**6b**). **p. S2 – S9**

Electronic Supplementary Information 7

Figure S15C. Panels (**c1**) – (**c8**) show the RP-Hplc and SMART™ RP-Hplc profiles at ½, 2, 3, 4, 8, 15, 27, 48 h of alkali digestion of native heptamer **5′-r(CACGAAC)-3′** (**7b**). **p. S2 – S9**

Electronic Supplementary Information 8

Figure S15D. Panels (**d1**) – (**d8**) show the RP-Hplc and SMART™ RP-Hplc profiles at ½, 2, 3, 4, 8, 15, 27, 48 h of alkali digestion of native heptamer **5′-r(CACGCAC)-3′** (**8b**). **p. S2 – S17**

Electronic Supplementary Information 9

Figure S15E. Panels (**e1**) – (**e7**) show the RP-Hplc and SMART™ RP-Hplc profiles at ½, 2, 3, 4, 8, 15, 27 of alkali digestion of $N^{1\text{-}Me}$-G containing heptamer . **5′-r(CAAG^{Me}AAC)-3′** (**5c**). **p. S2– S8**

Electronic Supplementary Information 10

Figure S15F. Panels (**f1**) – (**f7**) show the RP-Hplc and SMART™ RP-Hplc profiles at ½, 2, 3, 4, 8, 15, 27 of alkali digestion of $N^{1\text{-}Me}$-G containing heptamer . **5′-r(CACG^{Me}AAC)-3′** (**7c**). **p. S2– S14**

Electronic Supplementary Information 11

Figure S15G. Panels (**g1**) – (**g7**) show the RP-Hplc and SMART™ RP-Hplc profiles at ½, 2, 3, 4, 8, 15, 27 of alkali digestion of $N^{1\text{-}Me}$-G containing heptamer . **5′-r(CACG^{Me}CAC)-3′** (**8c**). **p. S2– S8**

Experimental Section

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

The pH dependant ^{31}P chemical shifts for the trimeric ssDNA and ssRNA compounds **1 – 4** as well as heptameric ssDNA and ssRNA compounds **5 – 8** were performed in Bruker DRX-500 and DRX-600 spectrometers. The NMR samples for compounds **1 – 8** (Figure 1 in the text) were prepared in D_2O solution (concentration of 1 mM in order to rule out any chemical shift change owing to self-association) with using $O{=}P(OMe)_3$ (δ = 0.0 ppm) as an external standard for ^{31}P chemical shifts. All pH-dependent NMR measurements have been performed at 298 K. The pH values shown in pH-dependent ^{31}P chemical shift plots in the Figure 2 in the text and Figures S1 – S3 already includes the correction for the deuterium effect from pH meter reading [pH = pD - 0.4]. The pH meter is equipped with a calomel microelectrode (in order to measure the pH inside the NMR tube) calibrated with standard buffer solutions (in H_2O) of pH 7 and 10. The pD of the sample has been adjusted by simple addition of micro liter volumes of NaOD solutions (0.01M, 0.1M and 0.5M). The assignments for all compounds **1 – 8** [see ref 8a] at 298 K have been performed by using ^{31}P decoupled ^{1}H COSY, ^{1}H TOCSY, ^{31}P - ^{1}H correlation spectroscopy and ^{1}H NOESY**.** For each FID of ^{31}P decoupled ^{1}H DQF-COSY and TOCSY spectra, 64 scans were recorded with delay of 2s and the data were zero-filled to 4 X 1 K in the t_1 and t_2 directions, then Fourier transformed, phase adjusted and baseline corrected in both dimensions using polynomial function. ^{31}P-^{1}H Correlation spectroscopy was performed in the absolute magnitude mode using 64 scans with delay of 2s and then were zero-filled to 1 X 1 K data points in the t_1 and t_2 directions, then Fourier transformed, phase adjusted and baseline corrected in both dimensions using polynomial functions. The ^{1}H NOESY spectra were also recorded at 278 K for the compounds **1** – **4** [ref 8a] or at 283 K for the compounds **5 – 8** [ref 8a] in the neutral pH. All

NOESY spectra for **1 – 8** were recorded in 500 and 600 MHz spectrometer with mixing time (τ_m) of 800 ms. For each FID of ^{1}H NOESY 64 scans were recorded with delay of 2s and the data were zero-filled to 4 X 1 K in the t_1 and t_2 directions, then Fourier transformed, phase adjusted and baseline corrected in both dimensions using polynomial function. All ^{31}P spectra have been recorded using 128 K data points and 128 scans.

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

The error in δ^{31}P can come from the following three sources:

(1) The error from the digital resolution of the NMR spectrometer at 600 MHz: When the chemical shift of ^{31}P resonance is determined by a peak picking software, error comes from digital resolution. This issue has been tested by measuring the ^{31}P chemical shift of the model compounds (d/r) EtpApEt (**9a**/**9b**), (d/r) EtpCpEt (**10a**/**10b**) and (d/r) EtpGpEt (**11a**/**11b**) three times each both in the neutral and alkaline pH (12.5). The Maximum variation of chemical shift from three experiments at the same pH was 0.001 ppm. Table S4 shows the ^{31}P chemical shifts of these compounds (**9 – 11**).

(2) The error from the line broadening of ^{31}P resonance: Unlike ^{1}H resonance ^{31}P peaks have more line broadening. This is because in a 600 MHz spectrometer the frequency for ^{1}H is 600 MHz but the frequency for ^{31}P is 243 MHz. Similarly in a 500 MHz spectrometer the frequency of ^{1}H is 500 MHz but the frequency for ^{31}P is 202 MHz. Hence there is less resolution for ^{31}P peaks compared to that of ^{1}H. To estimate the line broadening error we have estimated the line broadening of the phosphorus peaks both at the neutral and alkaline pH for the bisphosphate model compounds (d/r) EtpApEt (**9a**/**9b**), (d/r) EtpCpEt (**10a**/**10b**) and (d/r) EtpGpEt (**11a**/**11b**) as well as the heptamers (**5 – 8**). The line broadening (in ppm) values have been tabulated in Table S5. From the Table it is evident that the line broadening in case of monomers is not more than 0.02 ppm, and in heptamer it is not

more than 0.06 ppm at a particular pH. The line broadening changes with pH in going to a higher pH from the neutral both in cases of mononucleotide bisphosphates and heptamers. The maximum change in line broadening is 0.03 ppm in case of (p_1 of **5b**) and the line broadening becomes less at the higher pH. Thus the maximum error from ^{31}P line broadening can be 0.06 ppm.

(3) The error from the salt effect: In order to achieve the alkaline pH (from 6.6 to 12.5), NaOD has been added in a step-wise manner (30-40 pH points) to the sample solution. As a control study the chemical shift of phosphorus has been measured in the neutral and alkaline pH for model compounds (r/d)EtpApEt (**9a**/**9b**), (r/d)EtpCpEt (**10a**/**10b**), and (r/d) EtpGpEt (**11a**/**11b**). The nucleobases A and C in (**9a**/**9b**) and (**10a**/**10b**) do not have any deprotonation site. Hence the change in chemical shift of ^{31}P from pH 6.6 to 12.5 in these compounds should be entirely due to the salt effect as the pH changes to alkaline. The chemical shift values both in the neutral and alkaline pH for the compounds **9** – **11** are given in Table S4. The maximun change in ^{31}P chemical shift is 0.061 ppm for p_2 of Etp_1Ap_2Et (**9b**) over the total pH range of 6.6–12.5. Thus we have taken Etp_1Ap_2Et (**9b**) as a model compound and performed a blank titration on it over the same pH range (6.6 – 12.5), which is exactly similar to the pH range of the compounds (**1 – 8**) in our study. The δ^{31}P for the pH dependant titration plots of trimeric ssDNAs/ssRNAs (**1 – 4**) and heptameric ssDNAs/ssRNAs (**5 – 8**) have been prepared by subtracting the average value of δ^{31}P for p_1 and p_2 in Etp_1Ap_2Et (**9b**) from the ^{31}P chemical shift of the trimers and the heptamer at a particular pH. This difference in ^{31}P chemical shift has been plotted as a function of pH to obtain the pK_a from the inflection point of each titration plots. Thus, Etp_1Ap_2Et (**9b**) has been chosen as a model compound as adenosine has no deprotonation site over the pH range (6.6 – 12.5) in which the titration studies have been performed, so that the change in chemical shift can be attributed to salt effect. At the same time this had the maximum change in chemical shift value (0.061 ppm) in the pH range (6.6 – 12.5). So this is the maximum error that can come from the salt effect. In

our compounds the (d/r) trimers (**1 – 4**) and (d/r) heptamers (**5 – 8**), each internucleotidic phosphate in the sequence can be at the same time 3' to one nucleobase and 5' to the next nucleobase. As a result while subtracting the δ^{31}P of Etp_1Ap_2Et (**9b**) from that of concerned trimer and heptamer ^{31}P resonances, average of 3'(p_1) and 5'(p_2) phosphorus chemical shifts has been used. Thus by taking into account all the above three sources of error, the pK_a of 9-guaninyl from δ^{31}P chemical shift (pK_{a2}) has only been calculated when the change in chemical shift of a particular ^{31}P resonance, over the total pH range (6.6 – 12.5), is more than 0.10 ppm in order to avoid erroneous interpretation of our results.

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

(1) Accuracy of pK_a: The pK_a values for the ionization at N^1 center of **G** (obtained from δH8**G** marker protons of the 9-guaninyl residue) have been obtained in our earlier work by the Hill plot analysis of the pH-dependent ^{1}H chemical shifts measured by both 500 and 600 MHz NMR (see ref 8a). The pK_a values reported here are obtained by the pH dependant chemical shift of ^{31}P markers, which are neighboring to the 9-guaninyl ionization site in the pH range 6.6 – 12.5 for compounds **1 – 8**. The error in the ^{31}P chemical shift at 298K is discussed in the preceding section. The error in pK_a determination from ^{31}P markers is ±0.1 to ±0.9 in case of trimeric compounds **1 – 4** and ±0.0 to ±0.13 in heptameric ssDNAs/ssRNAs **5 – 8**. All individual errors of respective pK_a values are shown in Table 1 in text and S1. These accurate pK_a values allow us to safely attribute the observed pK_a differences larger than ±0.13 as significant differences owing to differential electronic environment around the phosphates. 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 of 6.6 to 12.5 for each compound), and the pH readings were found to vary only ±0.025, hence no buffer was used for our study.

(2) The pH range studied: The pH titration studies for d(Ap_1Gp_2A) (**1a**) (pH 6.61 – 12.45); d(Ap_1Gp_2C) (**2a**) (pH 6.73 – 12.15); d(Cp_1Gp_2A) (**3a**) (pH 6.94 – 12.11); d(Cp_1Gp_2C) (**4a**) (pH 7.21 – 11.31); d($Cp_1Ap_2Ap_3Gp_4Ap_5Ap_6C$) (**5a**) (pH 7.09 – 12.5); d($Cp_1Ap_2Ap_3Gp_4Cp_5Ap_6C$) (**6a**) (pH 6.91 – 12.58); d($Cp_1Ap_2Cp_3Gp_4Ap_5Ap_6C$) (**7a**) (pH 7.48 – 12.5); and for d($Cp_1Ap_2Cp_3Gp_4Cp_5Ap_6C$) (**8a**) (pH 7.43 – 12.33); consist of ~30 – 40 data points (Figure S1 – S3). Similarly the pH titration studies for r(Ap_1Gp_2A) (**1b**) (pH 6.79 – 11.8); r(Ap_1Gp_2C) (**2b**) (pH 7.05 – 11.95); r(Cp_1Gp_2A) (**3b**) (pH 6.72 – 12.02); r(Cp_1Gp_2C) (**4b**) (pH 7.2 – 12.16); r($Cp_1Ap_2Ap_3Gp_4Ap_5Ap_6C$) (**5b**) (pH 7.0 – 12.49); r($Cp_1Ap_2Ap_3Gp_4Cp_5Ap_6C$) (**6b**) (pH 7.23 – 12.50); r($Cp_1Ap_2Cp_3Gp_4Ap_5Ap_6C$) (**7b**) (pH 6.96 – 12.75); and for r($Cp_1Ap_2Cp_3Gp_4Cp_5Ap_6C$) (**8b**) (pH 6.75 – 12.47); consist of ~30 – 40 data points (Figure S1 – S3). The corresponding Hill plots for **1a** – **4a** and **5a – 8a** as well as **1b** – **4b** and **5b – 8b** are given in Figures S6 – S7 and the pK_as shown in Table 1 in text and S1 have been calculated from non-linear curve-fitting analyses and Hill plot analyses (see section D for details).

(D) pK_a Determination

The pH-dependent [over the range of pH 6.6 – 12.5, with an interval of pH 0.2 – 0.3] $\delta^{31}P$ for **1 - 8** show a sigmoidal behavior [Figure 2 in the text and Figures S1 – S3] for the deprotonation of 9-guaninyl moiety in trimeric and heptameric ssDNAs (**1a – 8a**) and ssRNAs (**1b – 8b**). In some ^{31}P resonances of ssRNA sequences (p2 and p4 of **6b** and p4 of **7b**) the pH-dependent $\delta^{31}P$ plots do not reach a distinct plateau at the end of deprotonation of 9-guaninyl. In other ^{31}P resonances of heptameric ssRNA sequences a plateau is reached for the deprotonation of 9-guaninyl but $\delta^{31}P$ shift at even higher pH values show deviation from the plateau. This observation is a result of the influence of ionization of vicinal 2'-OH group in ssRNAs on the phosphorus chemical shifts, which start before the complete deprotonation of 9-guaninyl in cases where plateau is not reached and after the complete

deprotonation of 9-guaninyl in cases where the plateau is reached. For sequence **5b** the pH titration was done in the pH range 7 – 12.49. The influence of ionization of vicinal 2'-OH group on δ^{31}P was found for pH values 12.4 – 12.49 in p_2, for pH values 12.25 – 12.49 in p_3 and for p_4 there was no such influence till pH 12.49. For sequence **6b** the pH titration was done in the pH range 7 – 12.5. The influence of ionization of vicinal 2'-OH group on δ^{31}P was found for pH values 12.37 – 12.5 in p_2, for pH values 12.12 – 12.5 in p_3, and for pH values 12.12 – 12.5 in p_4. For sequence **7b** the pH titration was done in the pH range 6.69 – 12.75. The influence of ionization of vicinal 2'-OH group on δ^{31}P was found for pH values 12.6 – 12.75 in p_2, for pH values 12.6 – 12.75 in p_3, and for pH values 12.5 – 12.75 in p_4. For sequence **8b** the pH titration was done in the pH range 6.75 – 12.47. The influence of ionization of vicinal 2'-OH group on δ^{31}P was found for pH value 12.47 in p_2, for p_3 there was no such influence till pH 12.47 and for pH values 12.14 – 12.47 in p_4. δ^{31}P data points were not considered for pK_a determination in these pH values. These points are marked as (▲) in Figure 2 and Figure S1. In an effort to understand if the δ^{31}P can be affected by the ionization of neighboring 2'-OH group, a model compound (**8c**) where the 9-guaninyl ionization was blocked by methylation of N^1 of G, was titrated on a pH range of 8 – 12.5. It was observed that the ^{31}P markers in **8c** showed no significant chemical shift change (Figure S1) below pH 11.6 which could be solely attributed to the influence of 2'-OH ionization on δ^{31}P at pH value in and above 11.6.

The pK_a determination in compounds **1 -8** is based on non-linear curve fitting and the Hill plot analysis. The Hill plot analysis was done using equation: $pH = \log((1-\alpha)/\alpha) + pK_a$, where α represents fraction of the protonated species. The value of α is calculated from the change of chemical shift relative to the deprotonated (D) state at a given pH ($\Delta_D = \delta_D - \delta_{obs.}$ for deprotonation, where δ_{obs} is the experimental chemical shift at a particular pH and δ_D is the chemical shift at the deprotonated state), divided by the total change in chemical shift between neutral (N) and deprotonated (D) state (Δ_T). So

the Henderson-Hasselbalch type equation can then be written as pH = log ((Δ_T - Δ_D)/Δ_D) + pK_a. The pK_a is calculated from the linear regression analysis of the Hill plot [Figure S6 – S7] .

The end point plateau at higher pH (δ_D) used in the Hill plot analysis, could not be measured clearly for the phosphate markers affected by 2'-OH ionization (p2 and p4 of **6b** and p4 of **7b**). Going to a pH value beyond 12.5 would also increase larger error due to salt effect. In an effort to treat all data uniformly, the ^{31}P chemical shift at a higher pH than pH 12.5 was determined in compounds **1** – **8** using the modified Henderson equation $\delta_{obs} = \delta_h + 1/K_a (\delta_l - \delta_{obs})a_{H+}$ where δ_h is the chemical shift at higher pH (pH > 12.5); δ_l is the chemical shift at low pH; and δ_{obs} is the chemical shift at observed pH. The above equation gives a straight line with a slope $1/K_a$ and δ_h value at the intercept (Figures S4 – S5). Value of intercept (δ_h) has been used in the Hill plot analysis as chemical shift at deprotonated state (δ_D).

As the 2'-OH ionization effect on δ^{31}P did not start before pH 11.6 in **8c** it would be advantageous to determine the end point plateau at higher pH and perform Hill plot analysis with δ^{31}P data points in the pH values between 10 – 11.6, to avoid effect of 2'-OH ionization on the pK_a data in case of ssRNAs. In the end point plateau determination at higher pH and Hill plot analysis data points are taken from the slope of the sigmoidal curve. Thus for ssRNAs where data points beyond 11.6 was used for end point plateau determination at higher pH and Hill plot analysis another set was done with data points in the pH between 10 – 11.6. The pK_a values were determined in three different ways in trimeric ssDNA and ssRNA (**1 – 4**) as well as heptameric ssDNAs and ssRNAs(**5 - 8**). In some ^{31}P resonances (p4 of **5b**, p2, p4 of **6b**, and p2 of **8b**) in case of ssRNAs it has been determined in five different ways. The five different ways of determining pK_a include (i) p$K_a^{\#}$ obtained from non-linear curve fitting of pH vs. experimental δ^{31}P plot. (ii) pK_a^{*} obtained from non-linear curve fitting of pH vs. experimental δ^{31}P including the calculated δ^{31}P at deprotonated state using δ^{31}P values between pH

10 - 11.6 in the calculation of the chemical shift at deprotonated state. To determine the pH value of the calculated chemical shift, the pK_a value obtained from the Hill plot analysis a ^{31}P resonance is taken and the pH value at the starting of the sigmoidal curve is deducted from it. This difference in pH is added to the pK_a value obtained from Hill plot analysis and the pH value of the calculated chemical shift is determined. (iii) pK_a^{∞} obtained from Hill plot analysis using δ^{31}P values between pH 10 - 11.6 in the calculation of the chemical shift at deprotonated state as well as Hill plot analysis. (iv) pK_a^{+} is obtained from non-linear curve fitting of pH vs. experimental δ^{31}P including the calculated δ^{31}P at deprotonated state using δ^{31}P values between pH 10 - 11.9 in the calculation of the δ^{31}P at deprotonated state and (v) pK_a^{∂} is obtained from Hill plot analysis using δ^{31}P at deprotonated state using δ^{31}P values between pH 10 - 11.9 in the calculation of the δ^{31}P at deprotonated state as well as Hill plot analysis. The final pK_a value obtained for any ^{31}P marker was an average of the different ways of pK_a determination discussed in (i) to (v) above.

(E) Alkaline Hydrolysis of the Heptameric ssRNAs

To the lyophilized solid oligonucleotides (5 od at λ_{260}) 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 taken out at suitable time intervals and immediately quenched with aqueous acetic acid (10 µL, 0.03N) to pH ~7. This was then stored at -20°C until Hplc analysed.

(F) Hplc separation of the Hydrolysis products at different time intervals

(1) The aliquots of the reaction were injected in RP-Hplc on a chromasil 100 C18 column (250 x 8 mm, 5 µm particle size). A mixture of acetonitrile in 0.1M triethylammonium acetate (TEAA) buffer was

used as eluent (buffer A: 5% CH_3CN in 0.1M TEAA, buffer B: 50% CH_3CN in 0.1M TEAA, pH 6.95). A linear gradient, starting with 0% B buffer + 100% A buffer to 25% B buffer + 75% A buffer in 40 minutes, with a constant flow rate of 1.5 ml min^{-1} was used.

(2) As the Maldi Tof mass-spectral analysis showed that all peaks separated by the above standard RP-Hplc system did not always give pure component, we used the Pharmacia-LKB Biotechnolgy's SMART™ RP-Hplc micro-purification system (Column: Jupiter 5µm C18 RP column, 150 mm x 2 mm, 300Å) for separation of those mixed fractions. Separations were performed at either 20º C or at 60º C. *See Section (G) for separation of each primary product as a pure component.* For the exact gradient and elution system employed were different for different mixtures, and they are documented in ESI 2, Fig S12A and ESI 3, Fig S12B under each elution profile.

(3) Elution profiles for 0h and 1h hydrolysis, which are used for complete analysis of the primary products formed after 1h of alkaline digestion, are given in Fig S12A in ESI 2 and Fig S 12B in ESI 3. All elution profiles at different time points (½, 1, 2, 3, 4, 8, 15, 27 and 48h) were however used in determining the degradation profile and the degradation rate constants for the heptamers (see both Fig S12A (ESI 2), S12B(ESI 3) and S15A (ESI 5) – S15G (ESI 11))

(G) Procedure for the separation of each primary product as a pure component after 1 h of alkaline hydrolysis

Detailed Hplc and mass-spec analysis were performed after 1h of alkaline digestion for all four heptameric ssRNAs in order to purify and characterize each of the primary products of alkaline hydrolysis. The 0 h and 1 h Hplc profiles used for this purpose are shown in ESI 2, Fig S12 (a1)-(a2) for **5b**. SMART™ Hplc profiles of the impure fractions (monitored by Maldi Tof mass-spectra) after first RP-Hplc are given in Fig S12 (a2i) and (a2j), ESI 2. The mass spectrum after the separation of all the

pure components after RP-Hplc and SMART™ Hplc are shown in the Fig S13 (a1)-(a8), ESI 4. Likewise, 0 h and 1 h Hplc profiles are shown in ESI 2, Fig S12 (b1)-(b2) for **6b**. SMART™ Hplc profiles of the impure fractions after first RP-Hplc are given in Fig S12 (b2i) and (b2j), ESI 2. The mass spectrum after the separation of all the pure components are shown in the Fig S13 (b1)-(b10), ESI 4. For **7b**, the 0 h and 1 h Hplc profiles are shown in ESI 2, Fig S12 (c1)-(c2). The SMART™ Hplc profiles are given in S12 (c2i) and (c2j). Mass spectrum after the separation of all the pure components are shown in the Fig S13 (c1)-(c8), ESI 4. For **8b**, the 0h and1h Hplc profiles are shown in ESI 2, Fig S12 (d1)-(d2). The SMART™ Hplc profiles are given in Fig S12 (d2i) - ($d2_l$), ESI 2. The mass spectrum after the separation of all the pure components are shown in the Fig S13 (d1)-(d12), ESI 4. Mass spectral analysis is given in Table S9 in ESI 5. For **5c**, the 0h and1h Hplc profiles are shown in ESI 3, Fig S12 (e1)-(e2). The SMART™ Hplc profiles are given in Fig S12 (e2i) - (e2j), ESI 3. The mass spectrum after the separation of all the pure components are shown in the Fig S13 (e1)-(e11), ESI 4. Mass spectral analysis is given in Table S9 in the ESI 4. For **7c**, the 0h and1h Hplc profiles are shown in ESI 3, Fig S12 (f1)-(f2). The SMART™ Hplc profiles are given in Fig S12 (f2i), ESI 3. The mass spectrum after the separation of all the pure components are shown in the Fig S13 (f1)-(f10), ESI 4. Mass spectral analysis is given in Table S9 in ESI 5. For **8c**, the 0h and 1h Hplc profiles are shown in ESI 3, Fig S12 (g1)-(g2). The SMART™ Hplc profiles are given in Fig S12 (g2i) - ($g2_j$), ESI 3. The mass spectrum after the separation of all the pure components are shown in the Fig S13 (g1)-(g8), ESI 4. Mass spectral analysis is given in Table S9 in the ESI 5.

(H) Procedure for Maldi Tof analysis

For Maldi Tof analysis of Hplc fractions an Ultraflex Tof/Tof instrument (Bruker Daltonics, Germany) operated in reflector mode was used. Fractions collected from the Hplc separations were dried and

redissolved in water (10 µl). 0.5 µl of sample was mixed on the target plate with 0.5 µl of 50 mg / ml diammonium citrate in water and 0.5 µl of 10 mg/ml of 2,4,6-Trihydroxy acetophenone in acetonitrile/water 50/50, v/v. The mass spectrometer was externally calibrated with a peptide mixture using alpha-cyano-4-hydroxycinnamic acid as matrix. The assigned masses are monoistopic. Mass-spectral analyses of all fractions after 1 h digestion are shown in Figure S13 (ESI 4), whereas Tables S9 (A-G) in ESI 5 show the mass-spectral analysis of various peaks in the Supporting Information.

(I) 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 (% area), at various incubation time, considering the heptamer peak as 100% area at 0.0 h. All of the four ssRNA heptamers were contaminated by small amount of both non-nucleot(s)idic impurities (at $R_T \approx 6.6'$ and $\approx 10.5'$) as well as nucleotidic impurities which have been defined for quantitation purposes (see below). Hence, all peaks formed upon degradation were recalculated after subtracting the impurities (taking the parent heptamer peak as 100% before addition of alkali) depending upon the retention times.

(1) 5'-r(CAAGAAC)-3' (**5b**) (Figure S12 (a1) for Hplc profile in the ESI 2): The heptameric peak was 97% pure before the alkali was added. A mixture of 5'-r(GAAC)-3' (m/z 1246.2) and 5'-r($CAAG_{2', 3'-cMP}$)-3' (m/z 1308.2) were present at R_T = 24.01' with an area of 0.5%. A mixture of 5'-r(AGAAC)-3' (m/z 1575.3) and 5'-r($CAAGA_{2', 3'-cMP}$)-3' (m/z 1636.3) was also present at R_T = 25.73'. A 0.9% impurity consisting of 5'-r(AAGAAC)-3' (m/z 1903.3) was present at R_T = 27.59'.

(2) 5'-r(CAAGCAC)-3' (**6b**) (Figure S12 (b1) for RP-Hplc profile in ESI 2): The heptameric peak was 94% pure before the alkali was added. An impurity of 5'-r(CAC)-3' (m/z 876.2) was present at R_T =

19.82' with an area of 1.4%. A mixture of 5'-r($CAA_{2'/3'-P}$)-3' (m/z 980.1) and 5'-r(GCAC)-3' (m/z 1221.2) was present at R_T = 24.01' with an area of 2.9%.

(3) 5'-r(CACGAAC)-3' (**7b**) (Figure S12 (c1) for RP-Hplc profile in ESI 2): The heptameric peak was 91.2% pure before the alkali was added. A mixture of 5'-r($CACGA_{2'/3'-P}$)-3' (m/z 1612.2) and 5'-r(ACGAAC)-3' (m/z 1879.2) was present at R_T = 26.51' with an area of 2.0%.The major impurities at R_T = 28.05' with an area of 4.3% were non-nucleot(s)idic impurities (m/z 1229.2, 1973.1).

(4) 5'-r(CACGCAC)-3' (**8b**) (Figure S12 (d1) for RP-Hplc profile in the ESI 2): The heptameric peak was ~95% pure before the alkali was added. A mixture of 5'-r($CACG_{2', 3'-cMP}$)-3' (m/z 1283.2) and 5'-r($CACGC_{2', 3'-cMP}$)-3' (m/z 1588.2) was present at R_T = 23.71' with an area of 1.4%. At R_T = 25.5' with an area of 1.5% was a non-nucleot(s)idic impurity (m/z 1949.2).

(5) 5'-r(CAA**G^{Me}**AAC)-3' (**5c**) (Figure S12 (e1) for Hplc profile in the ESI 3): The heptameric peak was 96.7% pure before the alkali was added. A 0.8% impurity containing of 5'-r(AAG^{Me}AAC)-3' (m/z 1917.2) was present at R_T = 24.35'.

(6) 5'-r($CACG^{Me}$AAC)-3' (**7c**) (Figure S12 (f1) for RP-Hplc profile in the ESI 3): The heptameric peak was 98.6% pure before the alkali was added. An impurity of 5'-r($CACG^{Me}A_{2', 3'-cMP}$) (m/z 1626.3) was present at R_T = 24.77' with an area of 1.3%.

(7) 5'-r($CACG^{Me}$CAC)-3' (**8c**) (Figure S12 (g1) for RP-Hplc profile in the ESI 3): The heptameric peak was ~99% pure before the alkali was added. Only a non-nucleot(s)idic impurity of 1.0% was present at R_T = 24.03'.

The hepatameric peak of **8b**, corresponding to each time point, and the heptameric peak together with the peak eluted after for **7c** corresponding to each time point of alkaline degradation, were further analyzed by SMART^{TM} RP-Hplc micro purification system as some smaller fragments (5'-ACGCAC-3' for **8b** and 5'-ACG^{Me}AAC-3', 5'-G^{Me}AAC-3', 5'-AAC-3' for 7c) were co-eluting together with the parent

heptamer [for the elution profiles see Fig S12 (d1)-(d2$_l$) in ESI 2, Fig S12 (f1)-(f2$_i$) in ESI 3, Fig S15 (d1)-(d8$_i$) in ESI 8 and S15 (f1)-(f7) in ESI 10; for the mass after separation of the hexameric peak from the heptamer, see ESI 4, Fig S13 (d10)-(d11) and S13 (f5)-(f8)]. A linear gradient, starting with 0% B buffer + 100% A buffer to 20% B buffer + 80% A buffer in 40 minutes with a constant flow rate of 100 μl m^{-1} was used for SMART™ RP-Hplc. The contents of buffer A and buffer B were the same as used in the first round of RP-Hplc.

(J) Calculation of percent cleavage at different sites of ssRNAs by alkaline degradation

A two-step procedure was adopted to take in to account of sequence-dependent hypochromicity contribution in the calculation of % area of each RP-Hplc /SMART™ RP-Hplc peak at time interval of 1 hour of the alkaline digestion.

(1) 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 Section H). Some of the single-peaks in the first round of RP-Hplc were found to contain more than one hydrolysis products (nucleotide fragments) when analyzed by Maldi Tof mass spectrometry. They were further analyzed by SMART™ RP-Hplc micro purification system. Those separated pure components/peaks were characterized again by Maldi Tof and were subsequently used in the next step (see below) for extinction-coefficient correction to determine the actual contribution of the different components in the % area shown in Figure 3 of the main text (see Figures S12 in ESI 2 and 3, Fig S13 in ESI 4, and Fig S15A(ESI 5) – Fig S15G(ESI 11)).

(2) Since the intrastrand stacking in the oligo RNA will essentially reduce the absorption of various oligonucleotides (hypochromic effect) depending upon their chain length as well as sequence context, we have initially defined the chain length and the sequence of each peak by Maldi Tof mass-spectral

analysis for calculation of extinction coefficients. Then the % area of each RP-Hplc/SMART™ RP-Hplc peak (after the first step) was further corrected by calculating the extinction coefficient at 260 nm at 20 °C at neutral pH for the single-strand RNA using the nearest-neighbor method. Average error of calculated extinction coefficients was shown to be around 4 %. See, for example, programs in http://www.owczarzy.net/emethod.htm and http://scitools.idtdna.com/Analyzer/. An example of the calculation considering the extinction co-efficients for **7b** is given for clarity in ESI 5, Figure S14.

Notes:

Note 1: Earlier, it has been observed[11g] that there is a change in the ^{31}P chemical shift value of neighboring phosphates of NADH on going to its oxidized state NAD. The phosphorus close to the pyridine ring of NADH undergoes 35 cps upfield shift upon oxidation. The phosphorus close to adenine ring undergoes an upfield shift of 10 cps. These upfield shifts have been explained as a result of an interaction between the positively charged nitrogen of pyridine ring of NAD and a phosphate next to it upon oxidation of NADH. This phosphate in turn interacts with the other phosphate near to adenine and thereby introducing a change in chemical shift of the second phosphorus also. We have also seen[8b] that due to distant interaction between 5'-phosphate and imidazole moiety of the 9-guaninyl in EtpG (**12**), the pK_a of 9-guaninyl becomes more basic compared to pK_a of 9-guaninyl in $G_{3'}$pEt (**13**). In our present work we have seen a situation when, upon ionization of a guanine-9-yl to guanylate ion, the deshielding of phosphorus resonances from the adjoining phosphates take place, possibly due to the charge repulsion interaction of the phosphate and the N1-deprotonated guanine-9-yl. This down-field ^{31}P chemical shift in the alkaline pH compared to those in the neutral is a well known effect, in general, in various types of phosphates[11a-h], phosphonates[11a,b] and aminophosphonates[11b], as in our case with the phosphates of the heptameric RNAs (Figure 2, see SI, Table S5), which also show pK_{a2}. (see Section C in the text for a discussion of this electrostatic charge repulsion effect).

Note 2: The upfield or downfield pH-dependent ^{31}P chemical shift is dictated by sum of three contributions[11b]: (a) variation of electronegativity of the oxygen atoms adjacent to phosphorus (inductive model), (b) variation of the O-P-O bond angles[11a], and (c) weaker ^{31}P screening for delocalization of charge into the phosphorus dπ orbitals. Variation of O-P-O bond angle is sensitive both to nature of *O*-alkyl substituents[11a] as well as owing to the involvement of dπ orbitals of phosphorus atom[11c]. A deshielding with decreasing O-P-O bond angles at the phosphorus atom has been established[11a,b] in a large series of alkylphosphates as well as for the five- and six-membered thioxophosphonates, phosphinates and phosphonates (δ^{31}P of the five-membered systems are more deshielded than that of the six-membered systems because O-P-O bond angle is smaller than that of the latter). This ^{31}P chemical shift change on the basis of O-P-O bond angle argument however does not apply[11a] to phosphines, symmetric phosphonium ions and some phosphate containing biological molecules[11i,j], which are likely similar to the well-known α, β and γ substituent effects in ^{13}C shifts[11a].

Note 3: Water is a highly polar molecule that can effectively stabilize developing negative charge on the salicylates through hydrogen bonding. On the other hand, DMSO is not capable of providing such effective solvation. Thus, the stabilization[12a] of the salicylate monoanion by the H-bond from water renders deprotonation of the carboxylic group more favorable thereby decreasing (more acidic) its pK_a value. On the other hand, the strength of an intramolecular H-bonding of salicylates in a hydrophobic environment (such as in DMSO) increases by several orders of magnitude than in the aqueous environment because of increase of the charge density on the donor/acceptor, which is likely to be true for the rate enhancement found in the enzyme's active site. Many examples[12b-k] of perturbed pK_a of carboxylic groups [in aspartate (Asp) and glutamate (Glu)] are found in protein owing to specific

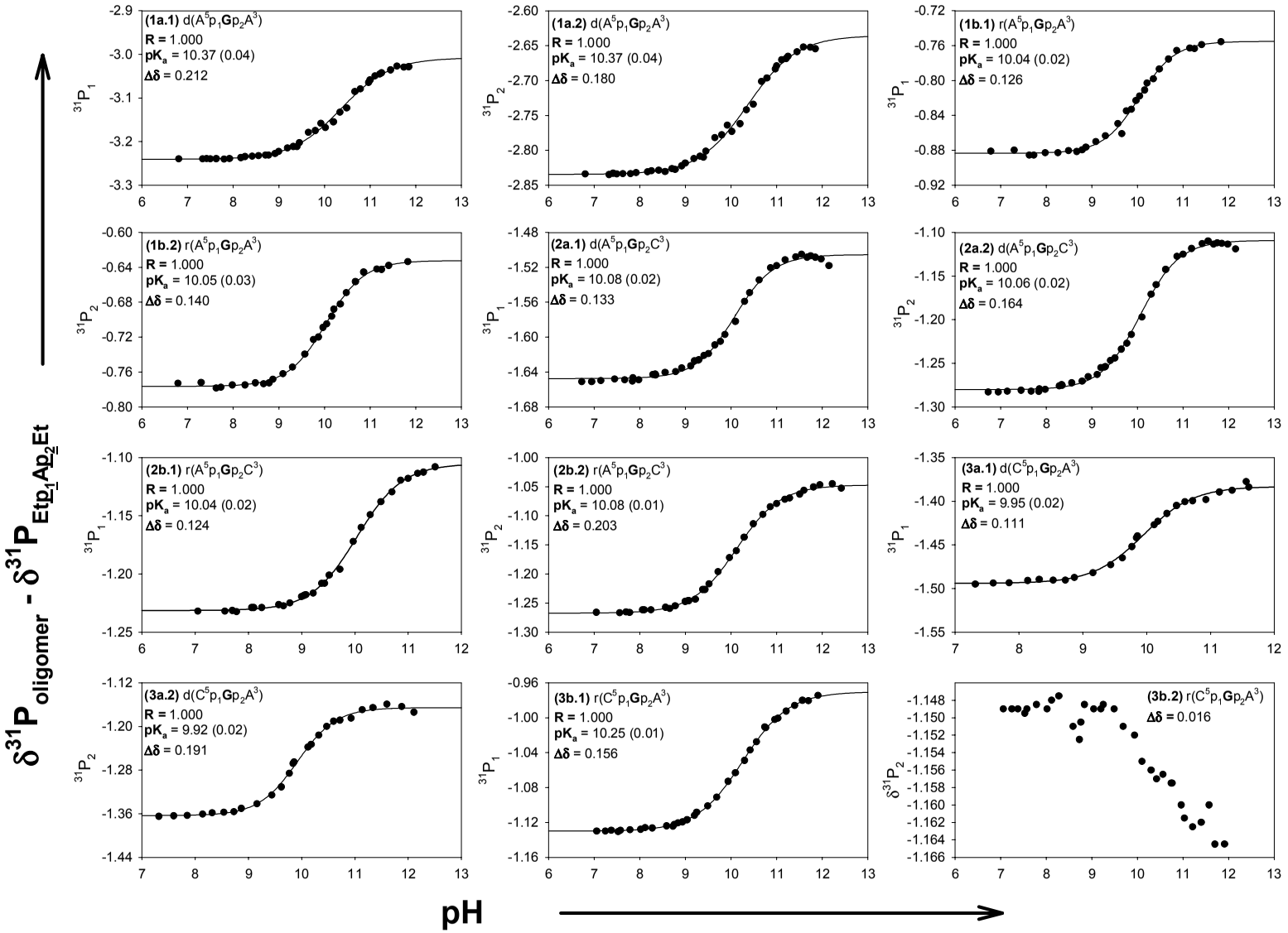
folding. Thus, those Asp or Glu, which are more buried inside protein in a rather hydrophobic pocket show elevated pK_a whereas those having more solvated microenvironment or having possibilities of forming H-bond stabilization through other neighboring group[12c-j] participation due to structural pre-organization show lower pK_a. A conserved buried Asp group inside a hydrophobic pocket of *Escherichia Coli* Thioredoxin[12b] in folded state show abnormally high pK_a of 7.5 compared to pK_a 3.9 – 4.0 for fully solvated Asp. Similarly, carboxyl group of Glu41 in the N-terminal domain of Rat T-lymphocyte glycoprotein CD2 shows[12c] unusually elevated pK_a of 6.73 compared to pK_a 4.3 – 4.4 for fully solvated Glu. The crystal structure of this protein shows[12c] that side chain methylene groups of Glu41 and Glu29 are within the van der Waals contact, and the carboxyl groups are positioned such to develop strong charge-charge repulsion, which causes the pK_a elevation. On the other hand, pH dependent NMR studies for turkey ovomucoid third domain (OMTKY3) show[12d] the presence of solvent exposed H-bonds involving side chains of Asp7, Glu19 and Asp27 having pK_a of <2.6, 3.2 and <2.3 respectively, which may explain their low pK_a values. NMR measurement along with comparative crystal studies for hen and turkey lysozymes shows[12e] that except Asp101 and Glu35, carboxylate groups of all acidic residue form H-bonds to stabilize the charged form thereby lowering their pK_a values (varies between pK_a 2.68 – 3.78). On the other hand, the hydrophobic microenvironment and interaction between carboxyl and neighboring groups in Glu35 in hen and turkey lysozymes[12e] contribute to its elevated pK_a of 6.06 and the solvated nature of Asp101 show its non-perturbed pK_a of 4.09. By combining pH titration and site-directed mutagenesis, it has been shown[12f] that low pK_a (3.54) for Glu19 of $P_{1'}$ site ($P_{1'}$-Glu^{19}) in OMTKY3 protein is due to H-bonding between carboxylate of $P_{1'}$-Glu^{19} and hydroxyl group of P_2-Thr^{17} (intermolecular) as well as between carboxylate of $P_{1'}$-Glu^{19} and amide of $P_{1'}$-Glu^{19} (intramolecular) and the charge-charge interaction between $P_{1'}$-Glu^{19} and $P_{3'}$-Arg^{21}. NMR and neutron diffraction studies with serine protease showed[7d]

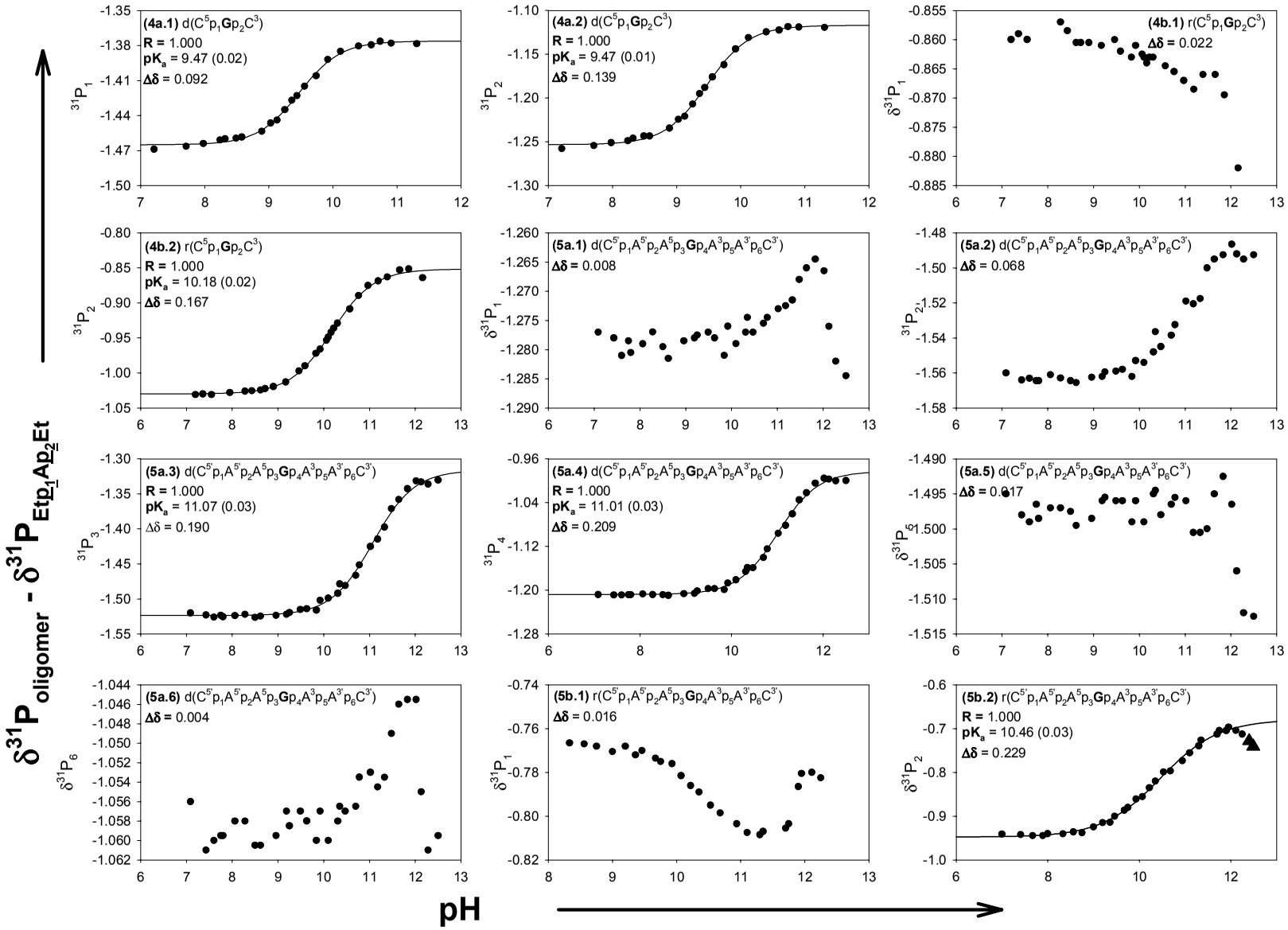
that buried Asp102 having perturbed pK_a of <2 at the active site of chymotrypsin due to the H-bonded network from neighboring histidine and serine acting as catalytic triad. The Asp32 (pK_a <3) and Asp215 (pK_a >5) are the catalytically active residue in pepsine, which have been shown[12l] to be involved in H-bonding between themselves as found from kinetics and crystal data, like in maleic acid[12l] where the two H-bonded carboxylate moieties have pK_a values 1.9 and 6.2.

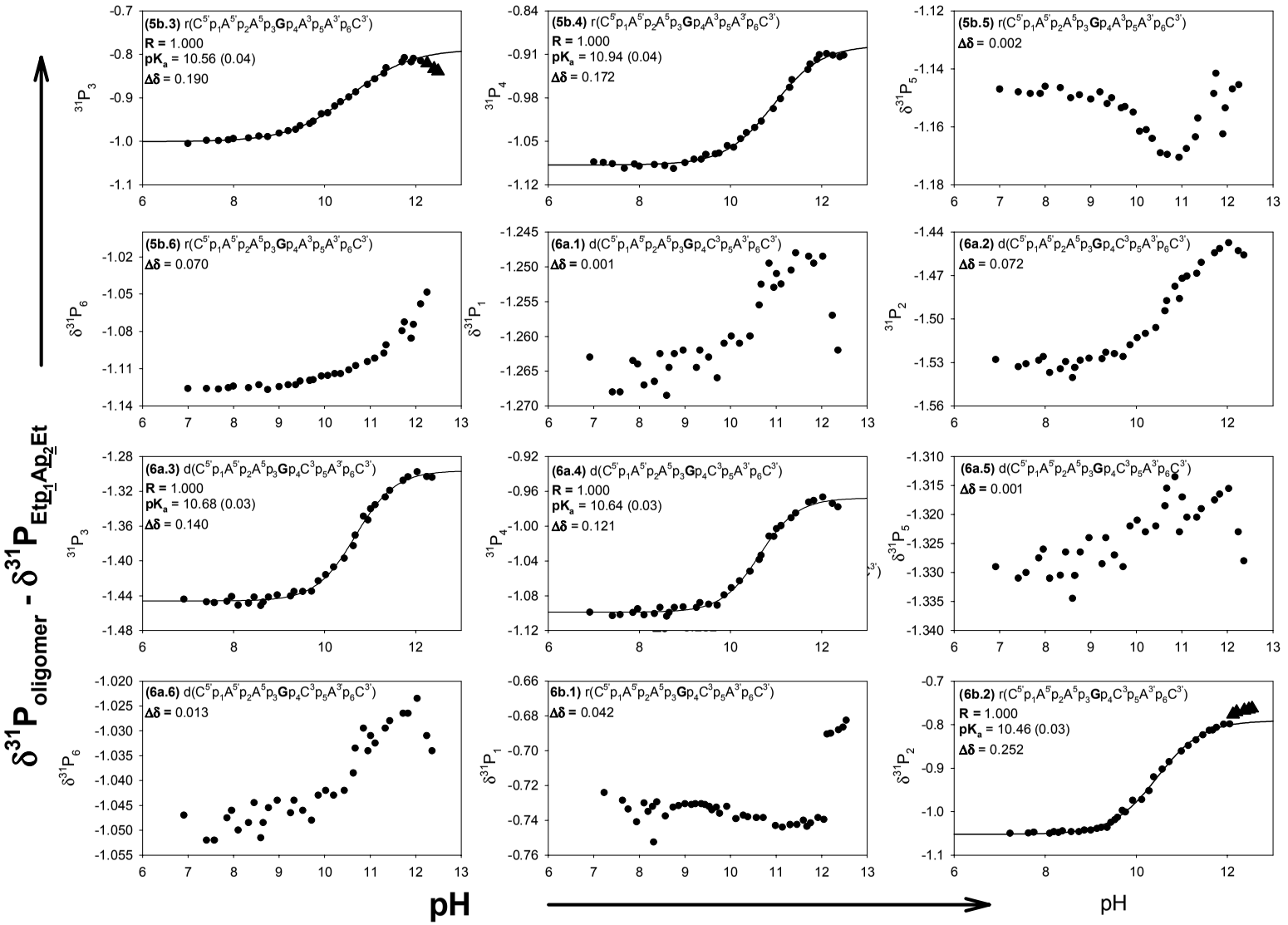
Note 4: The sugar conformations both at the neutral and the deprotonated states have been estimated for heptameric ssDNAs and ssRNAs from their respective $\Sigma J_{H1'}$, and $^3J_{1'2'}$coupling constants (Table S2). We have observed that for heptameric ssDNA the sugar conformation [North (N)(C2'-*exo*-C3'-*endo*) ⇆ South (S)(C2'-*endo*-C3'-*exo*)][13] for the sugar attached to guanine (maximum change of +9% more N-type with pH in **8a**) and the sugars attached to neighboring nucleobases (maximum change of +10% more N-type with pH in nucleobase C3' in **5a**) does not vary significantly over the pH range (6.6 – 12.5). For heptameric ssRNA on the other hand there is a significant change (maximum change is 39% more S-type with pH in case of **5b** and minimum change is 0% in **8b** among heptameric ssRNAs) in the sugar conformation for the sugar attached to guanine as well as the sugars attached to neighboring nucleobases of guanine (maximum change is 38% more S-type with pH in case of A3' of **5b**. and minimum change is 10% more S-type with pH in C3' of **8b** and 10% more N-type with pH in A5 of **6b** among heptameric ssRNAs, Table S2). With the generation of a negative charge in the oligomers (**1 – 8**) while doing the pH titration in the pH range 6.6 – 12.5 would produce repulsion between the newly formed $\underline{\mathbf{G}}^-$ and the negatively charged phosphate (pK_a 1.5). Hence stretching of the oligomers is expected on going from a neutral to alkaline pH. In C2'-*endo* conformation of sugar the 3' and the 5' phosphates remain apart by 7.0Å, whereas this distance decrease to 5.9Å in C3'-*endo* sugar conformation[13c]. Thus, if there is any stretching of the backbone leading to increase in distance between

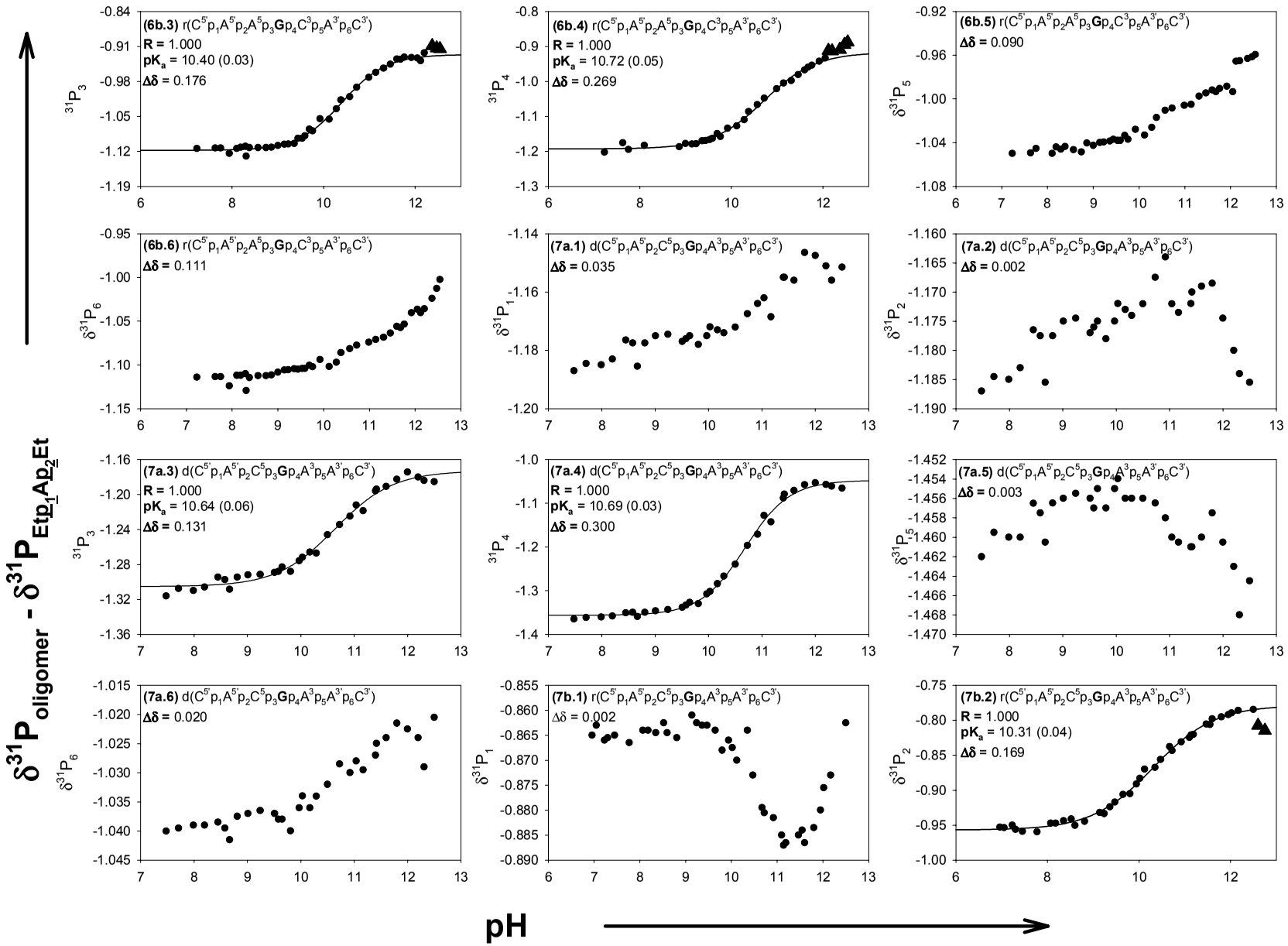
3' and the 5' phosphates due to deprotonation of guanine it can be reflected in the respective sugar conformation. For heptameric ssDNA, there is no appreciable change in endocyclic sugar conformation with pH. However, in heptameric ssRNAs the sugar conformations, apart from central guanine in a sequence, becomes more S-type in most cases with formation of $\underline{\mathbf{G}}^{-}$, which in turn causes expansion of the phosphate backbone. For sugar residue of A^{5} of **5b**, $A^{5'}$ and A^{5} of **6b**, the sugar conformation becomes more N-type with the ionization of $\underline{\mathbf{G}}$, causing contraction of phosphate backbone. In case of heptameric ssRNAs, there are appreciable changes of sugar conformation (error is ~2%) for some sugar moieties. However, in many cases (sugar residues of $C^{5'}$, $A^{5'}$ and $C^{3'}$ of **5b**; $C^{5'}$, C^{3}, $A^{3'}$ and $C^{3'}$ of **6b**; $A^{5'}$, A^{3}, $A^{3'}$ and $C^{3'}$ of **7b**; $A^{5'}$, C^{5}, C^{3} and $A^{3'}$ of **8b)** the sugar conformational change is marginal with increase in pH as well. Thus while ssDNAs are marginally contracted, ssRNAs undergoes stretching on ionization of guanine. However, in overall, it is obvious that the role of conformational effect due to sugar-phosphate backbone pre-organization is minimal in both heptameric ssRNAs and ssDNAs.

Figures S1:


(1a.1) d(A5p1Gp2A3)
R = 1.000
pKa = 10.37 (0.04)
Δδ = 0.212
(1a.2) d(A5p1Gp2A3)
R = 1.000
pKa = 10.37 (0.04)
Δδ = 0.180
(1b.1) r(A5p1Gp2A3)
R = 1.000
pKa = 10.04 (0.02)
Δδ = 0.126
(1b.2) r(A5p1Gp2A3)
R = 1.000
pKa = 10.05 (0.03)
Δδ = 0.140
(2a.1) d(A5p1Gp2C3)
R = 1.000
pKa = 10.08 (0.02)
Δδ = 0.133
(2a.2) d(A5p1Gp2C3)
R = 1.000
pKa = 10.06 (0.02)
Δδ = 0.164
(2b.1) r(A5p1Gp2C3)
R = 1.000
pKa = 10.04 (0.02)
Δδ = 0.124
(2b.2) r(A5p1Gp2C3)
R = 1.000
pKa = 10.08 (0.01)
Δδ = 0.203
(3a.1) d(C5p1Gp2A3)
R = 1.000
pKa = 9.95 (0.02)
Δδ = 0.111
(3a.2) d(C5p1Gp2A3)
R = 1.000
pKa = 9.92 (0.02)
Δδ = 0.191
(3b.1) r(C5p1Gp2A3)
R = 1.000
pKa = 10.25 (0.01)
Δδ = 0.156
(3b.2) r(C5p1Gp2A3)
Δδ = 0.016
δ31Poligomer - δ31PEtp1Ap2Et
pH


(4a.1) d(C5p1Gp2C3)
R = 1.000
pKa = 9.47 (0.02)
Δδ = 0.092
(4a.2) d(C5p1Gp2C3)
R = 1.000
pKa = 9.47 (0.01)
Δδ = 0.139
(4b.1) r(C5p1Gp2C3)
Δδ = 0.022
(4b.2) r(C5p1Gp2C3)
R = 1.000
pKa = 10.18 (0.02)
Δδ = 0.167
(5a.1) d(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
Δδ = 0.008
(5a.2) d(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
Δδ = 0.068
(5a.3) d(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 11.07 (0.03)
Δδ = 0.190
(5a.4) d(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 11.01 (0.03)
Δδ = 0.209
(5a.5) d(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
Δδ = 0.017
(5a.6) d(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
Δδ = 0.004
(5b.1) r(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
Δδ = 0.016
(5b.2) r(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 10.46 (0.03)
Δδ = 0.229
δ31Poligomer - δ31PEtp1Ap2Et
pH


(5b.3) r(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 10.56 (0.04)
Δδ = 0.190
(5b.4) r(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 10.94 (0.04)
Δδ = 0.172
(5b.5) r(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
Δδ = 0.002
(5b.6) r(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
Δδ = 0.070
(6a.1) d(C5'p1A5'p2A5p3Gp4C3p5A3'p6C3')
Δδ = 0.001
(6a.2) d(C5'p1A5'p2A5p3Gp4C3p5A3'p6C3')
Δδ = 0.072
(6a.3) d(C5'p1A5'p2A5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.68 (0.03)
Δδ = 0.140
(6a.4) d(C5'p1A5'p2A5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.64 (0.03)
Δδ = 0.121
(6a.5) d(C5'p1A5'p2A5p3Gp4C3p5A3'p6C3')
Δδ = 0.001
(6a.6) d(C5'p1A5'p2A5p3Gp4C3p5A3'p6C3')
Δδ = 0.013
(6b.1) r(C5'p1A5'p2A5p3Gp4C3p5A3'p6C3')
Δδ = 0.042
(6b.2) r(C5'p1A5'p2A5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.46 (0.03)
Δδ = 0.252
δ31P3
δ31P4
δ31P5
δ31P6
δ31P1
δ31P2
δ31Poligomer - δ31PEtp1Ap2Et
pH
pH


δ31Poligomer - δ31PEtp1Ap2Et
pH
(6b.3) r(C5'p1A5'p2A5'p3Gp4C3'p5A3'p6C3')
R = 1.000
pKa = 10.40 (0.03)
Δδ = 0.176
(6b.4) r(C5'p1A5'p2A5'p3Gp4C3'p5A3'p6C3')
R = 1.000
pKa = 10.72 (0.05)
Δδ = 0.269
(6b.5) r(C5'p1A5'p2A5'p3Gp4C3'p5A3'p6C3')
Δδ = 0.090
(6b.6) r(C5'p1A5'p2A5'p3Gp4C3'p5A3'p6C3')
Δδ = 0.111
(7a.1) d(C5'p1A5'p2C5'p3Gp4A3'p5A3'p6C3')
Δδ = 0.035
(7a.2) d(C5'p1A5'p2C5'p3Gp4A3'p5A3'p6C3')
Δδ = 0.002
(7a.3) d(C5'p1A5'p2C5'p3Gp4A3'p5A3'p6C3')
R = 1.000
pKa = 10.64 (0.06)
Δδ = 0.131
(7a.4) d(C5'p1A5'p2C5'p3Gp4A3'p5A3'p6C3')
R = 1.000
pKa = 10.69 (0.03)
Δδ = 0.300
(7a.5) d(C5'p1A5'p2C5'p3Gp4A3'p5A3'p6C3')
Δδ = 0.003
(7a.6) d(C5'p1A5'p2C5'p3Gp4A3'p5A3'p6C3')
Δδ = 0.020
(7b.1) r(C5'p1A5'p2C5'p3Gp4A3'p5A3'p6C3')
Δδ = 0.002
(7b.2) r(C5'p1A5'p2C5'p3Gp4A3'p5A3'p6C3')
R = 1.000
pKa = 10.31 (0.04)
Δδ = 0.169

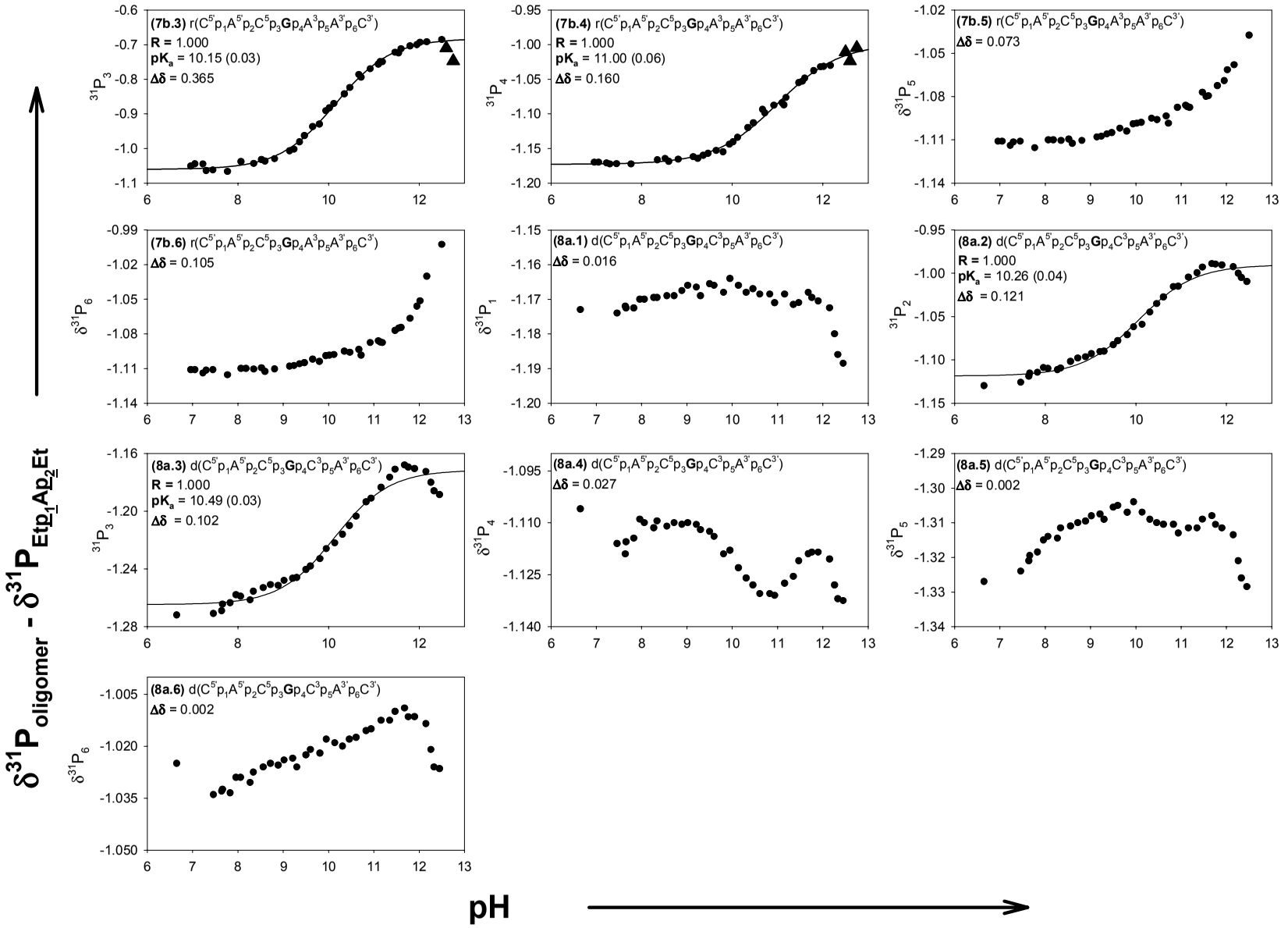
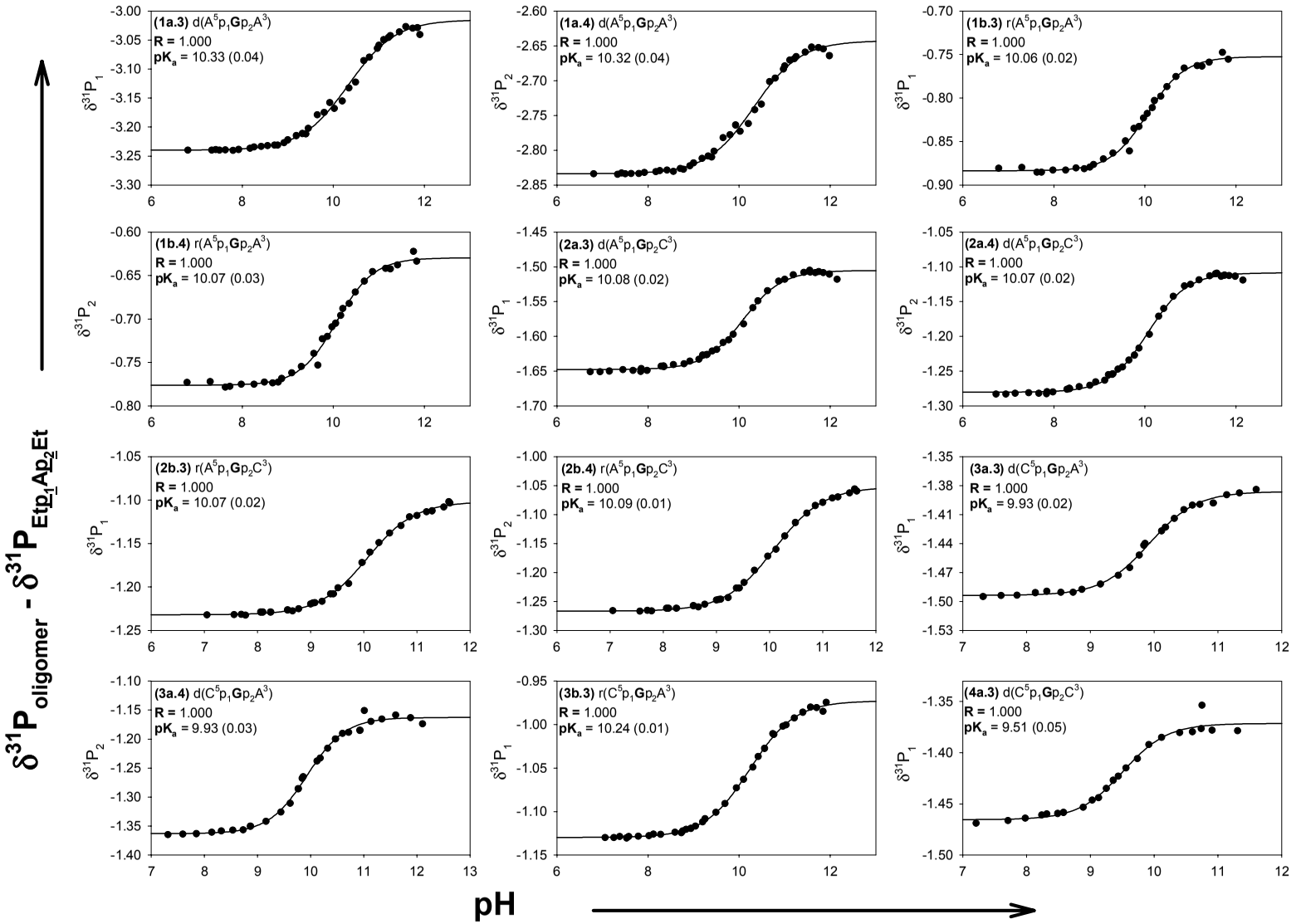
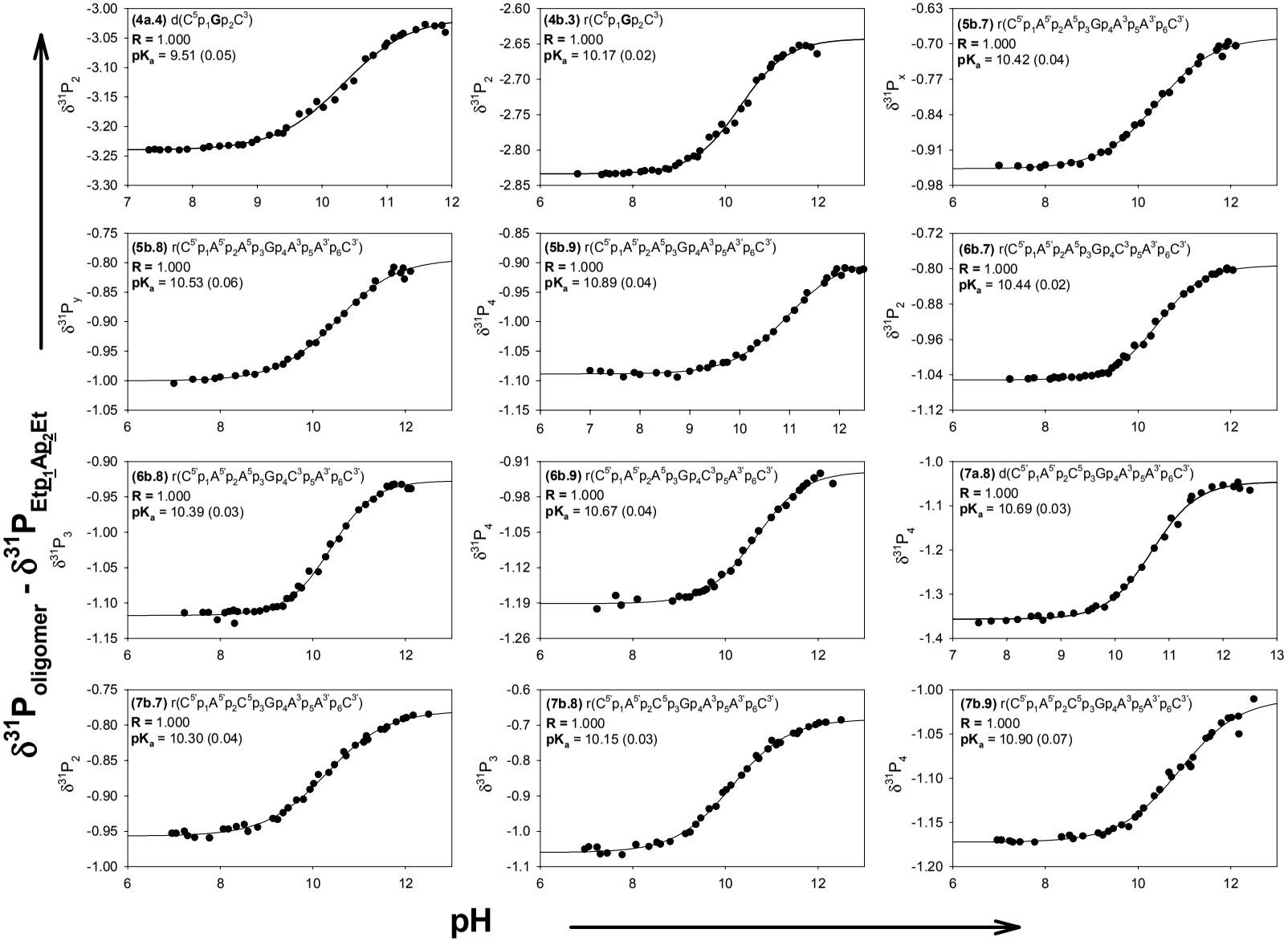

(7b.3) r(C5'p1A5'p2C5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 10.15 (0.03)
Δδ = 0.365
(7b.4) r(C5'p1A5'p2C5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 11.00 (0.06)
Δδ = 0.160
(7b.5) r(C5'p1A5'p2C5p3Gp4A3p5A3'p6C3')
Δδ = 0.073
(7b.6) r(C5'p1A5'p2C5p3Gp4A3p5A3'p6C3')
Δδ = 0.105
(8a.1) d(C5'p1A5'p2C5p3Gp4C3p5A3'p6C3')
Δδ = 0.016
(8a.2) d(C5'p1A5'p2C5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.26 (0.04)
Δδ = 0.121
(8a.3) d(C5'p1A5'p2C5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.49 (0.03)
Δδ = 0.102
(8a.4) d(C5'p1A5'p2C5p3Gp4C3p5A3'p6C3')
Δδ = 0.027
(8a.5) d(C5'p1A5'p2C5p3Gp4C3p5A3'p6C3')
Δδ = 0.002
(8a.6) d(C5'p1A5'p2C5p3Gp4C3p5A3'p6C3')
Δδ = 0.002
δ31Poligomer - δ31PEtp1Ap2Et
pH

Figure S1. Plot of pH-dependent (pH 6.6 – 12.75, corrected for deuterium effect, see experimental section) ^{31}P chemical shifts (δ^{31}P) (corrected for salt effect by subtracting the average δ^{31}P of the 5' phosphorus (p_1) and 3' phosphorus (p_2) of Et$\underline{p}_1$Ap_2Et (**9b**) from the δ^{31}P shifts of different phosphates marker of the oligomers (**1 – 8**) at each pH values, see experimental section for details) at each pH values for different ^{31}P markers of the trimeric (**1a – 4a**) and heptameric (**5a – 8a**) ssDNA as well as trimeric (**1b – 4b**) and heptameric (**5b – 7b**) ssRNA, showing the pK_a at the inflection point. Chemical shift variations at 30 – 40 different pH values have been measured in an interval of 0.2 – 0.3 pH units to obtain the sigmoidal curves. Each panel shows chemical shift change with pH for one particular ^{31}P marker. It is to be noted that pK_a values have been calculated only for those ^{31}P markers where $\Delta\delta >$ 0.10 ppm except for ^{31}P resonances of some ssRNA sequence (p_6 in **6b** and p_6 in **7b** p_6) 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 as the pH vs. δ^{31}P did not show a proper sigmoidal behavior. The change in δ^{31}P was a result of the effect of 2'-OH ionization on the ^{31}P chemical shifts (see experimental section D). Data points in the panels 5b.2, 5b.3, 6b.2, 6b.3, 6b.4, 7b.2, 7b.3 and 7b.4 marked with the symbol (▲) were not considered while doing curve-fitting analysis. These δ^{31}P chemical shifts were influenced by the vicinal 2'-OH ionization and hence not used for pK_a determination. The name of the compounds, correlation coefficient (R) obtained from each curve fitting through the experimental NMR titration points, pK_a values obtained after non-linear curve fitting as well as the total change in chemical shift over the pH range ($\Delta\delta$) are shown in each graphs.

Figures S2:


(1a.3) d(A5p1Gp2A3)
R = 1.000
pKa = 10.33 (0.04)
(1a.4) d(A5p1Gp2A3)
R = 1.000
pKa = 10.32 (0.04)
(1b.3) r(A5p1Gp2A3)
R = 1.000
pKa = 10.06 (0.02)
(1b.4) r(A5p1Gp2A3)
R = 1.000
pKa = 10.07 (0.03)
(2a.3) d(A5p1Gp2C3)
R = 1.000
pKa = 10.08 (0.02)
(2a.4) d(A5p1Gp2C3)
R = 1.000
pKa = 10.07 (0.02)
(2b.3) r(A5p1Gp2C3)
R = 1.000
pKa = 10.07 (0.02)
(2b.4) r(A5p1Gp2C3)
R = 1.000
pKa = 10.09 (0.01)
(3a.3) d(C5p1Gp2A3)
R = 1.000
pKa = 9.93 (0.02)
(3a.4) d(C5p1Gp2A3)
R = 1.000
pKa = 9.93 (0.03)
(3b.3) r(C5p1Gp2A3)
R = 1.000
pKa = 10.24 (0.01)
(4a.3) d(C5p1Gp2C3)
R = 1.000
pKa = 9.51 (0.05)
δ31P1
δ31P2
δ31Poligomer - δ31PEtp1Ap2Et
pH


δ31Poligomer - δ31PEtp1Ap2Et
pH
(4a.4) d(C5p1Gp2C3)
R = 1.000
pKa = 9.51 (0.05)
δ31P2
(4b.3) r(C5p1Gp2C3)
R = 1.000
pKa = 10.17 (0.02)
δ31P2
(5b.7) r(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 10.42 (0.04)
δ31Px
(5b.8) r(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 10.53 (0.06)
δ31Py
(5b.9) r(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 10.89 (0.04)
δ31P4
(6b.7) r(C5'p1A5'p2A5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.44 (0.02)
δ31P2
(6b.8) r(C5'p1A5'p2A5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.39 (0.03)
δ31P3
(6b.9) r(C5'p1A5'p2A5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.67 (0.04)
δ31P4
(7a.8) d(C5'p1A5'p2C5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 10.69 (0.03)
δ31P4
(7b.7) r(C5'p1A5'p2C5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 10.30 (0.04)
δ31P2
(7b.8) r(C5'p1A5'p2C5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 10.15 (0.03)
δ31P3
(7b.9) r(C5'p1A5'p2C5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 10.90 (0.07)
δ31P4

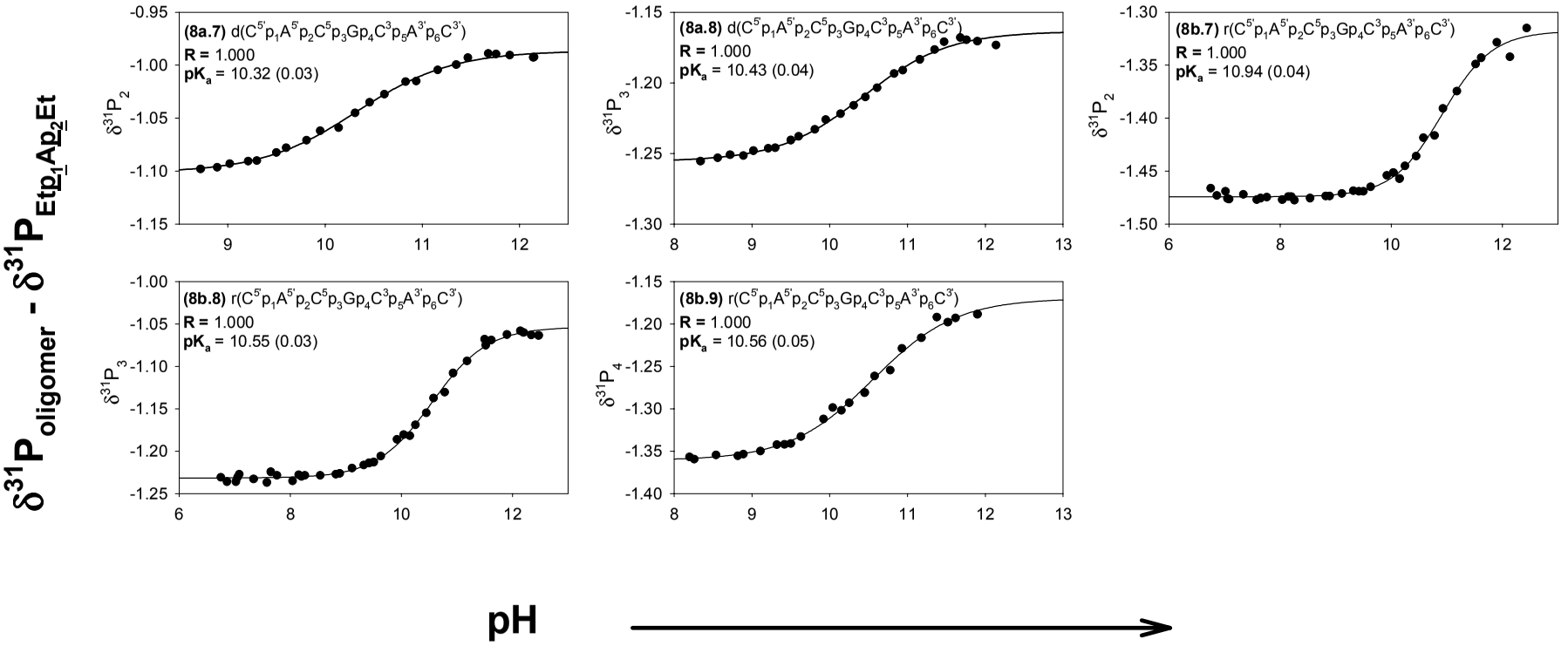

δ31Poligomer - δ31PEtp1Ap2Et
(8a.7) d(C5'p1A5'p2C5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.32 (0.03)
δ31P2
(8a.8) d(C5'p1A5'p2C5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.43 (0.04)
δ31P3
(8b.7) r(C5'p1A5'p2C5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.94 (0.04)
δ31P2
(8b.8) r(C5'p1A5'p2C5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.55 (0.03)
δ31P3
(8b.9) r(C5'p1A5'p2C5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.56 (0.05)
δ31P4
pH

Figure S2. Plot of pH-dependent (pH 6.6 – 12.5, corrected for deuterium effect, see experimental section) ^{31}P chemical shifts (δ^{31}P) (corrected for salt effect by subtracting the average δ^{31}P of the 5' phosphorus (p_1) and 3' phosphorus (p_2) of Et$\underline{p}_1$Ap_2Et (**9b**) from the δ^{31}P shifts of different phosphates marker of the oligomers (**1 – 8**). At each pH values, see experimental section for details) for different ^{31}P markers of the trimeric (**1a – 4a**) and heptameric (**5a – 8a**) ssDNA as well as trimeric (**1b – 4b**) and heptameric (**5b – 8b**) ssRNA, showing the pK_a at the inflection point. Each panel shows chemical shift change with pH for one particular ^{31}P marker. In these figures the calculated δ^{31}P at high pH (using data points between pH 10 – 11.6 in determining δ^{31}P at high pH) has been included along with the experimental δ^{31}P values to do a non-linear curve fitting (see experimental section D). The name of the compounds, correlation coefficient (R) obtained from each curve fitting through the experimental NMR titration points as well as the pK_a values obtained after non-linear curve fitting are shown in each graphs.

Figures S3:

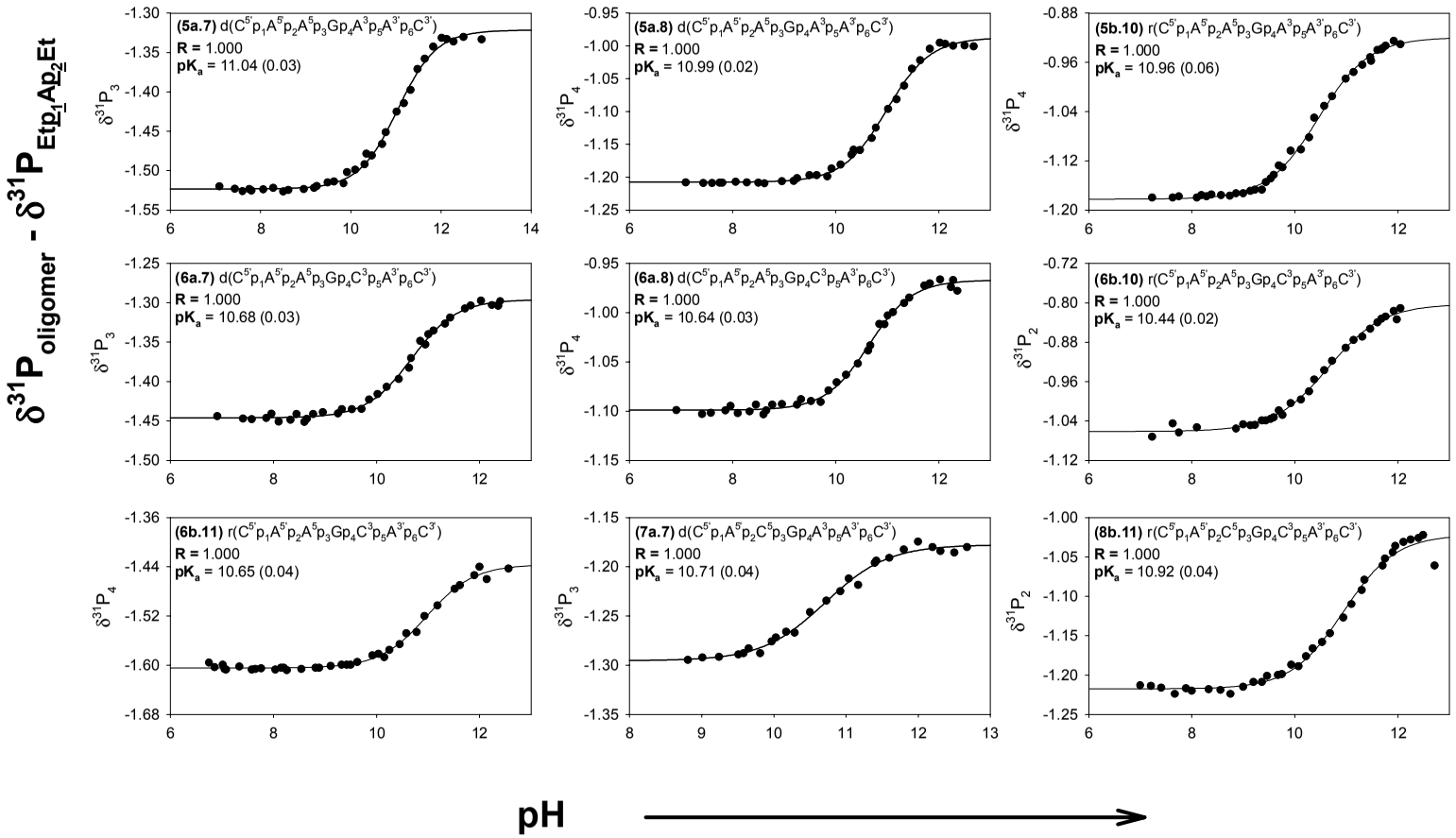
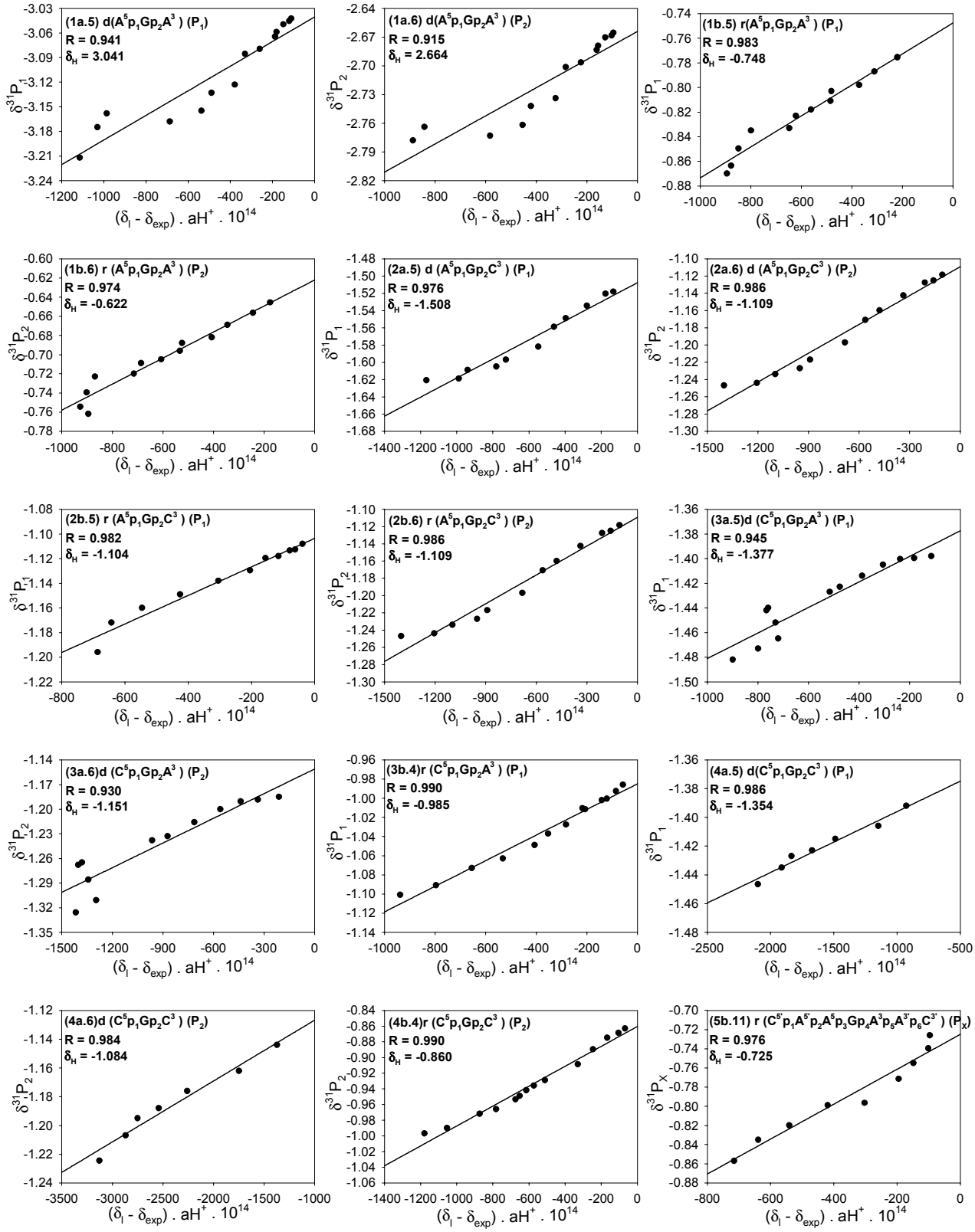
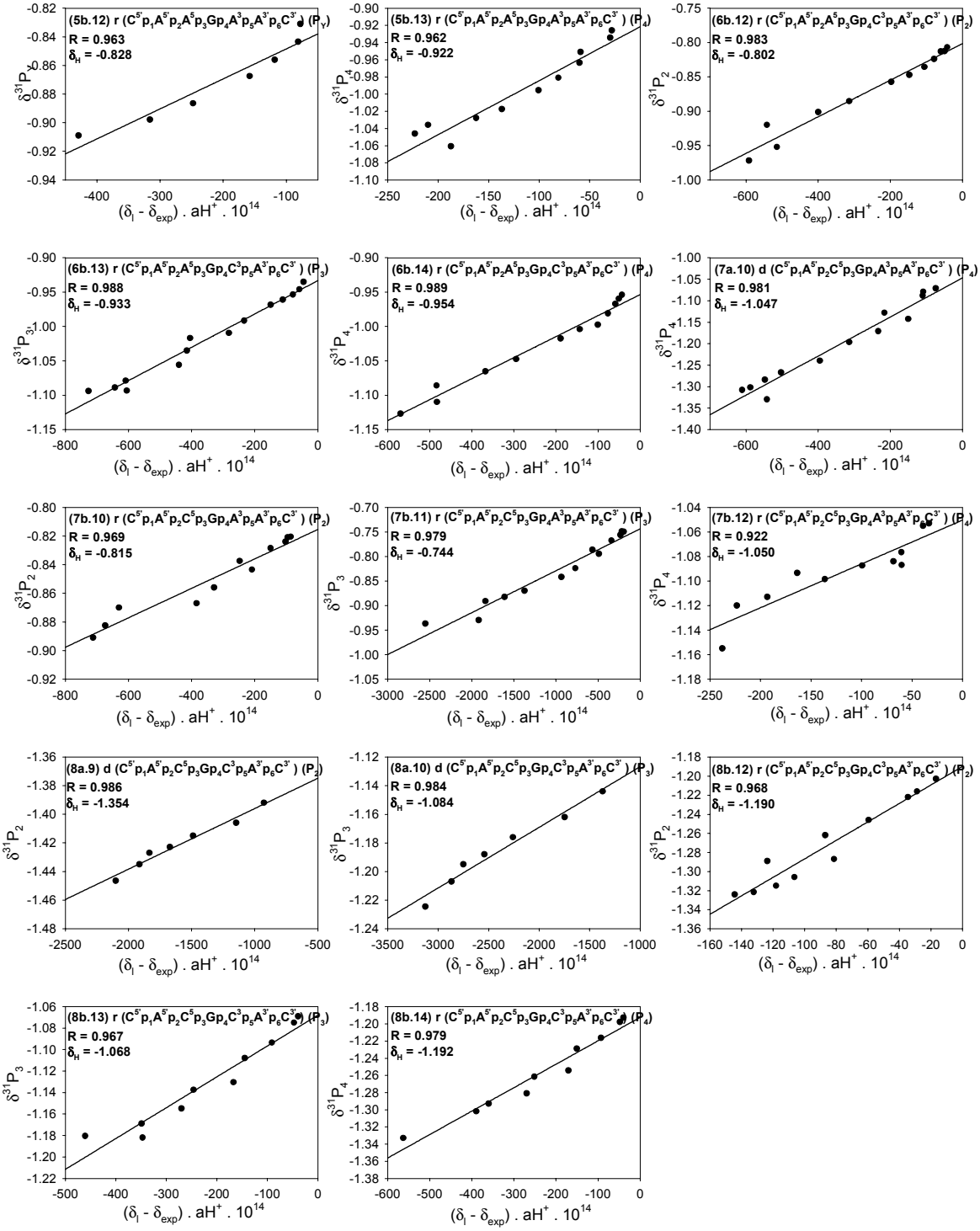

δ31Poligomer - δ31PEtp1Ap2Et
(5a.7) d(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 11.04 (0.03)
(5a.8) d(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 10.99 (0.02)
(5b.10) r(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 10.96 (0.06)
(6a.7) d(C5'p1A5'p2A5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.68 (0.03)
(6a.8) d(C5'p1A5'p2A5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.64 (0.03)
(6b.10) r(C5'p1A5'p2A5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.44 (0.02)
(6b.11) r(C5'p1A5'p2A5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.65 (0.04)
(7a.7) d(C5'p1A5'p2C5p3Gp4A3p5A3'p6C3')
R = 1.000
pKa = 10.71 (0.04)
(8b.11) r(C5'p1A5'p2C5p3Gp4C3p5A3'p6C3')
R = 1.000
pKa = 10.92 (0.04)
pH

Figure S3. Plot of pH-dependent (pH 6.6 – 12.5, corrected for deuterium effect, see experimental section) ^{31}P chemical shifts (δ^{31}P) (corrected for salt effect by subtracting the average δ^{31}P of the 5' phosphorus (p_1) and 3' phosphorus (p_2) of Et$\underline{p}_1$Ap_2Et (**9b**) from the δ^{31}P shifts of different phosphates marker of the oligomers (**5 – 8**) at each pH value, see experimental section for details) for different ^{31}P markers of heptameric (**5a – 8a**) ssDNA as well as heptameric (**5b – 8b**) ssRNA, showing the pK_a at the inflection point. Each panel shows chemical shift change with pH for one particular ^{31}P marker. In these figures the calculated δ^{31}P at high pH (using data points in pH values between 10 - 11.9 in determining δ^{31}P at high pH) has been included along with the experimental δ^{31}P values to do a non-linear curve fitting (see experimental section D). The name of the compounds, the correlation coefficient (R) obtained from each curve fitting through the experimental NMR titration points as well as the pK_a values obtained after non-linear curve fitting are shown in each graphs.

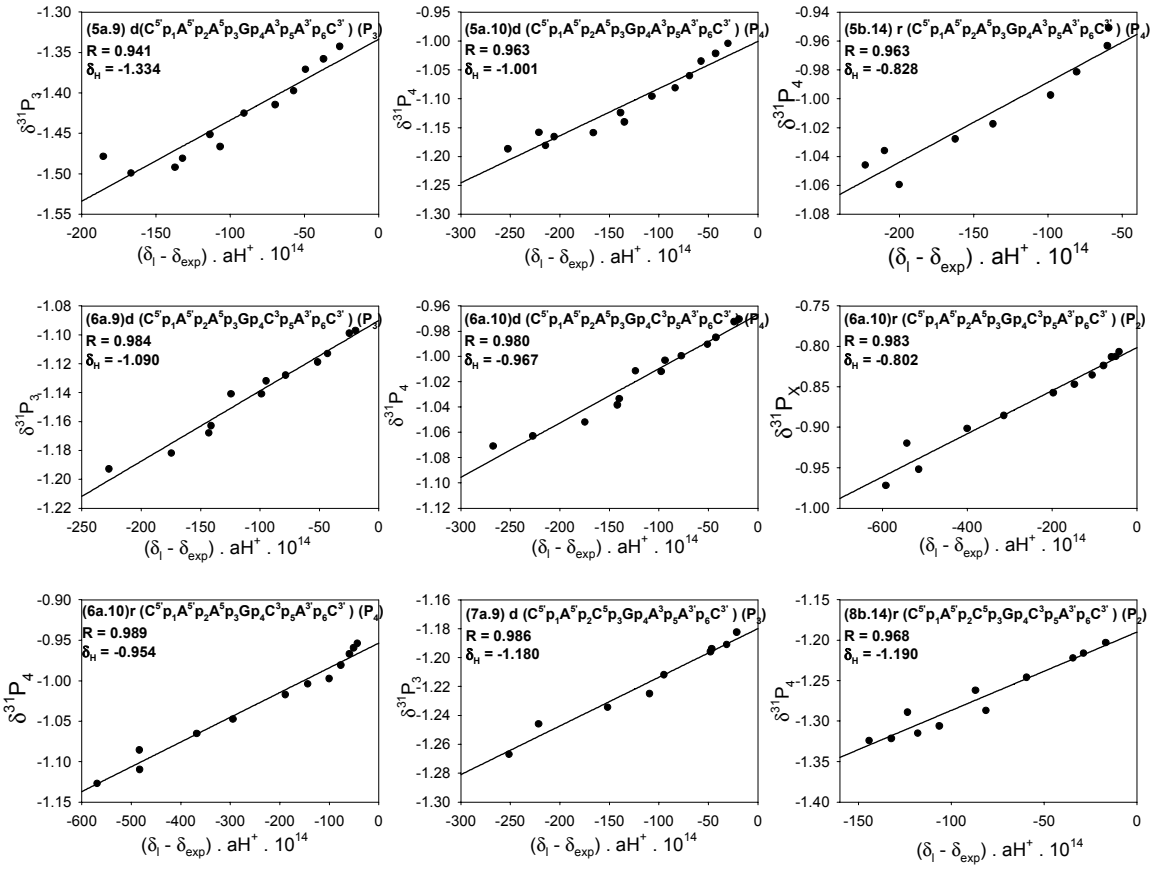
Figures S4:


(1a.5) d(A5p1Gp2A3) (P1)
R = 0.941
δH = 3.041
(1a.6) d(A5p1Gp2A3) (P2)
R = 0.915
δH = 2.664
(1b.5) r(A5p1Gp2A3) (P1)
R = 0.983
δH = -0.748
(1b.6) r (A5p1Gp2A3) (P2)
R = 0.974
δH = -0.622
(2a.5) d (A5p1Gp2C3) (P1)
R = 0.976
δH = -1.508
(2a.6) d (A5p1Gp2C3) (P2)
R = 0.986
δH = -1.109
(2b.5) r (A5p1Gp2C3) (P1)
R = 0.982
δH = -1.104
(2b.6) r (A5p1Gp2C3) (P2)
R = 0.986
δH = -1.109
(3a.5)d (C5p1Gp2A3) (P1)
R = 0.945
δH = -1.377
(3a.6)d (C5p1Gp2A3) (P2)
R = 0.930
δH = -1.151
(3b.4)r (C5p1Gp2A3) (P1)
R = 0.990
δH = -0.985
(4a.5) d(C5p1Gp2C3) (P1)
R = 0.986
δH = -1.354
(4a.6)d (C5p1Gp2C3) (P2)
R = 0.984
δH = -1.084
(4b.4)r (C5p1Gp2C3) (P2)
R = 0.990
δH = -0.860
(5b.11) r (C5'p1A5'p2A5p3Gp4A3p5A3'p6C3') (PX)
R = 0.976
δH = -0.725
δ31P1
δ31P2
δ31PX
(δI - δexp) . aH+ . 10^14


(5b.12) r (C5'p1A5'p2A5p3Gp4A3p5A3'p6C3') (P1)
R = 0.963
δH = -0.828
(5b.13) r (C5'p1A5'p2A5p3Gp4A3p5A3'p6C3') (P4)
R = 0.962
δH = -0.922
(6b.12) r (C5'p1A5'p2A5p3Gp4C3p5A3'p6C3') (P2)
R = 0.983
δH = -0.802
(6b.13) r (C5'p1A5'p2A5p3Gp4C3p5A3'p6C3') (P3)
R = 0.988
δH = -0.933
(6b.14) r (C5'p1A5'p2A5p3Gp4C3p5A3'p6C3') (P4)
R = 0.989
δH = -0.954
(7a.10) d (C5'p1A5'p2C5p3Gp4A3p5A3'p6C3') (P4)
R = 0.981
δH = -1.047
(7b.10) r (C5'p1A5'p2C5p3Gp4A3p5A3'p6C3') (P2)
R = 0.969
δH = -0.815
(7b.11) r (C5'p1A5'p2C5p3Gp4A3p5A3'p6C3') (P3)
R = 0.979
δH = -0.744
(7b.12) r (C5'p1A5'p2C5p3Gp4A3p5A3'p6C3') (P4)
R = 0.922
δH = -1.050
(8a.9) d (C5'p1A5'p2C5p3Gp4C3p5A3'p6C3') (P2)
R = 0.986
δH = -1.354
(8a.10) d (C5'p1A5'p2C5p3Gp4C3p5A3'p6C3') (P3)
R = 0.984
δH = -1.084
(8b.12) r (C5'p1A5'p2C5p3Gp4C3p5A3'p6C3') (P2)
R = 0.968
δH = -1.190
(8b.13) r (C5'p1A5'p2C5p3Gp4C3p5A3'p6C3') (P3)
R = 0.967
δH = -1.068
(8b.14) r (C5'p1A5'p2C5p3Gp4C3p5A3'p6C3') (P4)
R = 0.979
δH = -1.192
δ31P
(δl - δexp) . aH+ . 10^14

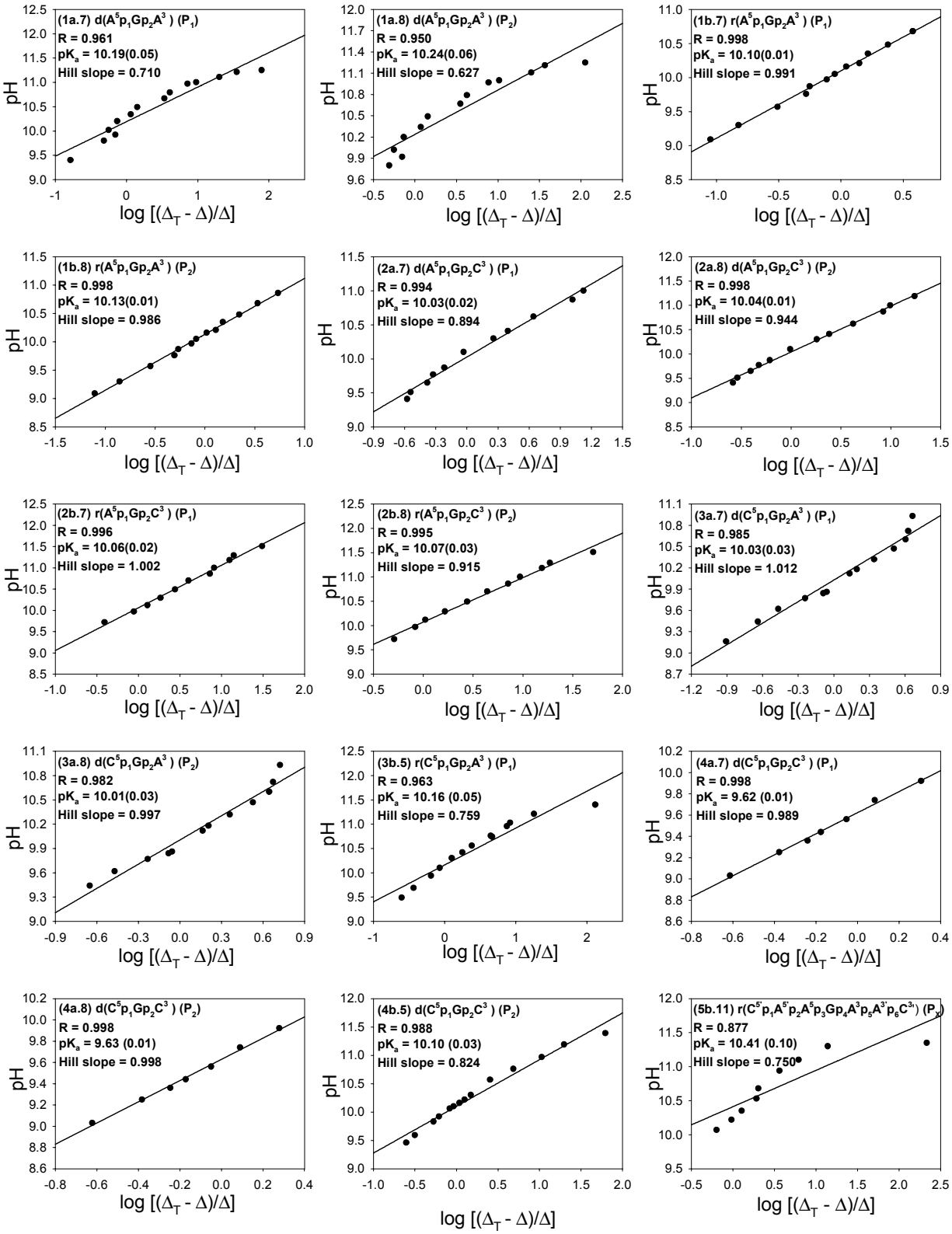
Figures S4: Graphical determination of ^{31}P chemical shift (δ_h) at the plateau at high pH for compounds **1 – 8**. The $\delta^{31}P$ is plotted against ($\delta_l - \delta_{exp}$) . aH^+ . 10^{14} (explained in experimental section) to obtain the theoretical chemical shift at highest pH for ^{31}P from the intercept of the correlation plot. Data points between pH 10 – 11.6 have been used in this plot for all ^{31}P resonances. The ^{31}P chemical shifts used here has been obtained by subtracting the average $\delta^{31}P$ of the 5' phosphorus (p_1) and 3' phosphorus (p_2) of Et$\underline{p_1}$Ap_2Et (**9b**) from the $\delta^{31}P$ shifts of different phosphates marker of the oligomers (**1 – 8**) Name of the compounds, value of the ^{31}P chemical shift at the highest pH and the correlation coefficient R has been shown in the respective graphs.

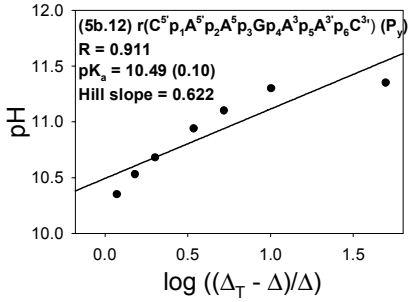
Figure S5.

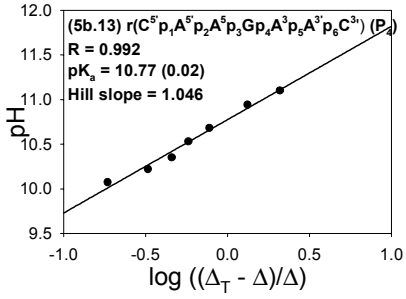


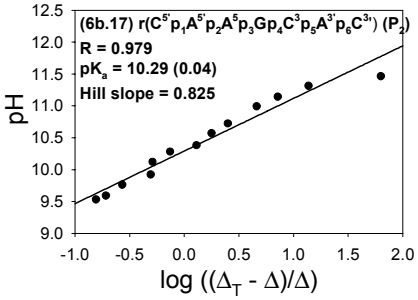
Figures S5: Graphical determination of ^{31}P chemical shift (δ_h) at the plateau at high pH for compounds **1 – 8**. The $\delta^{31}P$ is plotted against $(\delta_l - \delta_{exp})$. aH^+ . 10^{14} (explained in experimental section) to obtain the theoretical chemical shift at highest pH for ^{31}P from the intercept of the correlation plot. Data points between pH 10 – 11.9 have been used in this plot for all ^{31}P resonances. The ^{31}P chemical shifts used here has been obtained by subtracting the average $\delta^{31}P$ of the 5' phosphorus (p_1) and 3' phosphorus (p_2) of Etp_1Ap_2Et (**9b**) from the $\delta^{31}P$ shifts of different phosphates marker of the oligomers (**5 – 8**) Name of the compounds, value of the ^{31}P chemical shift at the highest pH and the correlation coefficient R has been shown in the respective graphs.

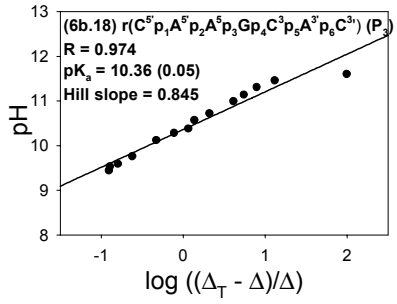
Figures S6:

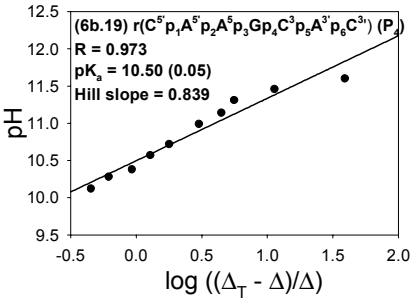

(1a.7) d(A5p1Gp2A3) (P1)
R = 0.961
pKa = 10.19(0.05)
Hill slope = 0.710
(1a.8) d(A5p1Gp2A3) (P2)
R = 0.950
pKa = 10.24(0.06)
Hill slope = 0.627
(1b.7) r(A5p1Gp2A3) (P1)
R = 0.998
pKa = 10.10(0.01)
Hill slope = 0.991
(1b.8) r(A5p1Gp2A3) (P2)
R = 0.998
pKa = 10.13(0.01)
Hill slope = 0.986
(2a.7) d(A5p1Gp2C3) (P1)
R = 0.994
pKa = 10.03(0.02)
Hill slope = 0.894
(2a.8) d(A5p1Gp2C3) (P2)
R = 0.998
pKa = 10.04(0.01)
Hill slope = 0.944
(2b.7) r(A5p1Gp2C3) (P1)
R = 0.996
pKa = 10.06(0.02)
Hill slope = 1.002
(2b.8) r(A5p1Gp2C3) (P2)
R = 0.995
pKa = 10.07(0.03)
Hill slope = 0.915
(3a.7) d(C5p1Gp2A3) (P1)
R = 0.985
pKa = 10.03(0.03)
Hill slope = 1.012
(3a.8) d(C5p1Gp2A3) (P2)
R = 0.982
pKa = 10.01(0.03)
Hill slope = 0.997
(3b.5) r(C5p1Gp2A3) (P1)
R = 0.963
pKa = 10.16 (0.05)
Hill slope = 0.759
(4a.7) d(C5p1Gp2C3) (P1)
R = 0.998
pKa = 9.62 (0.01)
Hill slope = 0.989
(4a.8) d(C5p1Gp2C3) (P2)
R = 0.998
pKa = 9.63 (0.01)
Hill slope = 0.998
(4b.5) d(C5p1Gp2C3) (P2)
R = 0.988
pKa = 10.10 (0.03)
Hill slope = 0.824
(5b.11) r(C5'p1A5'p2A5p3Gp4A3p5A3'p6C3') (Px)
R = 0.877
pKa = 10.41 (0.10)
Hill slope = 0.750
pH
log [(ΔT - Δ)/Δ]

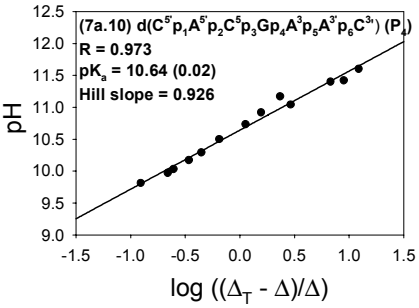

(5b.12) r(C5'p1A5'p2A5'p3Gp4A3'p5A3'p6C3'+) (Py)
R = 0.911
pKa = 10.49 (0.10)
Hill slope = 0.622
pH
log ((ΔT - Δ)/Δ)

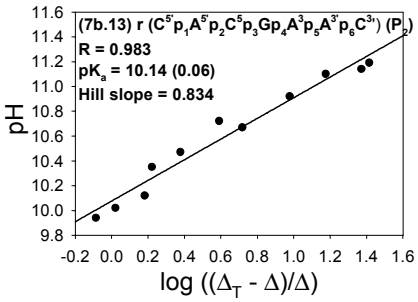

(5b.13) r(C5'p1A5'p2A5'p3Gp4A3'p5A3'p6C3'+) (P4)
R = 0.992
pKa = 10.77 (0.02)
Hill slope = 1.046
pH
log ((ΔT - Δ)/Δ)

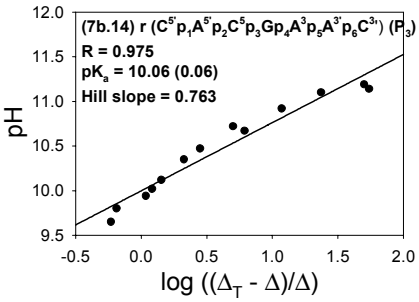

(6b.17) r(C5'p1A5'p2A5'p3Gp4C3'p5A3'p6C3'+) (P2)
R = 0.979
pKa = 10.29 (0.04)
Hill slope = 0.825
pH
log ((ΔT - Δ)/Δ)

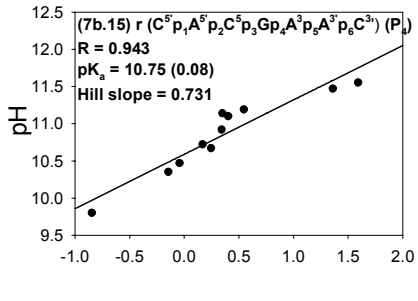

(6b.18) r(C5'p1A5'p2A5'p3Gp4C3'p5A3'p6C3'+) (P3)
R = 0.974
pKa = 10.36 (0.05)
Hill slope = 0.845
pH
log ((ΔT - Δ)/Δ)

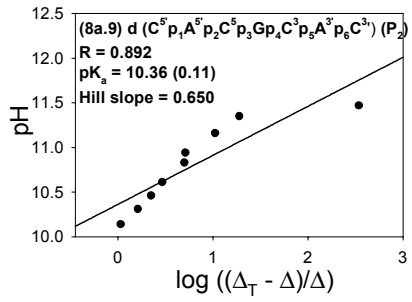

(6b.19) r(C5'p1A5'p2A5'p3Gp4C3'p5A3'p6C3'+) (P4)
R = 0.973
pKa = 10.50 (0.05)
Hill slope = 0.839
pH
log ((ΔT - Δ)/Δ)

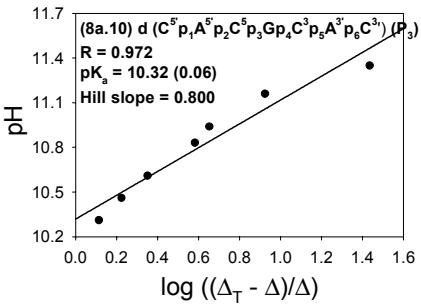

(7a.10) d(C5'p1A5'p2C5'p3Gp4A3'p5A3'p6C3'+) (P4)
R = 0.973
pKa = 10.64 (0.02)
Hill slope = 0.926
pH
log ((ΔT - Δ)/Δ)

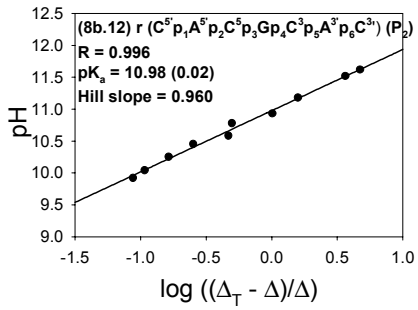

(7b.13) r (C5'p1A5'p2C5'p3Gp4A3'p5A3'p6C3'+) (P2)
R = 0.983
pKa = 10.14 (0.06)
Hill slope = 0.834
pH
log ((ΔT - Δ)/Δ)

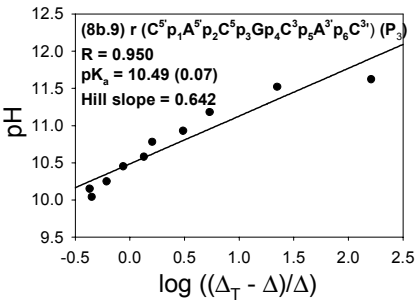

(7b.14) r (C5'p1A5'p2C5'p3Gp4A3'p5A3'p6C3'+) (P3)
R = 0.975
pKa = 10.06 (0.06)
Hill slope = 0.763
pH
log ((ΔT - Δ)/Δ)

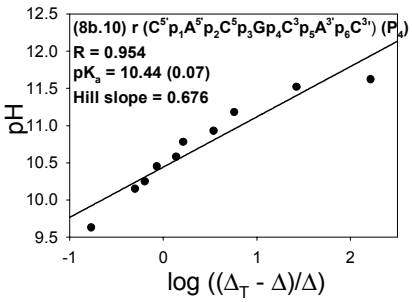

(7b.15) r (C5'p1A5'p2C5'p3Gp4A3'p5A3'p6C3'+) (P4)
R = 0.943
pKa = 10.75 (0.08)
Hill slope = 0.731
pH
log ((ΔT - Δ)/Δ)


(8a.9) d (C5'p1A5'p2C5'p3Gp4C3'p5A3'p6C3'+) (P2)
R = 0.892
pKa = 10.36 (0.11)
Hill slope = 0.650
pH
log ((ΔT - Δ)/Δ)


(8a.10) d (C5'p1A5'p2C5'p3Gp4C3'p5A3'p6C3'+) (P3)
R = 0.972
pKa = 10.32 (0.06)
Hill slope = 0.800
pH
log ((ΔT - Δ)/Δ)

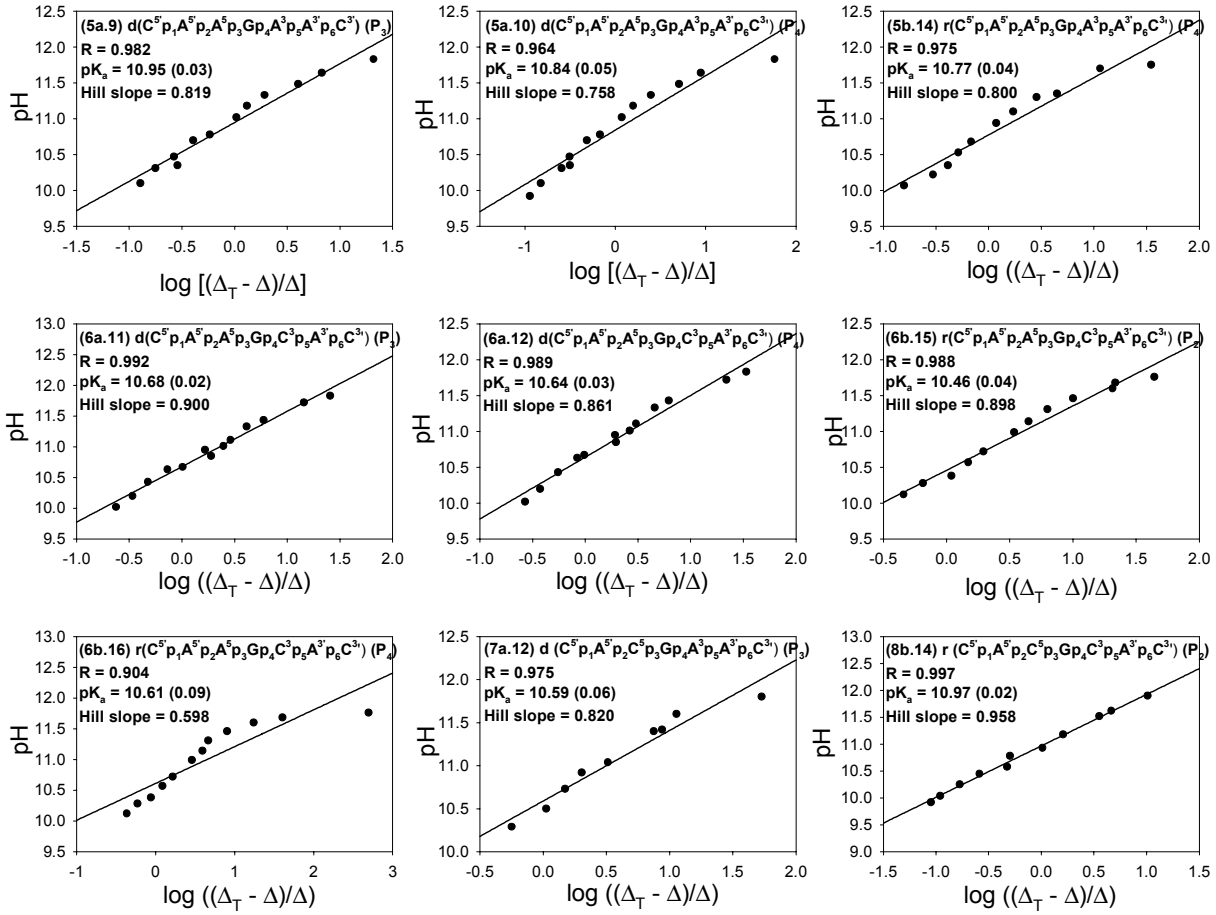

(8b.12) r (C5'p1A5'p2C5'p3Gp4C3'p5A3'p6C3'+) (P2)
R = 0.996
pKa = 10.98 (0.02)
Hill slope = 0.960
pH
log ((ΔT - Δ)/Δ)


(8b.9) r (C5'p1A5'p2C5'p3Gp4C3'p5A3'p6C3'+) (P3)
R = 0.950
pKa = 10.49 (0.07)
Hill slope = 0.642
pH
log ((ΔT - Δ)/Δ)


(8b.10) r (C5'p1A5'p2C5'p3Gp4C3'p5A3'p6C3'+) (P4)
R = 0.954
pKa = 10.44 (0.07)
Hill slope = 0.676
pH
log ((ΔT - Δ)/Δ)

Figures S6:. The Hill plots for compounds **1 - 8** The plot of log $((\Delta_T - \Delta)/\Delta)$ as a function of pH gave a straight line with Hill slope as slope and pK_a as intercept of the straight line. The ^{31}P chemical shifts used for calculating log $((\Delta_T - \Delta)/\Delta)$ (experimental section for details) here has been obtained by subtracting the average δ^{31}P of the 5' phosphorus (p_1) and 3' phosphorus (p_2) of Et$\underline{p_1}$Ap_2Et (9b) from the δ^{31}P shifts of different phosphates marker of the oligomers (**1** – **8**). Data points between pH 10 – 11.6 have been used in these plots for all ^{31}P resonances. Name of the compounds, values for the correlation coefficient (R), pK_a obtained from Hill plot analyses and the values of Hill slope are shown in the respective graphs.

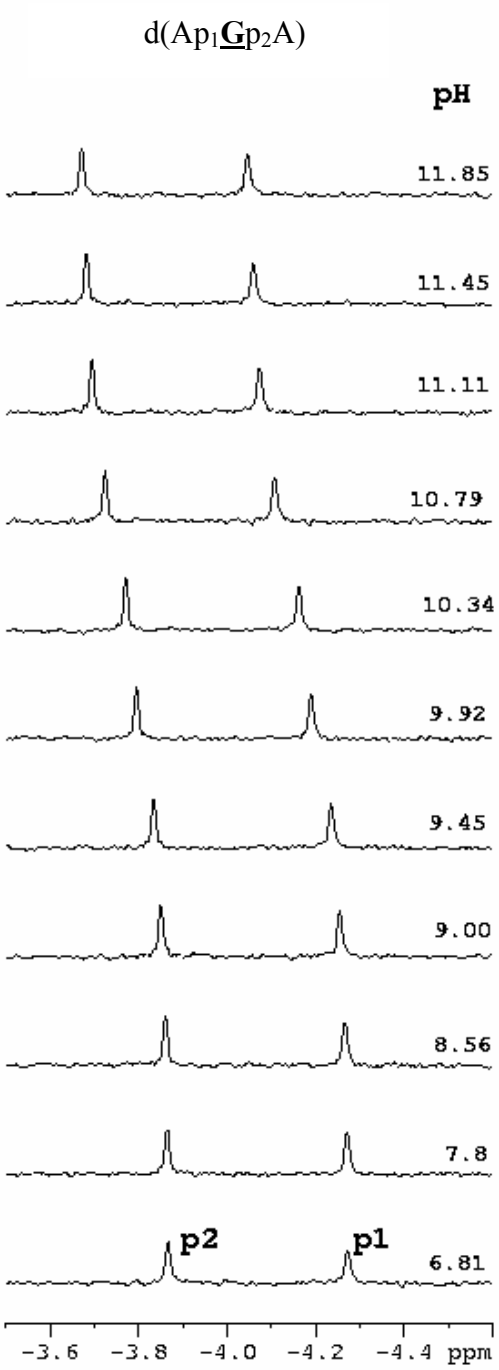
Figures S7:



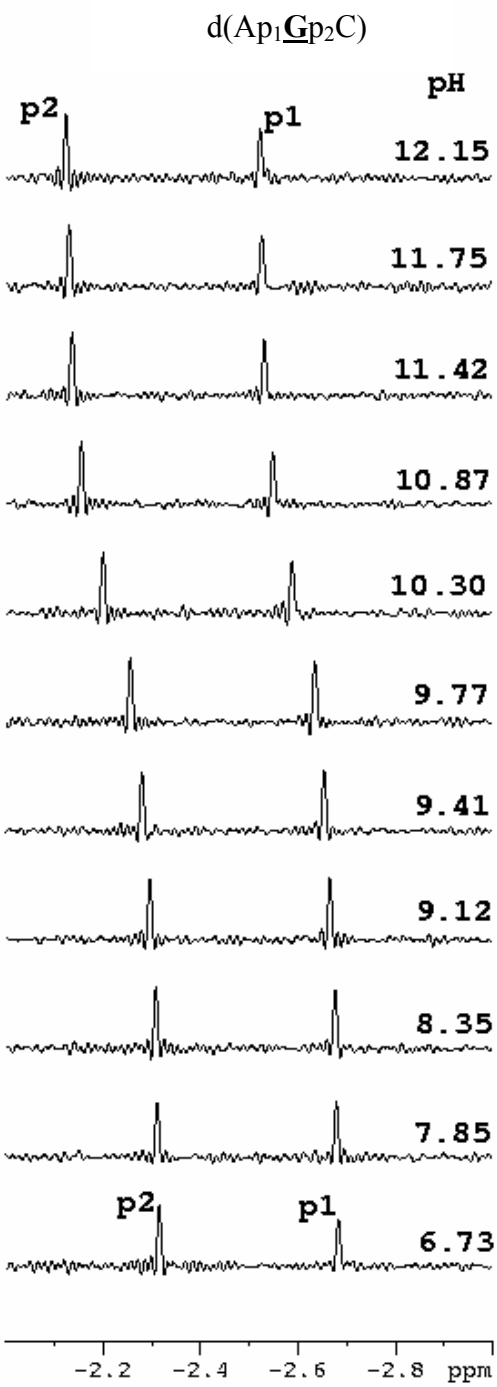
Figures S7: The Hill plots for compounds **1 - 8** The plot of log $((\Delta_T - \Delta)/\Delta)$ as a function of pH gave a straight line with Hill slope as slope and pK_a as intercept of the straight line. The ^{31}P chemical shifts used for calculating log $((\Delta_T - \Delta)/\Delta)$ (experimental section for details) here has been obtained by subtracting the average δ^{31}P of the 5' phosphorus (p_1) and 3' phosphorus (p_2) of Etp_1Ap_2Et (9b) from the δ^{31}P shifts of

different phosphates marker of the oligomers (**1** – **8**). Data points between pH 10 – 11.9 have been used in these plots for all ^{31}P resonances. Name of the compounds, values for the correlation coefficient (R), pK_a obtained from Hill plot analyses and the values of Hill slope are shown in the respective graphs.

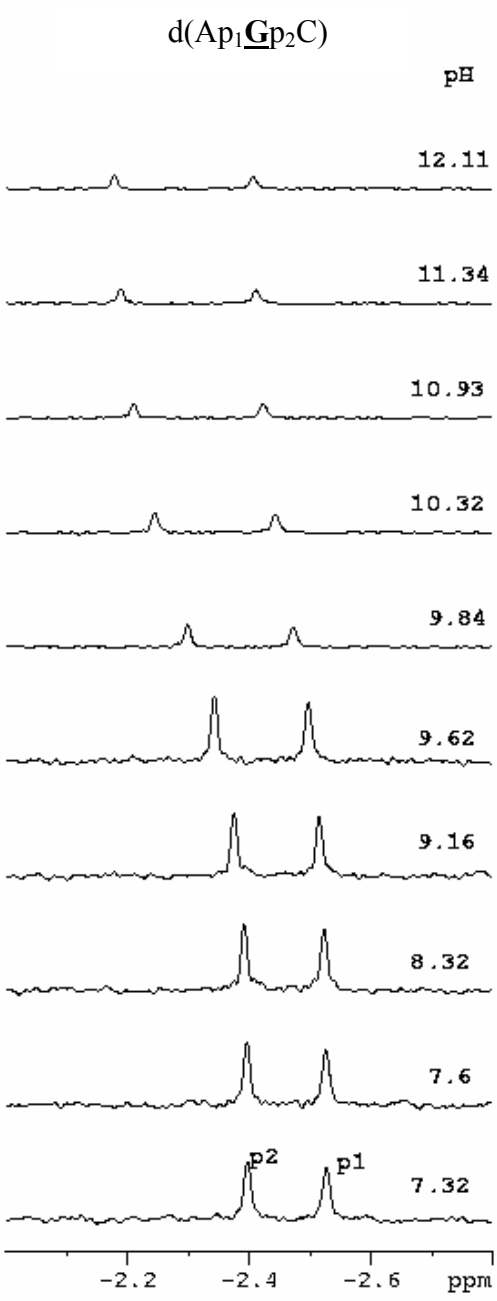
(A) The pH-dependent ^{31}P Chemical shift (in D_2O) of d(Ap$_1$**G**p$_2$A) (**1a**) at 298 K



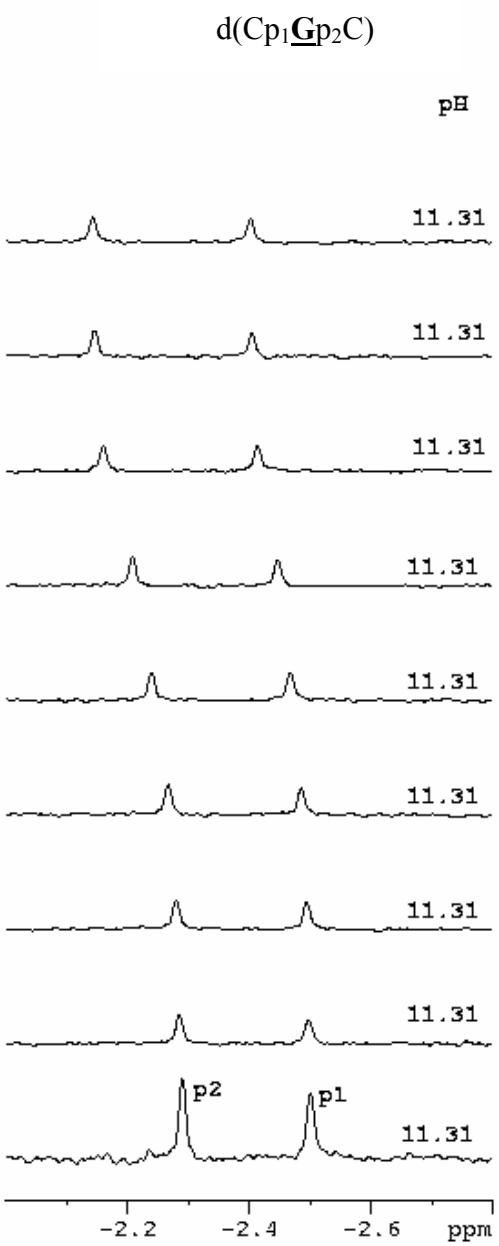
(B) The pH-dependent ^{31}P Chemical shift (in D_2O) of d($Ap_1\underline{G}p_2C$) (**2a**) at 298 K



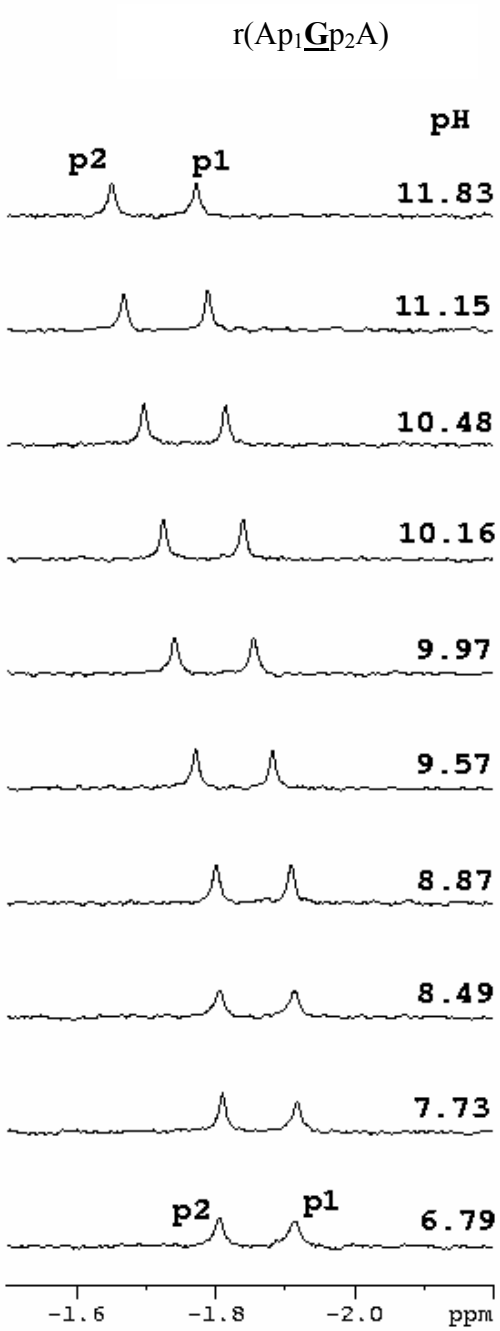
(C) The pH-dependent ^{31}P Chemical shift (in D_2O) of d(Cp$_1$**G**p$_2$A) (**3a**) at 298 K



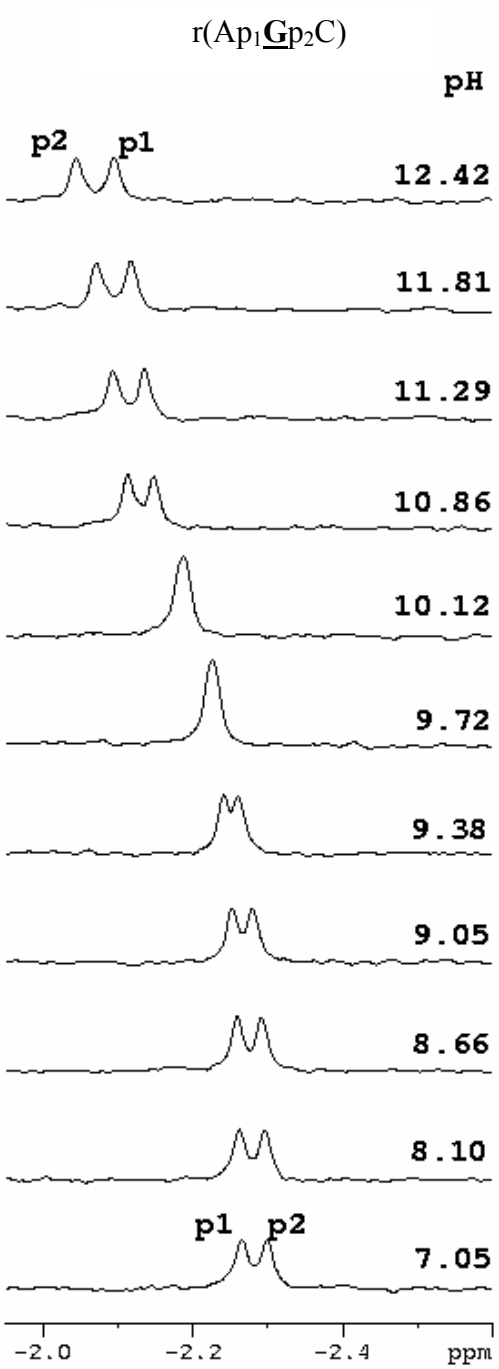
(D) The pH-dependent ^{31}P Chemical shift (in D_2O) of d(Cp_1**G**p_2C) (**4a**) at 298 K



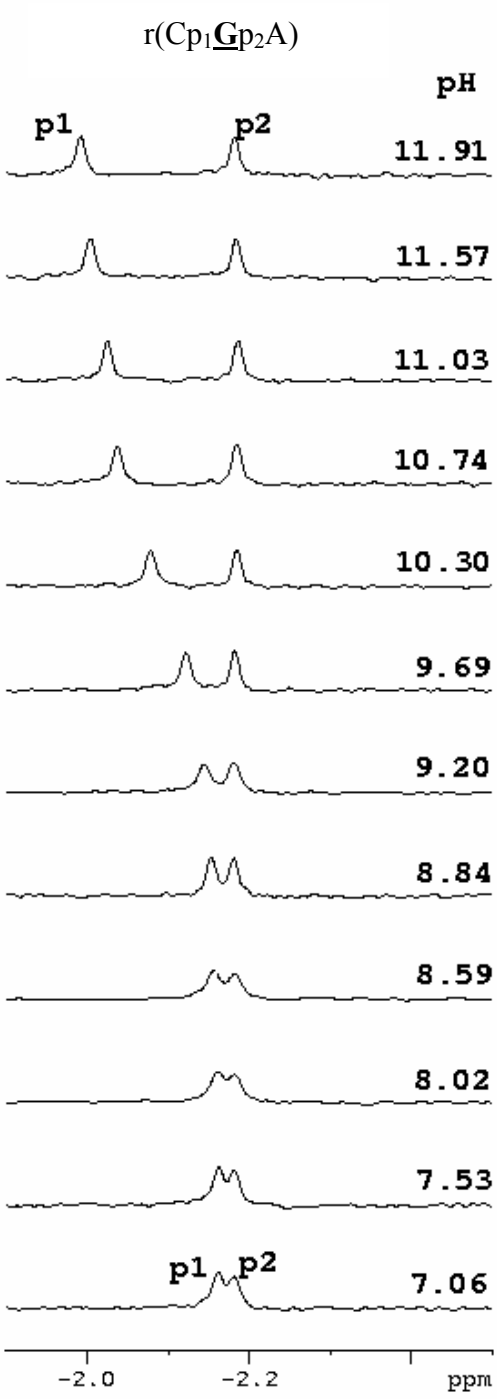
(E) The pH-dependent ^{31}P Chemical shift (in D_2O) of r(Ap_1**G**p_2A) (**1b**) at 298 K



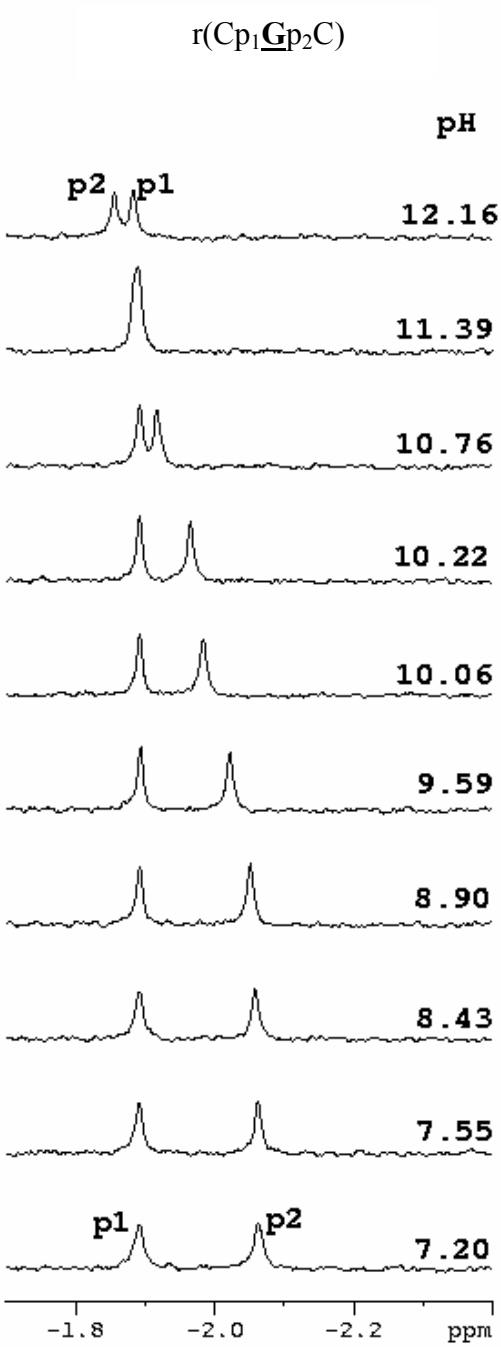
(F) The pH-dependent ^{31}P Chemical shift (in D_2O) of r(Ap$_1$**G**p$_2$C) (**2b**) at 298 K



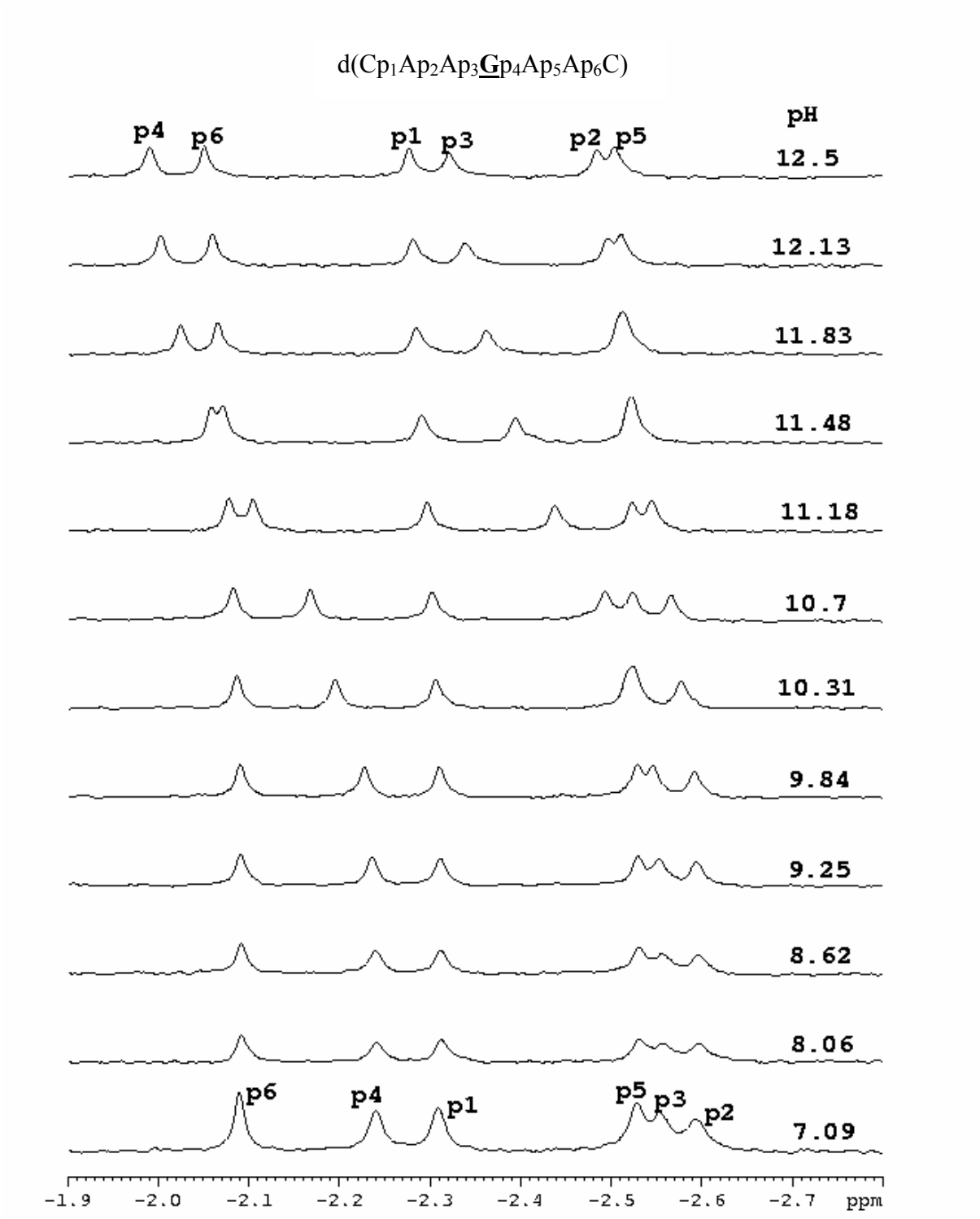
(G) The pH-dependent ^{31}P Chemical shift (in D_2O) of r(Cp_1**G**p_2A) (**3b**) at 298 K



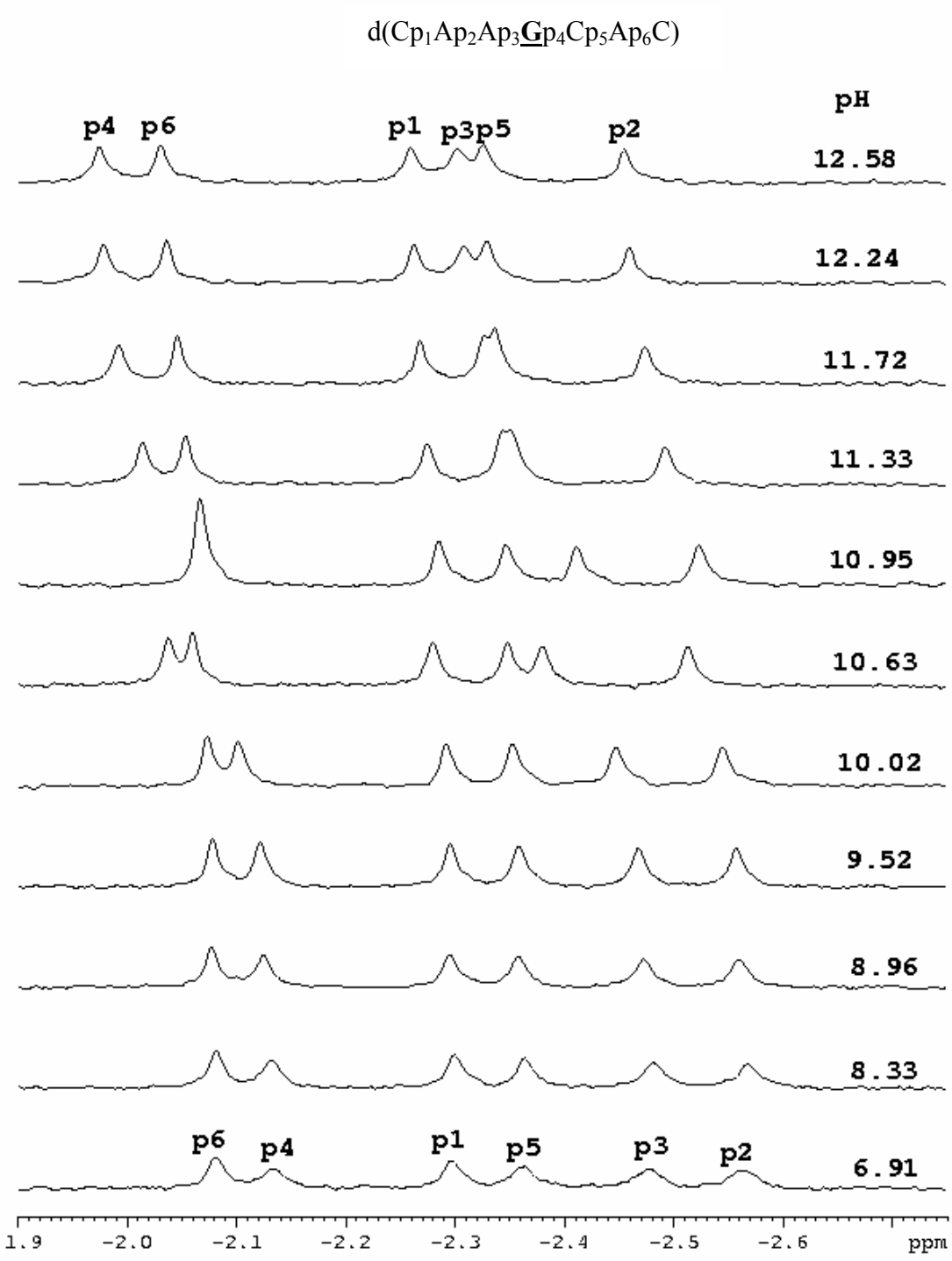
(H) The pH-dependent ^{31}P Chemical shift (in D_2O) of r(Cp_1**G**p_2C) (**4b**) at 298 K



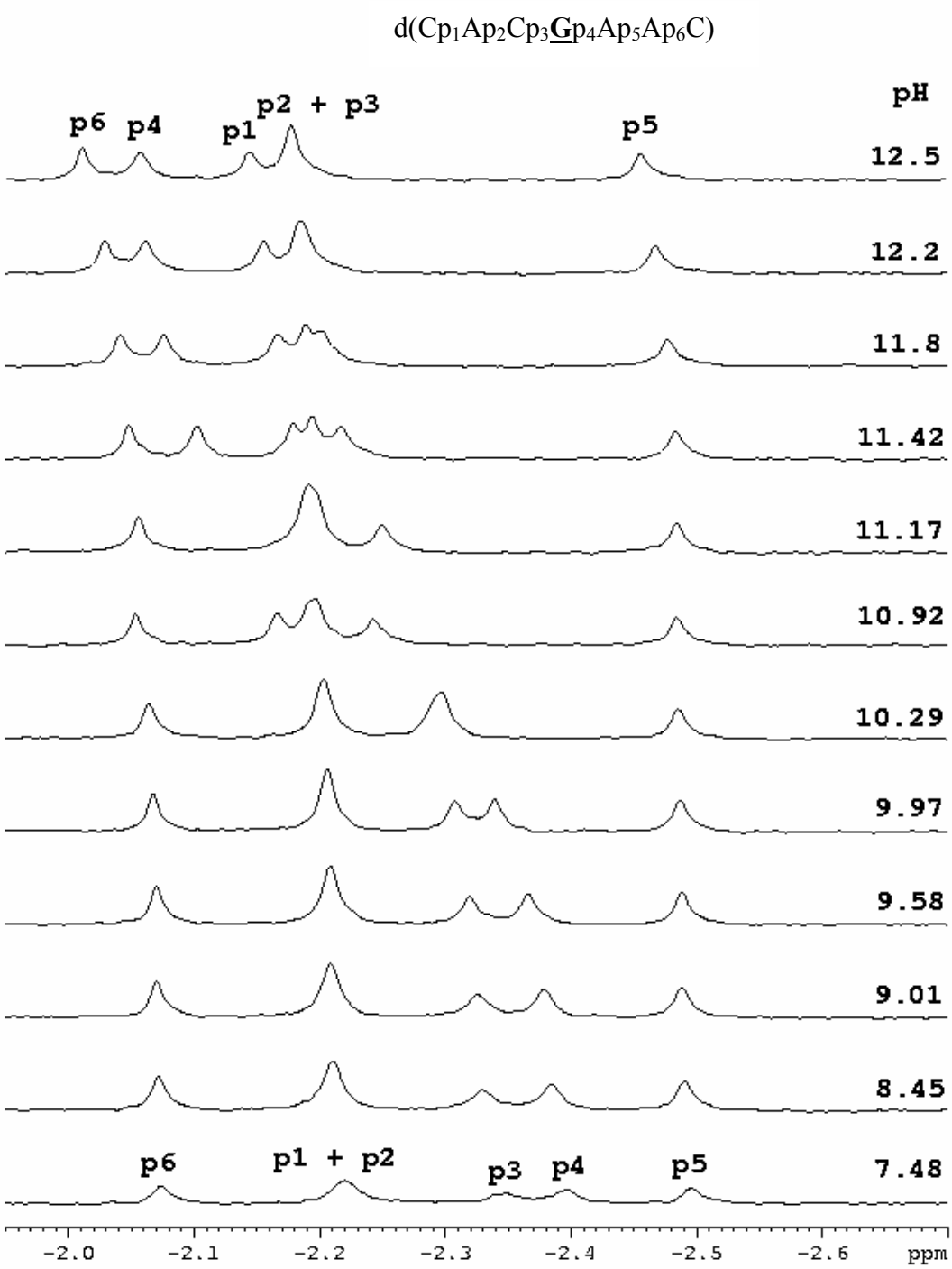
(I) The pH-dependent ^{31}P Chemical shift (in D_2O) of d($Cp_1Ap_2Ap_3\underline{G}p_4Ap_5Ap_6C$) (**5a**) at 298 K



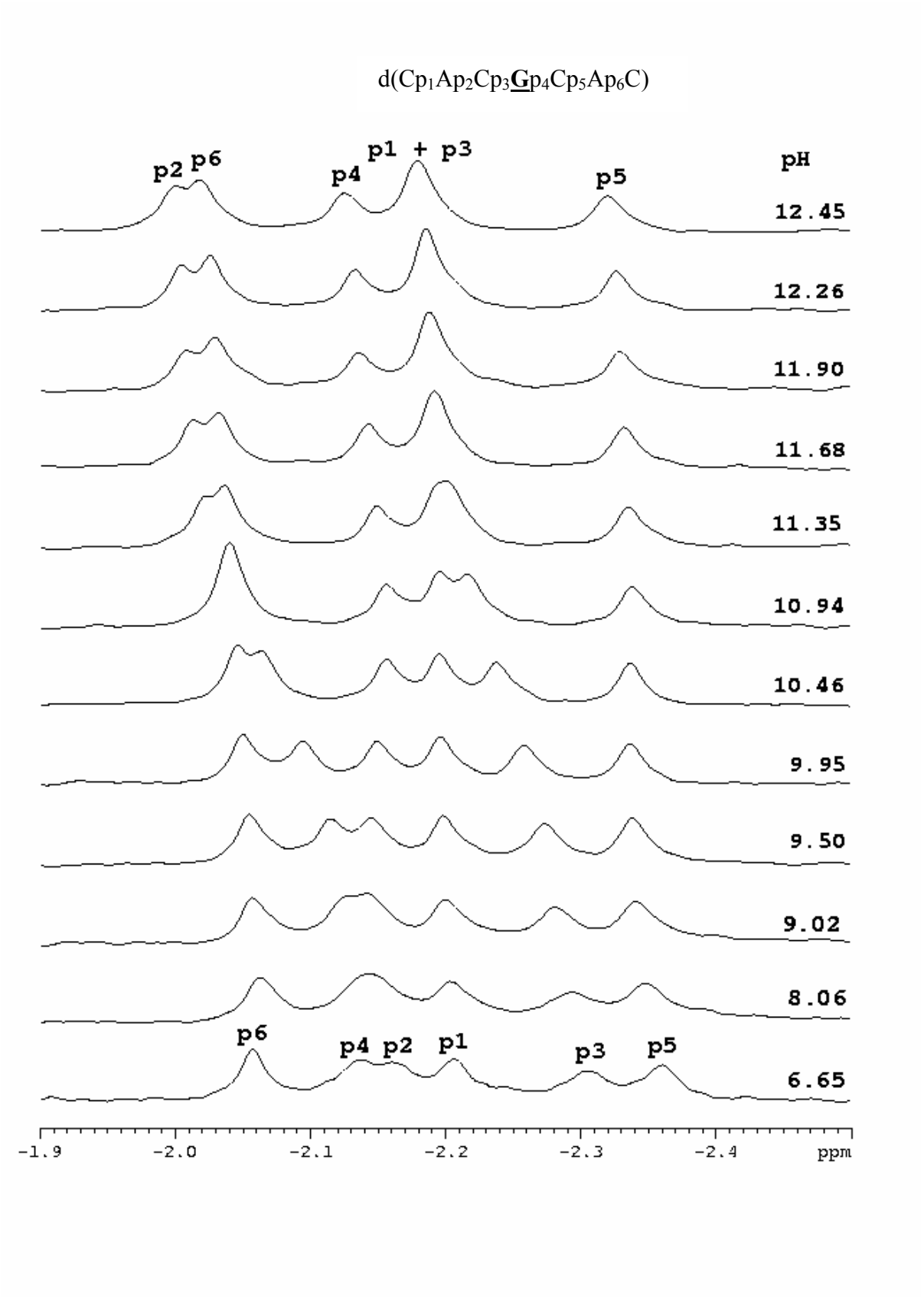
(J) The pH-dependent ^{31}P Chemical shift (in D_2O) of d($Cp_1Ap_2Ap_3\underline{G}p_4Cp_5Ap_6C$) (**6a**) at 298 K



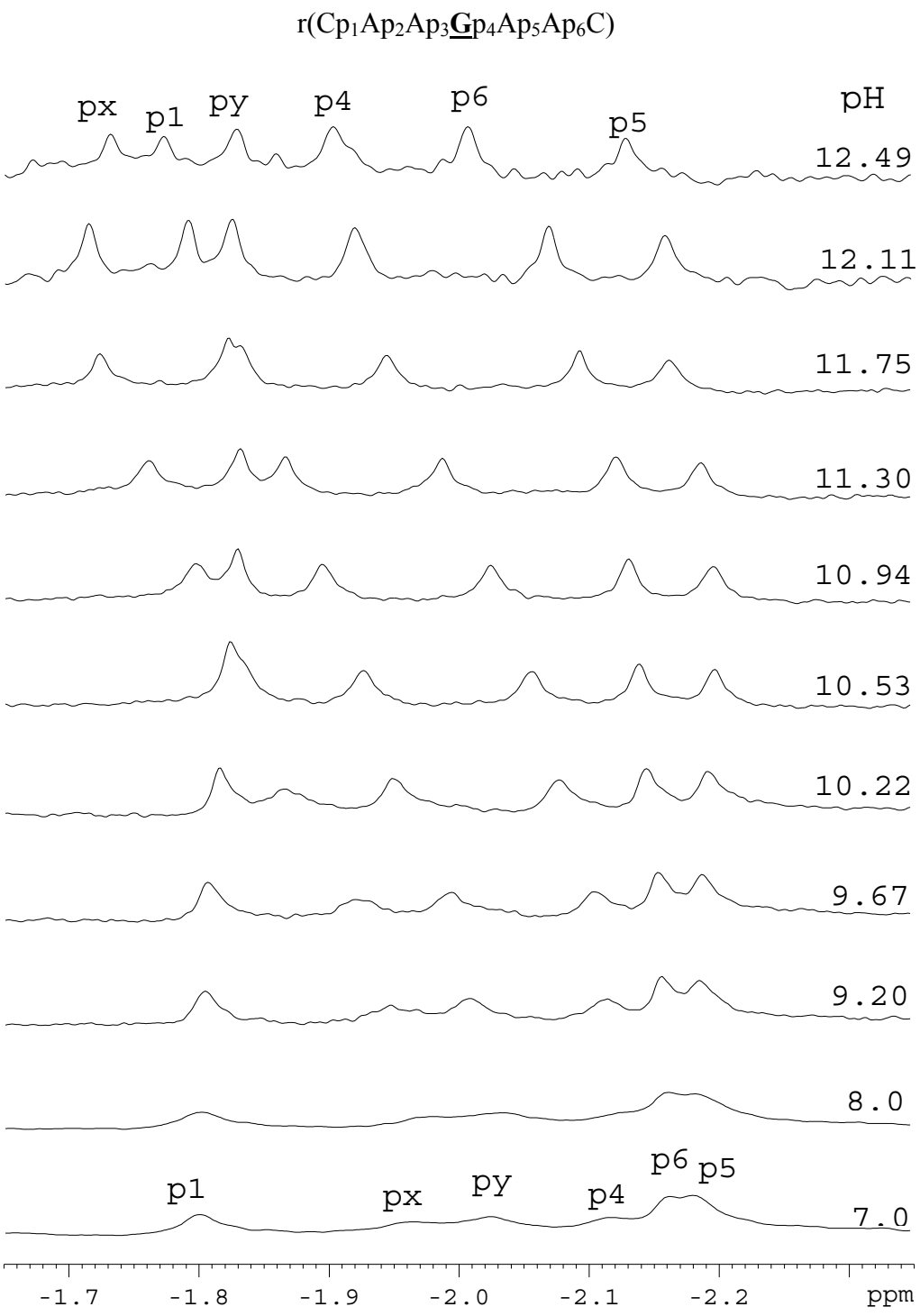
(K) The pH-dependent ^{31}P Chemical shift (in D_2O) of d($Cp_1Ap_2Cp_3\underline{G}p_4Ap_5Ap_6C$) (**7a**) at 298 K



(L) The pH-dependent ^{31}P Chemical shift (in D_2O) of d($Cp_1Ap_2Cp_3\underline{G}p_4Cp_5Ap_6C$) (**8a**) at 298 K



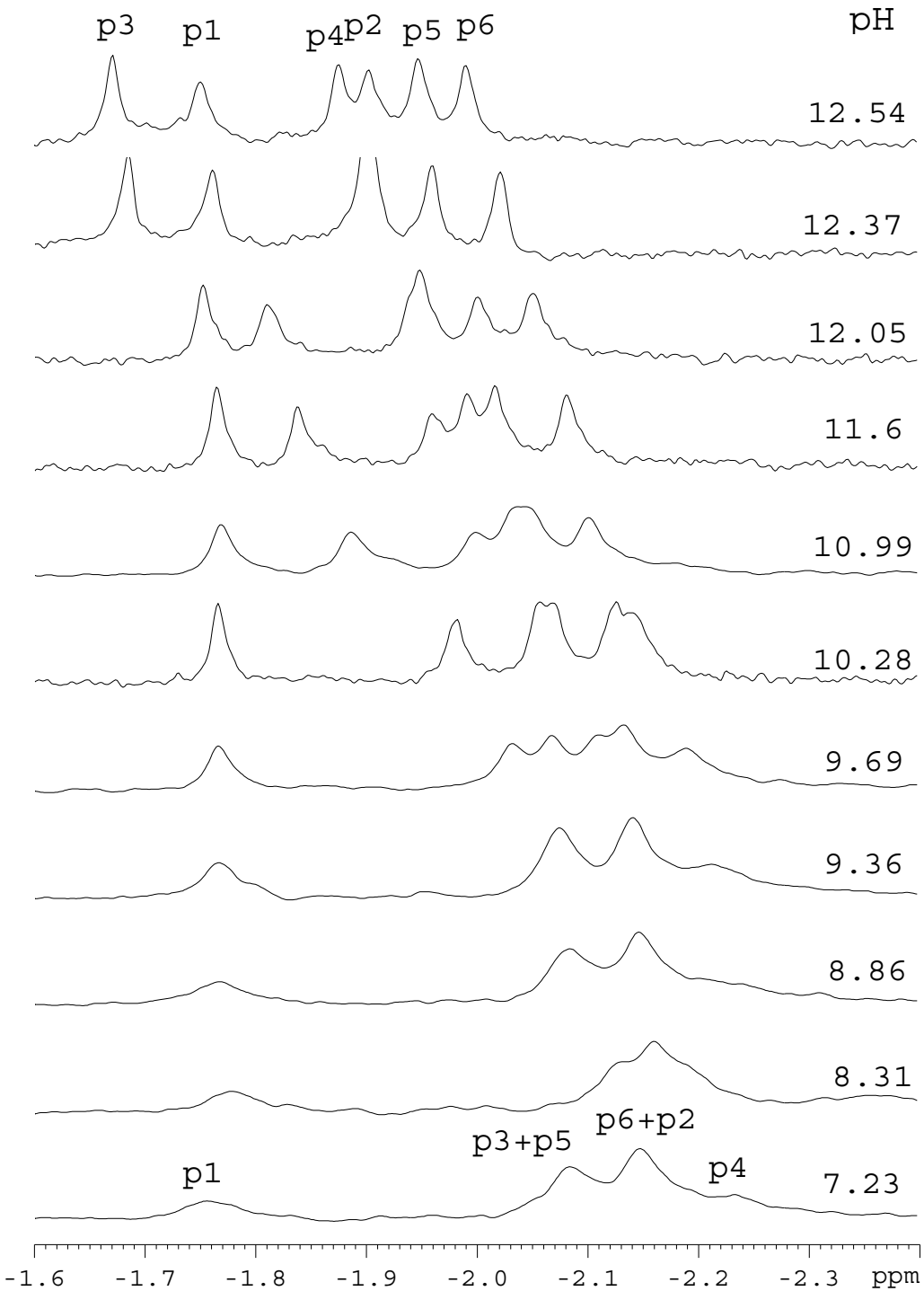
(M) The pH-dependent ^{31}P Chemical shift (in D_2O) of r($Cp_1Ap_2Ap_3\underline{\mathbf{G}}p_4Ap_5Ap_6C$) (**5b**) at 298 K



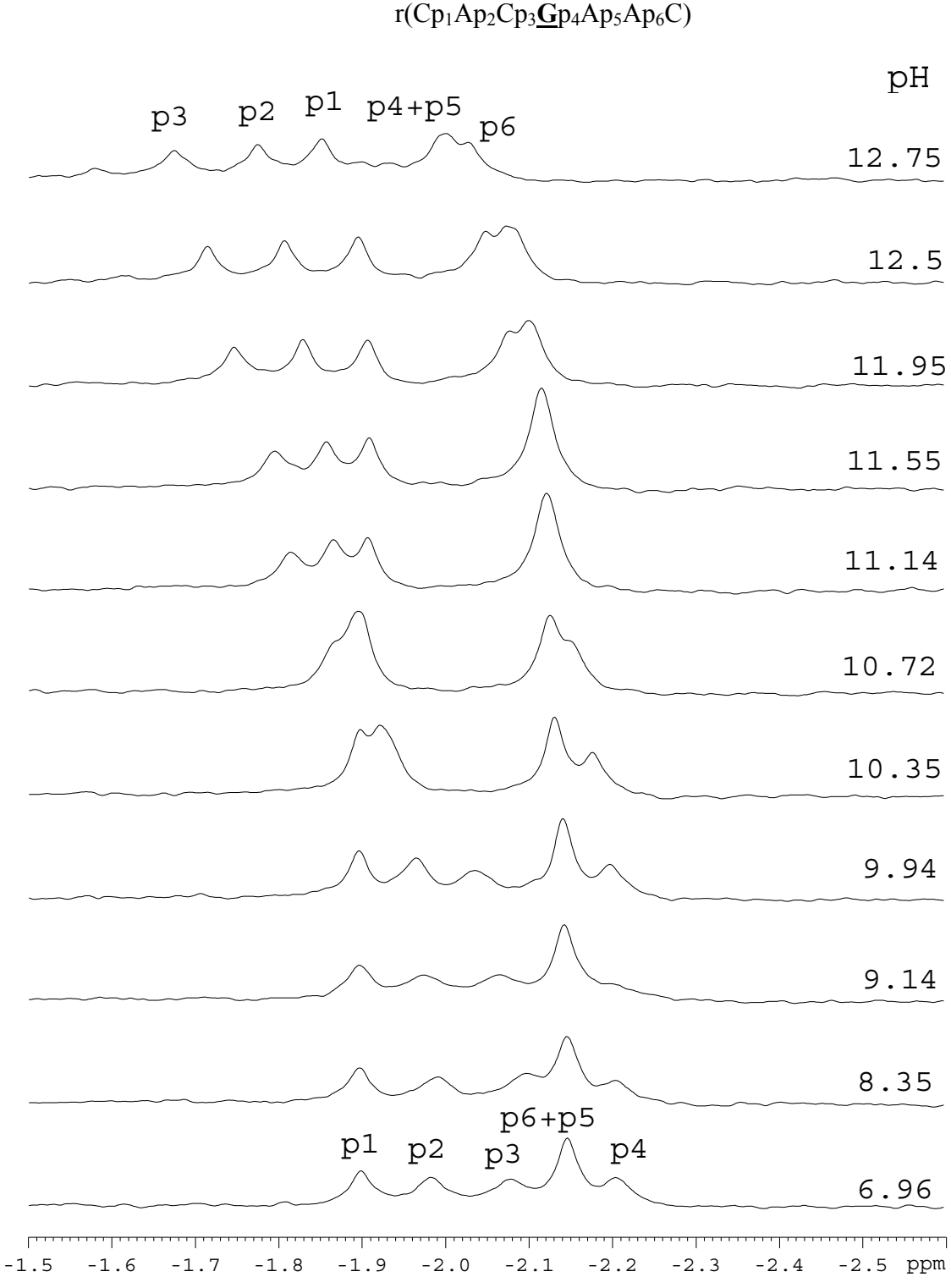
pX and **pY** are either p2 or p3, but could not be assigned exactly due to unavailability of $^{1}H - ^{31}P$ cross peaks. See the footnote in Table 1.

(N) The pH-dependent ^{31}P Chemical shift (in D_2O) of r($Cp_1Ap_2Ap_3\underline{\mathbf{G}}p_4Cp_5Ap_6C$) (**6b**) at 298 K

r($Cp_1Ap_2Ap_3\underline{\mathbf{G}}p_4Cp_5Ap_6C$)



(O) The pH-dependent ^{31}P Chemical shift (in D_2O) of r($Cp_1Ap_2Cp_3\underline{\mathbf{G}}p_4Ap_5Ap_6C$) (**7b**) at 298 K



(P) The pH-dependent ^{31}P Chemical shift (in D_2O) of r($Cp_1Ap_2Cp_3$**G**$p_4Cp_5Ap_6C$) (**8b**) at 298 K

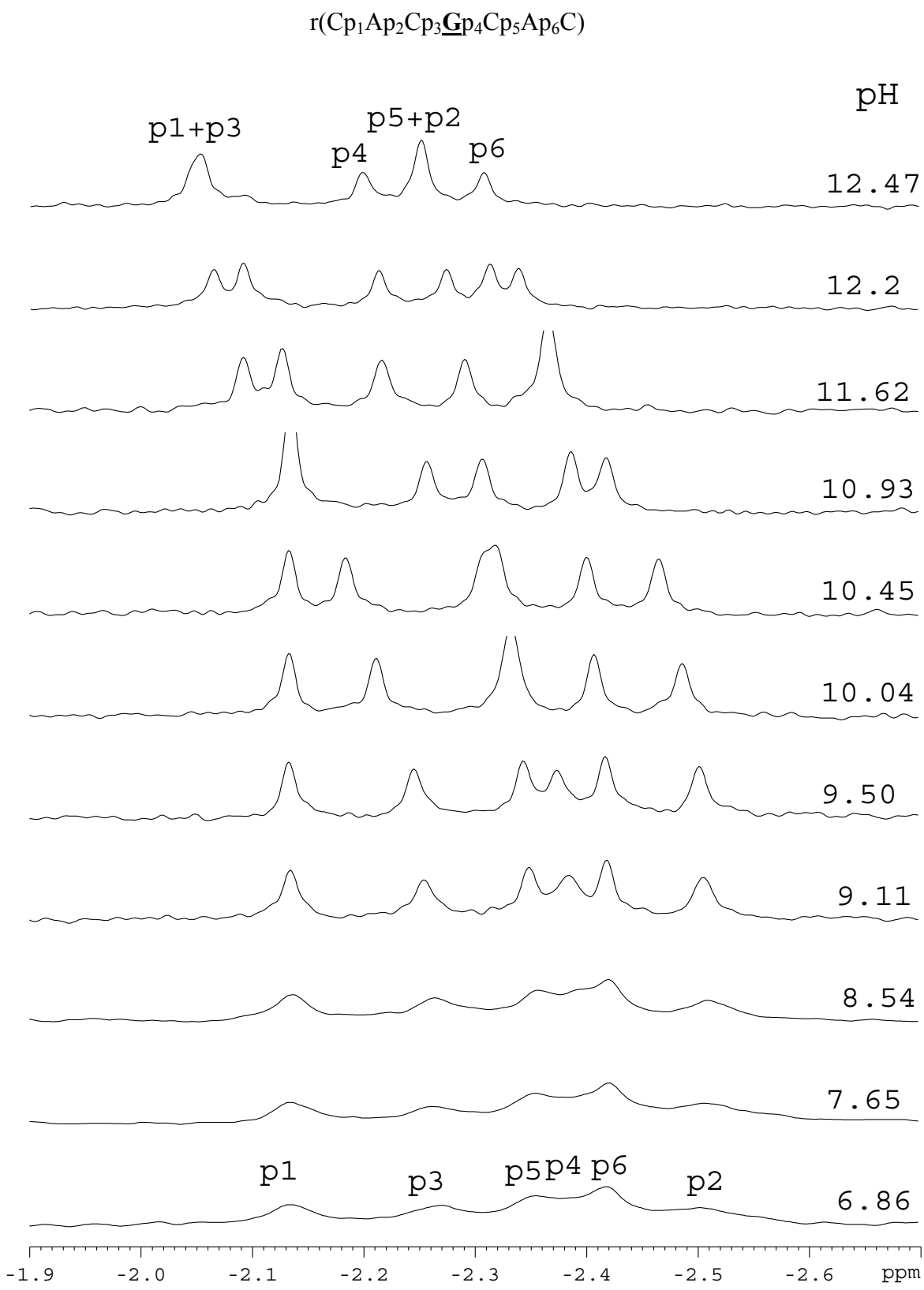
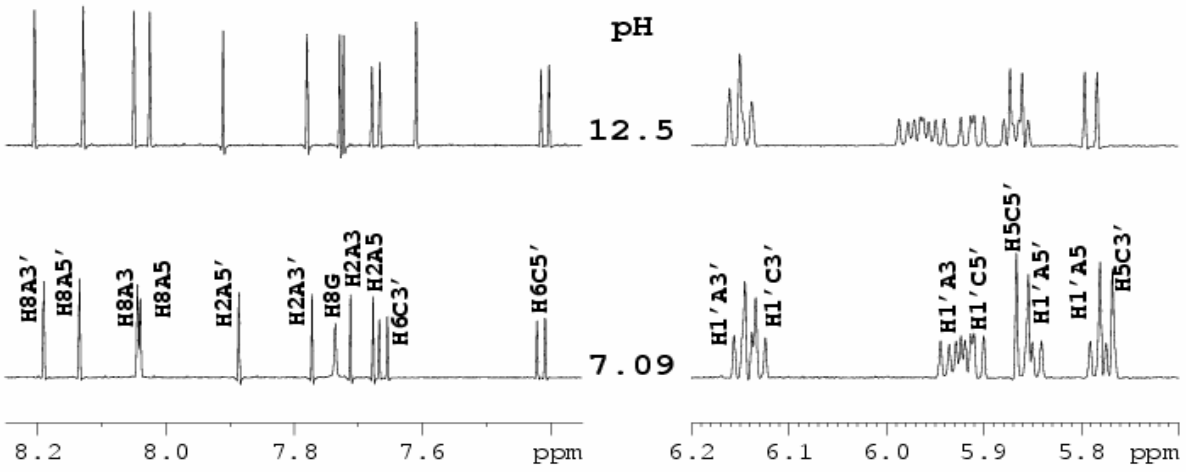
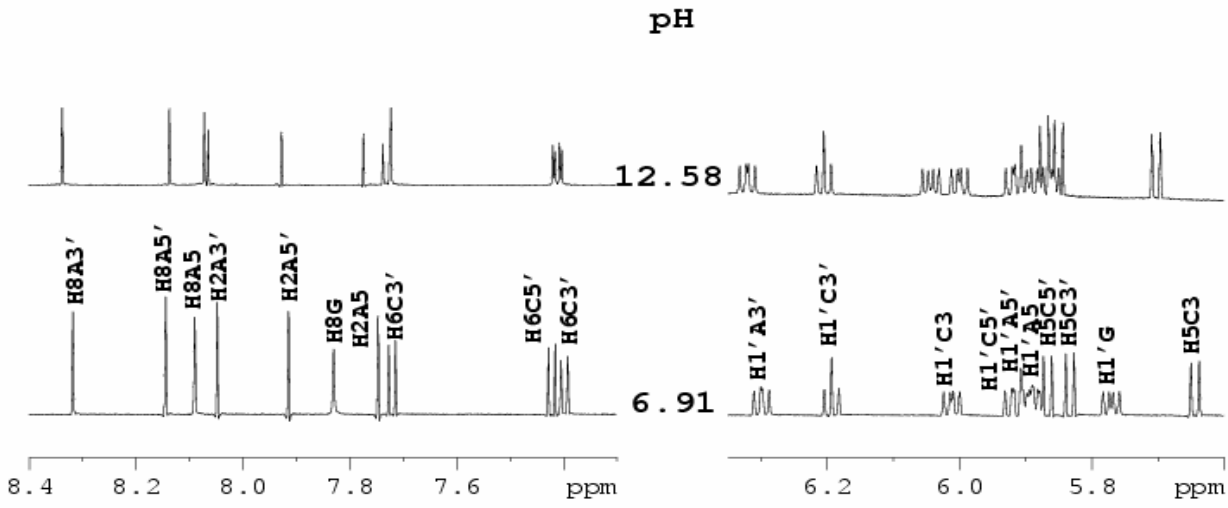


Figure S8. Panels (**A**) – (**P**) show stack plots of the pH-dependent ^{31}P NMR chemical shifts (δ^{31}P) in D_2O for isosequential ssDNAs (**1a** – **8a**) and ssRNA (**1b** – **8b**) at 298 K. It is noteworthy that peaks show artificial differences in line-widths because of different expansions of chemical shifts scale. It is only 10 – 12 pHs (including two plateaus at two extreme pHs) are shown out of total ~30 – 40 pHs (used for the titration plots) in panel (**A**) for d(A**G**A) (**1a**), panel (**B**) for d(A**G**C) (**2a**), panel (**C**) for d(C**G**A) (**3a**), panel (**D**) for d(C**G**C) (**4a**), panel (**E**) for r(A**G**A) (**1b**), panel (**F**) for r(A**G**C) (**2b**), panel (**G**) for r(C**G**A) (**3b**), panel (**H**) for r(C**G**C) (**4b**), panel (**I**) for d(CAA**G**AAC) (**5a**), panel (**J**) for d(CAA**G**CAC) (**6a**), panel (**K**) for d(CAC**G**AAC) (**7a**), panel (**L**) for d(CAC**G**CAC) (**8a**), (**M**) for r(CAA**G**AAC) (**5b**), panel (**N**) for r(CAA**G**CAC) (**6b**), panel (**O**) for r(CAC**G**AAC) (**7b**), panel (**P**) for r(CAC**G**CAC) (**8b**). All NMR measurements have been performed in 500 MHz and/or 600 MHz spectrometers. The pH values are corrected with the deuterium effect. The assignments for all compounds have been performed on the basis of ^{1}H NOESY, ^{1}H COSY, ^{31}P decoupled ^{1}H COSY, TOCSY and ^{31}P - ^{1}H correlation spectroscopy at 298 K and also ^{1}H NOESY at 283 K in the neutral pH[10a].

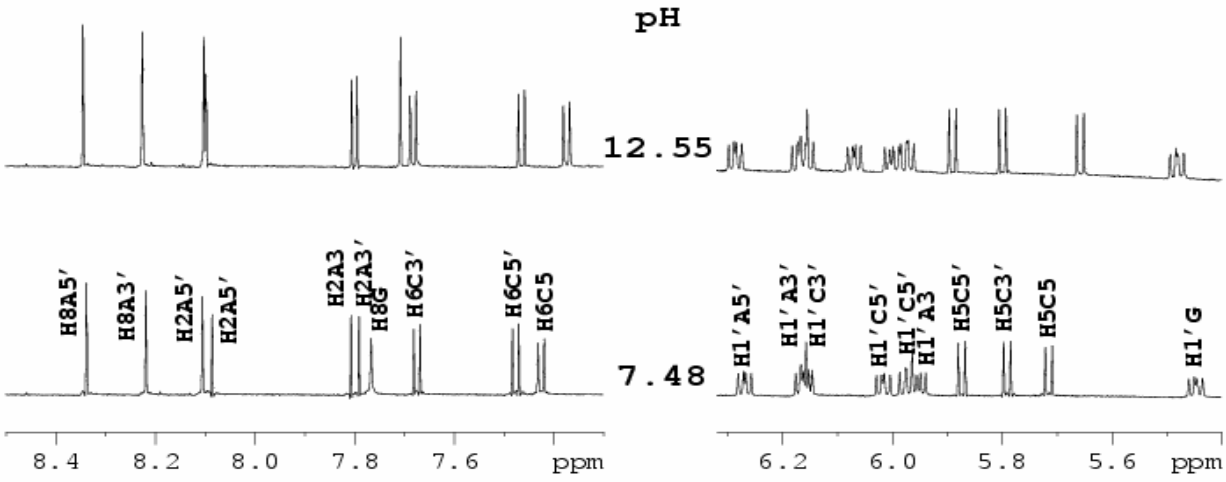
(A) The pH-dependent ^{1}H Chemical shift of aromatic (8.25-7.35 ppm) and anomeric (6.2-5.7 ppm) region (in D_2O) of d($C^{5'}p_1A^{5'}p_2A^{5}p_3\underline{G}p_4A^{3}p_5A^{3'}p_6C^{3'}$) (**5a**) at 298



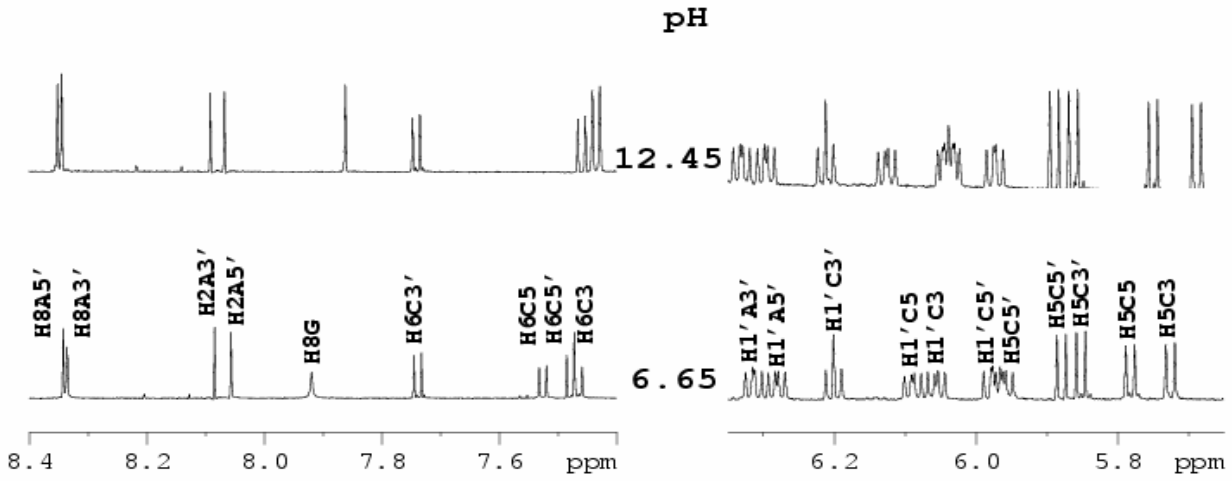
(B) The pH-dependent ^{1}H Chemical shift of aromatic (8.4-7.3 ppm) and anomeric (6.35-5.6 ppm) region (in D_2O) of d($C^{5'}p_1A^{5'}p_2A^{5}p_3\underline{G}p_4C^{3}p_5A^{3'}p_6C^{3'}$) (**6a**) at 298 K



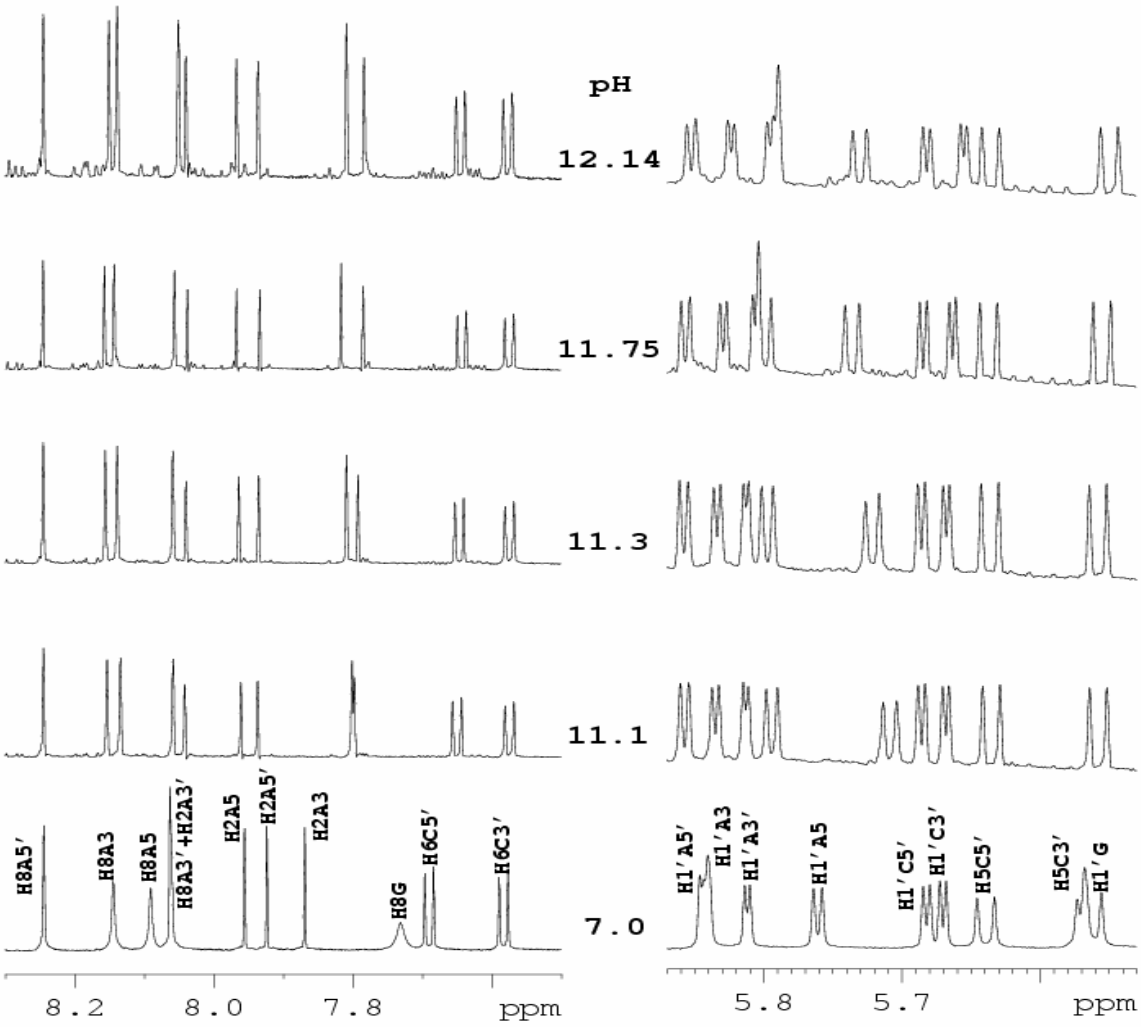
(C) The pH-dependent ^{1}H Chemical shift of aromatic (8.5-7.3 ppm) and anomeric (6.32-5.4 ppm) region (in D_2O) of d($C^{5'}p_1A^{5'}p_2C^{5}p_3\underline{G}p_4A^{3}p_5A^{3'}p_6C^{3'}$) (**7a**) at 298 K.



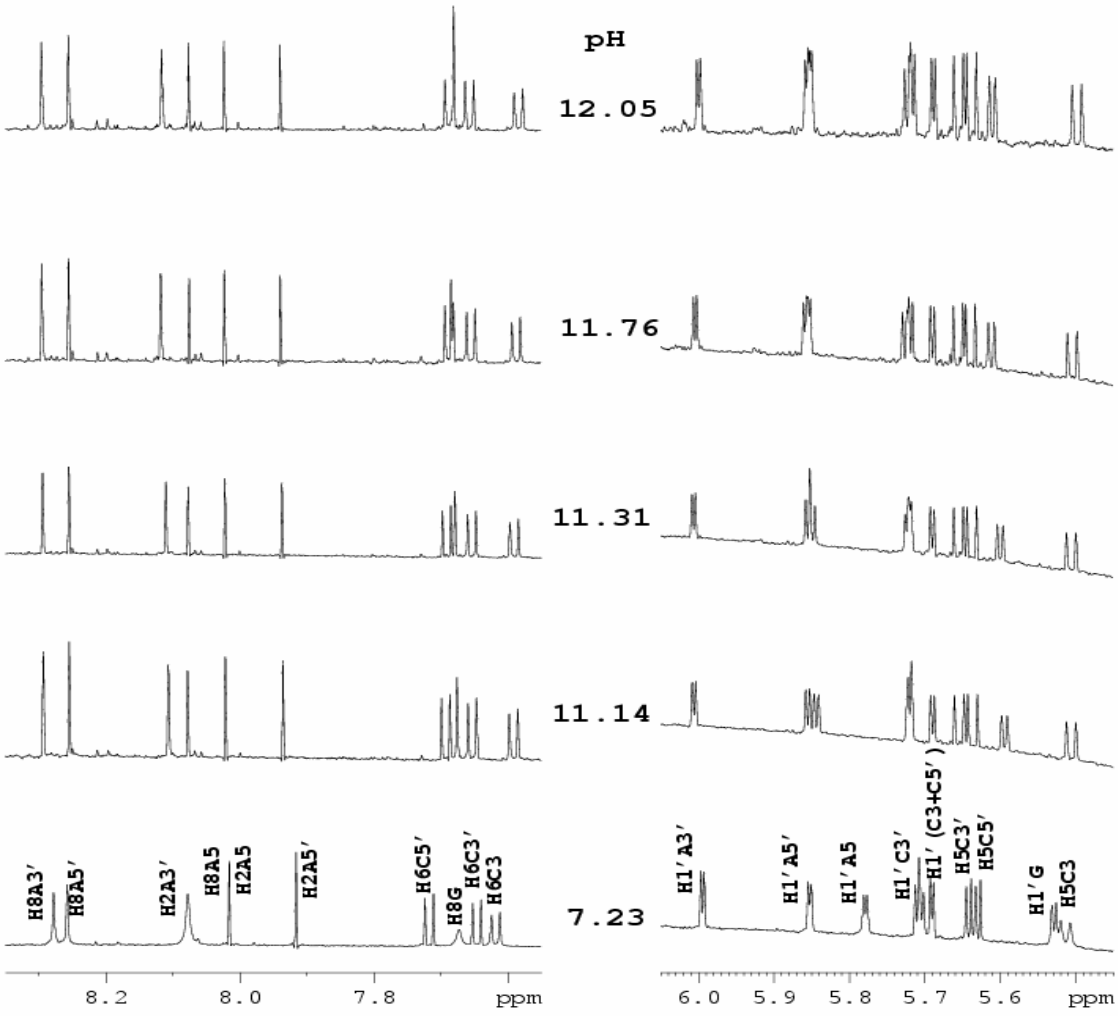
(D) The pH-dependent ^{1}H Chemical shift of aromatic (8.4-7.4 ppm) and anomeric (6.35-5.65 ppm) region (in D_2O) of d($C^{5'}p_1A^{5'}p_2C^{5}p_3\underline{G}p_4C^{3}p_5A^{3'}p_6C^{3'}$) (**8a**) at 298 K.



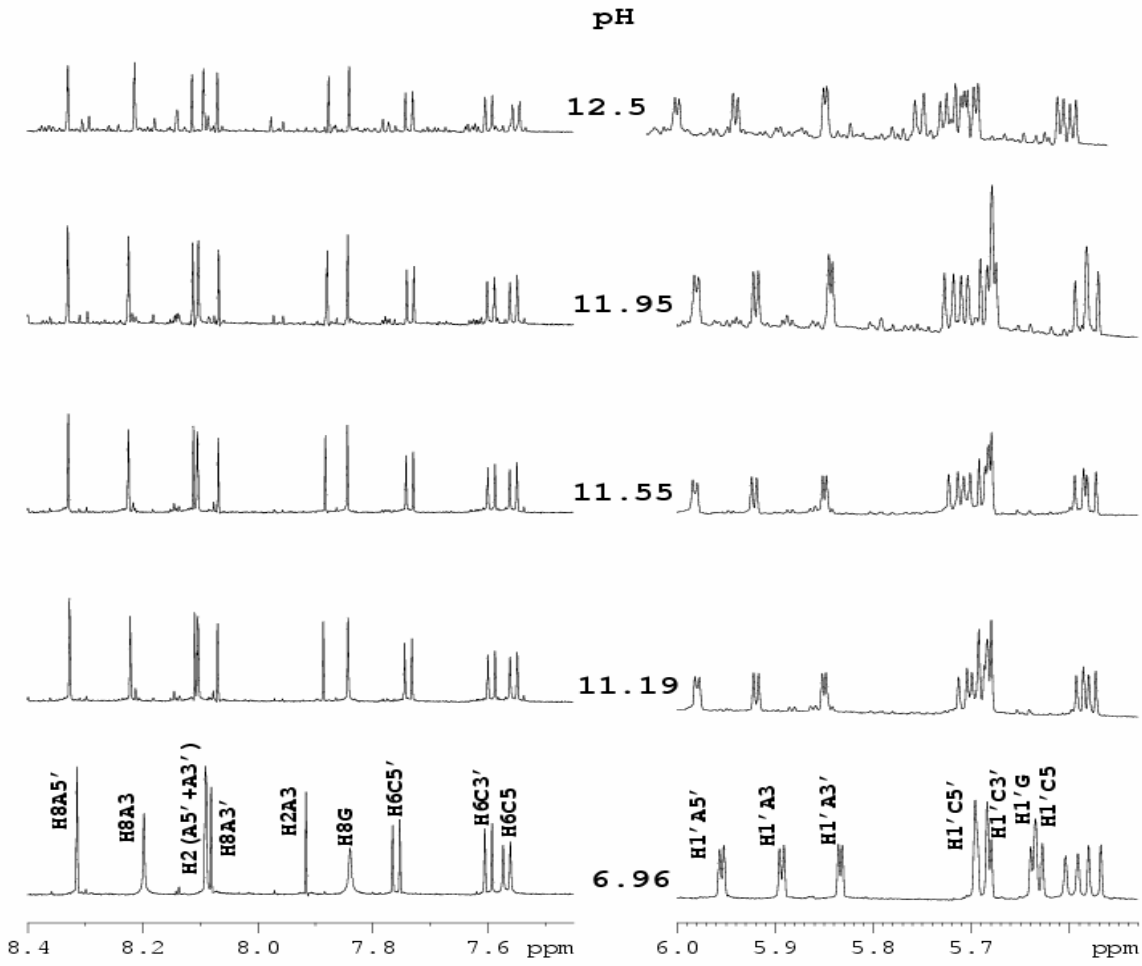
(E) The pH-dependent ^{1}H Chemical shift of aromatic (8.3-7.5 ppm) and anomeric (5.87-5.53 ppm) region (in D_2O) of r($C^{5'}p_1A^{5'}p_2A^{5}p_3\underline{G}p_4A^{3}p_5A^{3'}p_6C^{3'}$) (**5b**) at 298 K.



(F) The pH-dependent ^{1}H Chemical shift of aromatic (8.35-7.55 ppm) and anomeric (6.05-5.45 ppm) region (in D_2O) of r($C^{5'}p_1A^{5'}p_2A^{5}p_3\underline{G}p_4C^{3}p_5A^{3'}p_6C^{3'}$) (**6b**) at 298 K.



(G) The pH-dependent ^{1}H Chemical shift of aromatic (8.4-7.45 ppm) and anomeric (6-5.53 ppm) region (in D_2O) of r($C^{5'}p_1A^{5'}p_2C^{5}p_3\underline{\mathbf{G}}p_4A^{3}p_5A^{3'}p_6C^{3'}$) (**7b**) at 298 K.



(H) The pH-dependent ^{1}H Chemical shift of aromatic (8.4-7.5 ppm) and anomeric (6.05-5.52 ppm) region (in D_2O) of r($C^{5'}p_1A^{5'}p_2C^{5}p_3\underline{G}p_4C^{3}p_5A^{3'}p_6C^{3'}$) (**8b**) at 298 K.

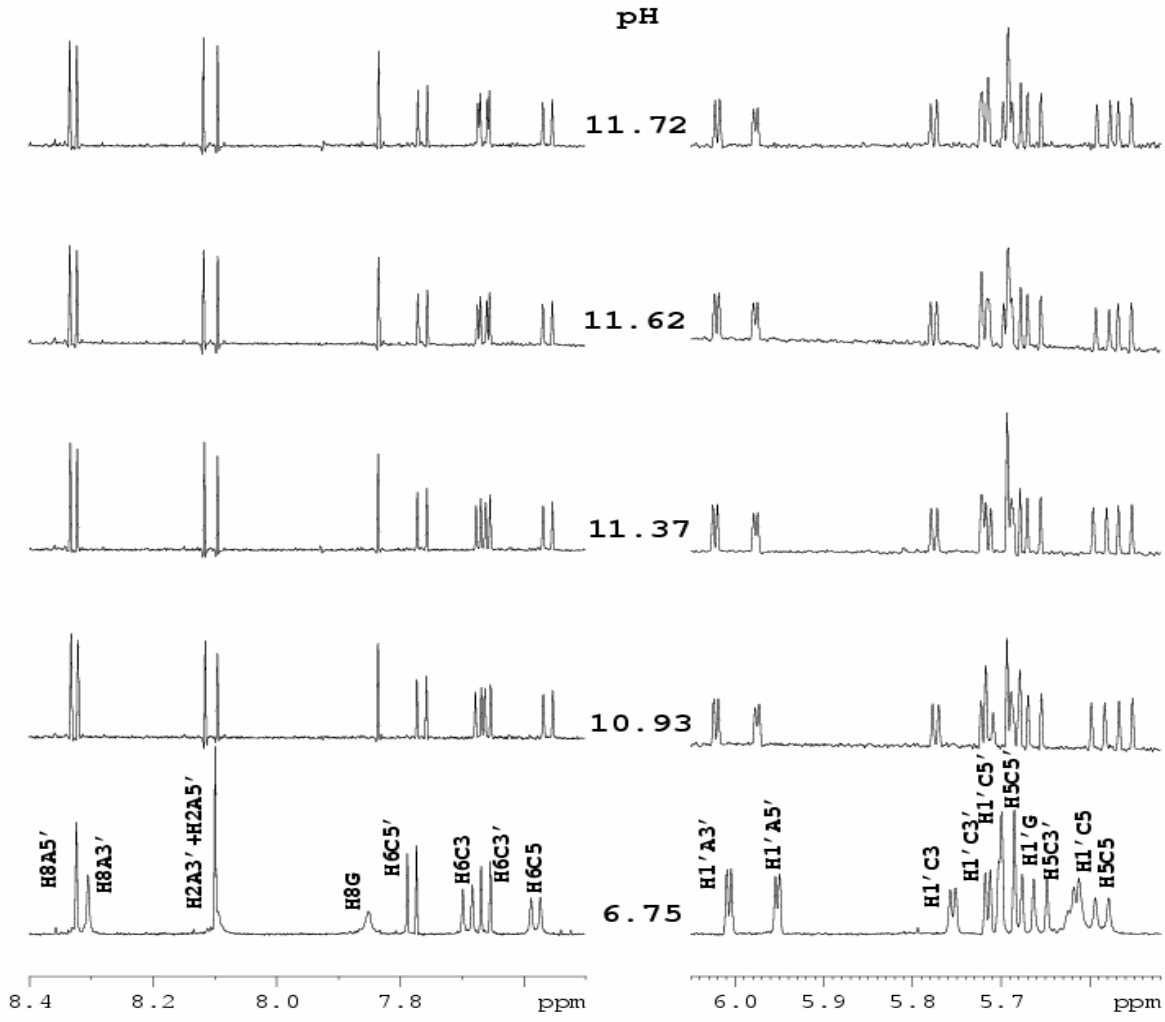


Figure S9. Panels (**A**) – (**H**) show stack plots of the pH-dependent ^{1}H NMR chemical shifts (δ^{1}H) in D_2O for isosequential ssDNAs (**5a** – **8a**) and ssRNA (**5b** – **8b**) at 298 K. The aromatic as well as anomeric regions both at the neutral and the alkaline pHs for (**5 - 8**) show that there is a small amount of decomposition only in case of heptameric ssRNAs (**5b – 8b**) at the alkaline pHs. The decomposition is not more than 2%-3% at the highest alkaline pH has been confirmed by integrating the aromatic and anomeric signals for the heptameric ssRNAs and the products formed. In case of heptameric ssDNAs (**5a – 8a**) as there is no scope of decomposition only the neutral and the alkaline pHs are shown. Panel (**A**) is for d($C^{5'}p_1A^{5'}p_2A^{5}p_3\underline{\mathbf{G}}p_4A^{3}p_5A^{3'}p_6C^{3'}$) (**5a**), panel (**B**) for d($C^{5'}p_1A^{5'}p_2A^{5}p_3\underline{\mathbf{G}}p_4C^{3}p_5A^{3'}p_6C^{3'}$) (**6a**), panel (**C**) for d($C^{5'}p_1A^{5'}p_2C^{5}p_3\underline{\mathbf{G}}p_4A^{3}p_5A^{3'}p_6C^{3'}$) (**7a**), panel (**D**) for d($C^{5'}p_1A^{5'}p_2C^{5}p_3\underline{\mathbf{G}}p_4C^{3}p_5A^{3'}p_6C^{3'}$) (**8a**), (**E**) for r($C^{5'}p_1A^{5'}p_2A^{5}p_3\underline{\mathbf{G}}p_4A^{3}p_5A^{3'}p_6C^{3'}$) (**5b**), panel (**F**) for r($C^{5'}p_1A^{5'}p_2^{5}p_3\underline{\mathbf{G}}p_4C^{3}p_5A^{3'}p_6C^{3'}$) (**6b**), panel (**G**) for r($C^{5'}p_1A^{5'}p_2C^{5}p_3\underline{\mathbf{G}}p_4A^{3}p_5A^{3'}p_6C^{3'}$) (**7b**), panel (**H**) for r($C^{5'}p_1A^{5'}p_2C^{5}p_3\underline{\mathbf{G}}p_4C^{3}p_5A^{3'}p_6C^{3'}$) (**8b**). All NMR measurements have been performed in 600 MHz spectrometers. The pH values are corrected with the deuterium effect. The assignments for all compounds have been performed on the basis of ^{1}H NOESY, ^{1}H COSY, ^{31}P decoupled ^{1}H COSY, TOCSY and ^{31}P - ^{1}H correlation spectroscopy at 298 K and also ^{1}H NOESY at 283 K in the neutral pH[8a].

Table S1. Comparison of pK_a of **G** obtained from H8G marker (pK_{a1}) and phosphorus markers (pK_{a2}) in ssDNA (**1a – 4a**) and ssRNA (**1b – 4b**)

Compound name		pK_a[#]	pK_a[*]	pK_a[∝]	average pK_a	pK_a from δH8G
d(Ap$_1$Gp$_2$A) **1a**	P_1	10.37 ± 0.04	10.33 ± 0.04	10.19 ± 0.05	10.30 ± 0.09	10.24 ± 0.01
	P_2	10.37 ± 0.04	10.32 ± 0.04	10.24 ± 0.06	10.31 ± 0.07	
r(Ap$_1$Gp$_2$A) **1b**	P_1	10.04 ± 0.02	10.06 ± 0.02	10.10 ± 0.01	10.07 ± 0.03	10.03 ± 0.01
	P_2	10.05 ± 0.03	10.07 ± 0.03	10.13 ± 0.01	10.08 ± 0.04	
d(Ap$_1$Gp$_2$C) **2a**	P_1	10.08 ± 0.02	10.08 ± 0.02	10.03 ± 0.02	10.06 ± 0.03	10.10 ± 0.01
	P_2	10.06 ± 0.02	10.07 ± 0.02	10.04 ± 0.01	10.06 ± 0.02	
r(Ap$_1$Gp$_2$C) **2b**	P_1	10.04 ± 0.02	10.07 ± 0.02	10.06 ± 0.02	10.06 ± 0.02	‡
	P_2	10.08 ± 0.01	10.09 ± 0.01	10.07 ± 0.03	10.08 ± 0.01	
d(Cp$_1$Gp$_2$A) **3a**	P_1	9.95 ± 0.02	9.93 ± 0.02	10.03 ± 0.03	9.97 ± 0.05	10.01 ± 0.01
	P_2	9.92 ± 0.02	9.93 ± 0.03	10.01 ± 0.03	9.95 ± 0.05	
r(Cp$_1$Gp$_2$A) **3b**	P_1	10.25 ± 0.01	10.24 ± 0.01	10.16 ± 0.05	10.22 ± 0.05	10.25 ± 0.01
	P_2	-	-	-	-	
d(Cp$_1$Gp$_2$C) **4a**	P_1	9.47 ± 0.02	9.51 ± 0.05	9.62 ± 0.01	9.53 ± 0.08	9.49 ± 0.01
	P_2	9.47 ± 0.01	9.51 ± 0.05	9.63 ± 0.01	9.54 ± 0.08	
r(Cp$_1$Gp$_2$C) **4b**	P_1	-	-	-	-	
	P_2	10.18 ± 0.02	10.17 ± 0.02	10.10 ± 0.03	10.15 ± 0.04	10.13 ± 0.01

All pK_a values and corresponding errors are calculated from non-linear curve-fitting analysis and Hill plot analyses (see experimental section D for details). pK_a[#] is obtained from non-linear curve fitting of pH vs. experimental δ^{31}P plot. pK_a[*] is obtained from non-linear curve fitting of pH vs. experimental δ^{31}P including the calculated δ^{31}P at deprotonated state using δ^{31}P values between pH 10 - 11.6 in the calculation of the chemical shift at deprotonated state. pK_a[∝] is obtained from Hill plot analysis using δ^{31}P values between pH 10 - 11.6 in the calculation of the chemical shift at deprotonated state as well as Hill plot analysis. ‡No titration profile found (see ref 8d).

Table S2. The endocyclic $\Sigma^3 J_{H1'}$ (in Hz) from ^{1}H NMR at 298 K for the respective pentofuranose moiety in each nucleotidic unit in **5a** – **8a** and **5b – 8b** in both the neutral (N) and the deprotonated (D) states. [a]

Oligo-ssDNA/ssRNA		$\Sigma^3 J_{1'}$ in the neutral (N) and deprotonated (D) states						
		$C^{5'}$	$A^{5'}$	A^5/C^5	G	A^3/C^3	$A^{3'}$	$C^{3'}$
$d(C^{5'}p_1A^{5'}p_2A^5p_3\underline{G}p_4A^3p_5A^{3'}p_6C^{3'})$ (**5a**)	**N**	14.2	15.4	15.3	14.0	15.0	14.2	13.4
	D	14.2	15.0	14.8	14.0	15.0	13.7	12.8
$r(C^{5'}p_1A^{5'}p_2A^5p_3\underline{G}p_4A^3p_5A^{3'}p_6C^{3'})$ (**5b**)	**N**	2.9	3.2	3.7	3.3	2.5	2.2	2.7
	D	3.1	2.8	2.6	6.0	3.8	4.8	2.6
$d(C^{5'}p_1A^{5'}p_2Ap_3\underline{G}p_4C^3p_5A^{3'}p_6C^{3'})$ (**6a**)	**N**	14.2	N.A	N.A	14.5	14.3	13.7	13.2
	D	14.2	N.A	N.A	14.5	14.5	13.9	13.2
$r(C^{5'}p_1A^{5'}p_2A^5p_3\underline{G}p_4C^3p_5A^{3'}p_6C^{3'})$ (**6b**)	**N**	2.5	2.5	2.8	3.4	3.5	2.5	3.1
	D	2.9	1.6	2.1	4.7	3.5	2.7	3.0
$d(C^{5'}p_1A^{5'}p_2C^5p_3\underline{G}p_4Ap_5A^{3'}p_6C^{3'})$ (**7a**)	**N**	13.9	14.2	14.4	14.9	15.1	14.7*	13.4*
	D	14.1	14.2	14.4	14.7	14.8	14.3*	13.6*
$r(C^{5'}p_1A^{5'}p_2C^5p_3\underline{G}p_4A^3p_5A^{3'}p_6C^{3'})$ (**7b**)	**N**	1.3	2.7	4.2	2.9	2.6	2.3	2.5
	D	2.5	2.6	5.6	4.0	3.0	2.1	2.9
$d(C^{5'}p_1A^{5'}p_2C^5p_3\underline{G}p_4C^3p_5A^{3'}p_6C^{3'})$ (**8a**)	**N**	13.9	14.2	14.3	14.5	14.4	13.8	13.2
	D	N.A	14.1	14.3	14.0	N.A	13.8	13.2
$r(C^{5'}p_1A^{5'}p_2C^5p_3\underline{G}p_4C^3p_5A^{3'}p_6C^{3'})$ (**8b**)	**N**	1.6	2.3	2.9	4.3	3.1	2.5	2.8
	D	2.5	2.4	2.9	4.3	3.5	2.6	3.5

[a] All spectra are taken in 600 MHz spectrometer. The error in measurement of $^3J_{1'2'}$ is ± 0.1 Hz.

* Coupling constant measured in 1D rows from the DQF-COSY{31P} spectrum at 298 K, 4K by 1K complex datapoints, zerofilled to 32K in the f2 dimension. The error in measurement of $\sum^3 J_{1'2'}$ is ± 0.6 – 0.9 Hz

N.A are written for the cases where the coupling constants could not be determined due to overlapping of signals

Table S3. The percentage of North-type sugar population was estimated with the relation %N = 100*(1 -($\Sigma J_{H1'}$ - 9.8) / 5.9), using the $\Sigma^3 J_{H1'}$ from ^{1}H NMR at 298 K for the respective pentofuranose moiety in each nucleotydic units of heptameric ssDNA (**5a** – **8a**). The percentage of North-type sugar population was also estimated with the relation %N = 100*(7.9 - $^3J_{1'2'}$) / 6.9, using the $^3J_{1'2'}$ from ^{1}H NMR at 298 K for the respective pentofuranose moiety in each nucleotydic units of heptameric ssRNA (**5b** – **8b**).

Oligo-ssDNA/ssRNA		%N in the neutral (N) and deprotonated (D) states						
		$C^{5'}$	$A^{5'}$	A^5/C^5	G	A^3/C^3	$A^{3'}$	$C^{3'}$
d($C^{5'}p_1A^{5'}p_2A^5p_3\underline{G}p_4A^3p_5A^{3'}p_6C^{3'}$) (5a)	N	25	5	7	29	12	25	39
	D	25	12	15	29	12	34	49
r($C^{5'}p_1A^{5'}p_2A^5p_3\underline{G}p_4A^3p_5A^{3'}p_6C^{3'}$) (5b)	N	72	68	61	67	78	83	75
	D	70	74	77	28	59	45	77
d($C^{5'}p_1A^{5'}p_2Ap_3\underline{G}p_4C^3p_5A^{3'}p_6C^{3'}$) (6a)	N	25	N.A	N.A	20	24	34	42
	D	25	N.A	N.A	20	20	31	42
r($C^{5'}p_1A^{5'}p_2A^5p_3\underline{G}p_4C^3p_5A^{3'}p_6C^{3'}$) (6b)	N	78	78	74	65	64	78	70
	D	72	91	84	46	64	75	71
d($C^{5'}p_1A^{5'}p_2C^5p_3\underline{G}p_4Ap_5A^{3'}p_6C^{3'}$) (7a)	N	31	25	22	14	10	17*	39*
	D	27	25	22	17	15	24*	36*
r($C^{5'}p_1A^{5'}p_2C^5p_3\underline{G}p_4A^3p_5A^{3'}p_6C^{3'}$) (7b)	N	96	75	54	72	77	81	78
	D	78	77	33	57	71	84	72
d($C^{5'}p_1A^{5'}p_2C^5p_3\underline{G}p_4C^3p_5A^{3'}p_6C^{3'}$) (8a)	N	31	25	24	20	22	32	42
	D	N.A	27	24	29	N.A	32	42
r($C^{5'}p_1A^{5'}p_2C^5p_3\underline{G}p_4C^3p_5A^{3'}p_6C^{3'}$) (8b)	N	91	81	72	52	70	78	74
	D	78	80	72	52	64	77	64

[a] All spectra are taken in 600 MHz spectrometer. The error in measurement of $^3J_{1'2'}$ is ± 0.1 Hz with %N error of 2%

* Coupling constant measured in 1D rows from the DQF-COSY{31P} spectrum at 298 K, 4K by 1K complex datapoints, zerofilled to 32K in the f2 dimension. The error in measurement of $\sum^3 J_{1'2'}$ is ± 0.6 – 0.9 Hz with %S error of 12 - 18%

N.A are written for the cases where the coupling constants could not be determined due to overlapping of signals

Table S4. The ^{31}P chemical shifts ($\delta^{31}P$) in ppm for the monomeric bisphosphate model compounds (**9** – **11**) at neutral pH (~7.25) and alkaline pH (~12.5). At each pH values chemical shift has been determined at three different experiments to estimate the error in digital resolution.

Compounds	$\delta^{31}P$ at pH ~7.25		$\delta^{31}P$ at pH ~12.5	
	p_1 (in ppm)	p_2 (in ppm)	p_1 (in ppm)	p_2 (in ppm)
Etp$_1$**dA**p$_2$Et (**9a**)	-0.921 -0.921 -0.921	-1.546 -1.546 -1.546	-0.907 -0.906 -0.907	-1.535 -1.534 -1.534
Etp$_1$**dC**p$_2$Et (**10a**)	-0.888 -0.888 -0.887	-1.595 -1.594 -1.594	-0.862 -0.861 -0.861	-1.578 -1.578 -1.578
Etp$_1$**dG**p$_2$Et (**11a**)	-0.903 -0.903 -0.902	-1.520 -1.521 -1.520	-0.871 -0.871 -0.871	-1.505 -1.505 -1.506
Etp$_1$**A**p$_2$Et (**9b**)	-0.946 -0.946 -0.946	-1.205 -1.205 -1.205	-0.924 -0.924 -0.924	-1.144 -1.144 -1.144
Etp$_1$**C**p$_2$Et (**10b**)	-0.970 -0.969 -0.969	-1.278 -1.277 -1.277	-0.946 -0.946 -0.946	-1.223 -1.221 -1.221
Etp$_1$**G**p$_2$Et (**11b**)	-0.939 -0.938 -0.938	-1.186 -1.186 -1.185	-0.935 -0.934 -0.934	-1.140 -1.140 -1.140

Table S5. The line broadening values for ^{31}P chemical shifts ($\delta^{31}P$) in ppm for the monomeric bisphosphate model compounds (**9** – **11**) as well as heptameric ssDNAs/ssRNAs (**5** – **8**) at neutral pH (~7.25) and alkaline pH (~12.5) pH.

Compounds		Line Broadening for $\delta^{31}P$ in ppm					
		p_1	p_2	p_3	p_4	p_5	p_6
Etp$_1$**dA**p$_2$Et (**9a**)	**N**	0.01	0.01	–	–	–	
	D	0.01	0.01	–	–	–	–
Etp$_1$**dC**p$_2$Et (**10a**)	**N**	0.01	0.01	–	–	–	
	D	0.01	0.01	–	–	–	–
Etp$_1$**dG**p$_2$Et (**11a**)	**N**			–	–	–	
	D			–	–	–	–
Etp$_1$**A**p$_2$Et (**9b**)	**N**	0.01	0.01	–	–	–	
	D	0.01	0.01	–	–	–	–
Etp$_1$**C**p$_2$Et (**10b**)	**N**	0.01	0.01	–	–	–	
	D	0.02	0.02	–	–	–	–
Etp$_1$**G**p$_2$Et (**11b**)	**N**	0.01	0.01	–	–	–	–
	D	0.01	0.01	–	–	–	–
d(Cp$_1$Ap$_2$Ap$_3$**G**p$_4$Ap$_5$Ap$_6$C) (**5a**)	**N**	0.02	–	–	0.02	–	0.01
	D	0.01	–	–	0.02	–	0.01
d(Cp$_1$Ap$_2$Ap$_3$**G**p$_4$Cp$_5$Ap$_6$C) (**6a**)	**N**	0.02	0.03	0.03	0.02	0.02	0.02
	D	0.02	0.02	–	0.01	–	0.01
d(Cp$_1$Ap$_2$Cp$_3$**G**p$_4$Ap$_5$Ap$_6$C) (**7a**)	**N**	0.03	0.03	–	0.02	0.02	0.02
	D	0.02	0.02	–	0.02	0.02	0.01
d(Cp$_1$Ap$_2$Cp$_3$**G**p$_4$Cp$_5$Ap$_6$C) (**8a**)	**N**	0.03	–	0.04	–	0.03	0.03
	D	0.03	–	0.03	0.03	0.03	–
r(Cp$_1$Ap$_2$Ap$_3$**G**p$_4$Ap$_5$Ap$_6$C) (**5b**)	**N**	0.05	–	–	–	–	–
	D	0.03	0.03	0.03	0.03	0.03	0.02
r(Cp$_1$Ap$_2$Ap$_3$**G**p$_4$Cp$_5$Ap$_6$C) (**6b**)	**N**	0.06	–	–	–	–	–
	D	0.03	0.03	0.03	0.03	–	–
r(Cp$_1$Ap$_2$Cp$_3$**G**p$_4$Ap$_5$Ap$_6$C) (**7b**)	**N**	0.04	0.06	–	–	–	–
	D	0.03	0.04	0.04	–	–	–
r(Cp$_1$Ap$_2$Cp$_3$**G**p$_4$Cp$_5$Ap$_6$C) (**8b**)	**N**	0.05	–	–	–	–	–
	D	0.02	0.02	0.02	0.02	0.02	0.02

Table S6: ^{31}P chemical shifts [δ^{31}_{P}, in ppm] at the neutral (N) and the deprotonated (D) states at 298 K for ssDNA trimers **1a** – **4a**, ssDNA heptamers **5a** – **8a** as well as ssRNA trimers **1b** – **4b**, ssRNA heptamers **5b** – **8b.**

Compounds		^{31}P Chemical Shift [δ^{31}_{P}] at neutral (N) and *deprotonated* (D) states at 298 K					
		p_1	p_2	p_3	p_4	p_5	p_6
d(Ap$_1$Gp$_2$A) (1a)	**N**	-4.2730	-3.8670				
	D	*-4.0460*	*-3.6720*				
r(Ap$_1$Gp$_2$A) (1b)	**N**	-1.9140	-1.8060				-
	D	*-1.7730*	*-1.6510*				-
d(Ap$_1$Gp$_2$C) (2a)	**N**	-2.9640	-2.8350				
	D	*-2.4070*	*-2.1790*				
r(Ap$_1$Gp$_2$C) (2b)	**N**	-2.2650	-2.2990				
	D	*-2.0950*	*-2.0440*				
d(Cp$_1$Gp$_2$A) (3a)	**N**	-2.5270	-2.3970				
	D	*-2.4070*	*-2.1790*				
r(Cp$_1$Gp$_2$A) (3b)	**N**	-2.1630	-2.1820				
	D	*-1.9920*	*-2.1820*				
d(Cp$_1$Gp$_2$C) (4a)	**N**	-2.5010	-2.2900				
	D	*-2.4020*	*-2.1430*				
r(Cp$_1$Gp$_2$C) (4b)	**N**	-1.8920	-2.0630				
	D	*-1.8870*	*-1.8690*				
d(Cp$_1$Ap$_2$Ap$_3$Gp$_4$Ap$_5$Ap$_6$C) (5a)	**N**	-2.3100	-2.5930	-2.5530	-2.2410	-2.5280	-2.0890
	D	*-2.2760*	*-2.4840*	*-2.3220*	*-1.9910*	*-2.5040*	*-2.0510*
r(Cp$_1$Ap$_2$Ap$_3$Gp$_4$Ap$_5$Ap$_6$C) (5b)	**N**	-1.8000	-1.9740	-2.0380	-2.1160	-2.1800	-2.1590
	D	*-1.8250*	*-1.7320*	*-1.8370*	*-1.9540*	*-2.1680*	*-2.0990*
d(Cp$_1$Ap$_2$Ap$_3$Gp$_4$Cp$_5$Ap$_6$C) (6a)	**N**	-2.2960	-2.5610	-2.4770	-2.1320	-2.3620	-2.0800
	D	*-2.2590*	*-2.4530*	*-2.3010*	*-1.9750*	*-2.3250*	*-2.0310*
r(Cp$_1$Ap$_2$Ap$_3$Gp$_4$Cp$_5$Ap$_6$C) (6b)	**N**	-1.7560	-2.0820	-2.1460	-2.2340	-2.0820	-2.1460
	D	*-1.7530*	*-1.8120*	*-1.9470*	*-1.9470*	*-2.0070*	*-2.0500*
d(Cp$_1$Ap$_2$Cp$_3$Gp$_4$Ap$_5$Ap$_6$C) (7a)	**N**	-2.2200	-2.2200	-2.3490	-2.3980	-2.4950	-2.0730
	D	*-2.1430*	*-2.1770*	*-2.1770*	*-2.0570*	*-2.4560*	*-2.0120*
r(Cp$_1$Ap$_2$Cp$_3$Gp$_4$Ap$_5$Ap$_6$C) (7b)	**N**	-1.8980	-1.9860	-2.0840	-2.2030	-2.1440	-2.1440
	D	*-1.8540*	*-1.7760*	*-1.6770*	*-2.0020*	*-2.0290*	*-1.9940*
d(Cp$_1$Ap$_2$Cp$_3$Gp$_4$Cp$_5$Ap$_6$C) (8a)	**N**	-2.2060	-2.1630	-2.3050	-2.1390	-2.3600	-2.0580
	D	*-2.1800*	*-2.0010*	*-2.1800*	*-2.1240*	*-2.3200*	*-2.0180*
r(Cp$_1$Ap$_2$Cp$_3$Gp$_4$Cp$_5$Ap$_6$C) (8b)	**N**	-2.1310	-2.4990	-2.2640	-2.3910	-2.3530	-2.417
	D	*-2.1170*	*-2.3460*	*-2.0800*	*-2.2060*	*-2.2800*	*-2.3460*

Table S7: Oligomerization shift estimated from ^{1}H chemical shift at the neutral (N) state at 298K for aromatic protons ssRNA using appropriate monomeric reference compounds

Sequence	$C^{5'}$	$A^{5'}$	A^{5}/C^{5}	G	A^{3}/C^{3}	$A^{3'}$	$C^{3'}$
r($C^{5'}pA^{5'}pA^{5}p\underline{\mathbf{G}}pA^{3'}pA^{3'}pC$) **5b**	0.430 (H5C) 0.161 (H6C)	0.249 (H8A) 0.360 (H2A)	0.400 (H8A) 0.330 (H2A)	0.366 (H8G)	0.349 (H8A) 0.415 (H2A)	0.432 (H8A) 0.221 (H2A)	0.538 (H5C) 0.351 (H6C)
r($C^{5'}pA^{5'}pA^{5}p\underline{\mathbf{G}}pC^{3'}pA^{3'}pC$) **6b**	0.436 (H5C) 0.131 (H6C)	0.235 (H8A) 0.368 (H2A)	0.415 (H8A) 0.268 (H2A)	0.424 (H8G)	0.613 (H5C) 0.303 (H6C)	0.215 (H8A) 0.206 (H2A)	0.461 (H5C) 0.288 (H6C)
r($C^{5'}pA^{5'}pC^{5}p\underline{\mathbf{G}}pC^{3'}pA^{3'}pC$) **7b**	0.379 (H5C) 0.091 (H6C)	0.178 (H8A) 0.193 (H2A)	0.529 (H5C) 0.355 (H6C)	0.258 (H8G)	0.295 (H8A) 0.369 (H2A)	0.412 (H8A) 0.193 (H2A)	0.526 (H5C) 0.337 (H6C)
r($C^{5'}pA^{5'}pC^{5}p\underline{\mathbf{G}}pC^{3'}pA^{3'}pC$) **8b**	0.380 (H5C) 0.071 (H6C)	0.172 (H8A) 0.185 (H2A)	0.547 (H5C) 0.345 (H6C)	0.245 (H8G)	0.525 (H5C) 0.237 (H6C)	0.190 (H8A) 0.191 (H2A)	0.453 (H5C) 0.279 (H6C)

Table S8: Oligomerization shift estimated from ^{1}H chemical shift at the deprotonated (D) state at 298K for aromatic protons ssRNA using appropriate monomeric reference compounds

Sequence	$C^{5'}$	$A^{5'}$	A^{5}/C^{5}	G	A^{3}/C^{3}	$A^{3'}$	$C^{3'}$
r($C^{5'}$p$A^{5'}$pA^{5}p**G**p$A^{3'}$p$A^{3'}$pC) **5b**	0.431 (H5C) 0.204 (H6C)	0.248 (H8A) 0.316 (H2A)	0.351 (H8A) 0.348 (H2A)	0.195 (H8G)	0.336 (H8A) 0.497 (H2A)	0.454 (H8A) 0.226 (H2A)	0.543 (H5C) 0.358 (H6C)
r($C^{5'}$p$A^{5'}$pA^{5}p**G**p$C^{3'}$p$A^{3'}$pC) **6b**	0.431 (H5C) 0.161 (H6C)	0.237 (H8A) 0.343 (H2A)	0.376 (H8A) 0.259 (H2A)	0.330 (H8G)	0.622 (H5C) 0.337 (H6C)	0.197 (H8A) 0.206 (H2A)	0.443 (H5C) 0.276 (H6C)
r($C^{5'}$p$A^{5'}$pC^{5}p**G**p$C^{3'}$p$A^{3'}$pC) **7b**	0.388 (H5C) 0.114 (H6C)	0.163 (H8A) 0.169 (H2A)	0.552 (H5C) 0.371 (H6C)	0.171 (H8G)	0.267 (H8A) 0.407 (H2A)	0.425 (H8A) 0.178 (H2A)	0.530 (H5C) 0.336 (H6C)
r($C^{5'}$p$A^{5'}$pC^{5}p**G**p$C^{3'}$p$A^{3'}$pC) **8b**	0.384 (H5C) 0.085 (H6C)	0.158 (H8A) 0.165 (H2A)	0.567 (H5C) 0.360 (H6C)	0.177 (H8G)	0.543 (H5C) 0.255 (H6C)	0.170 (H8A) 0.188 (H2A)	0.437 (H5C) 0.270 (H6C)