Nucleic Acids 2, 237-318.
Robertus, J. D., Ladner, J. E., Finch, J. J., Rhodes, D., Brown, R. S., Clark, B. F., and Klug, A. (1974), *Nucleic Acid Res. 1,* 927-932.
Scruggs, R. L., and Ross, P. D. (1970), *J. Mol. Biol. 47,* 29-40.
Shulman, R. G., Hilbers, C. W., Kearns, D. R., Reid, B. R., and Wong, Y. P. (1973), *J. Mol. Biol. 78,* 57-69.
Simuth, J., Scheit, K. H., and Gottschalk (1970), *Biochem. Biophys. Acta 204,* 317-380.
Stevens, C. L., and Felsenfeld, G. (1964), *Biopolymers 2,* 293-314.
Teitelbaum, H., and Englander, S. W. (1975a), *J. Mol. Biol. 92,* 55-78.
Teitelbaum, H., and Englander, S. W. (1975b), *J. Mol. Biol. 92,* 79-92.
Ueda, T., Imazawa, M., Mivra, K., Iwata, R., and Odajima, K. (1971), *Tetrahedron Lett. 27,* 2507-2510.
Wong, K. L., Bolton, P. H., and Kearns, D. R. (1975), *Biochim. Biophys. Acta 383,* 446-451.
Zimmerman, S. B., Cohen, G. H., and Davies, D. R. (1975), *J. Mol. Biol. 92,* 181-192.

Hydrogen-Bonded Complexes of the Ribodinucleoside Monophosphates in Aqueous Solution. Proton Magnetic Resonance Studies†

Thomas R. Krugh,* Jean Wilcox Laing,‡ and Michael A. Young

ABSTRACT: A proton magnetic resonance study of the chemical shifts of a series of ribodinucleoside monophosphates in neutral H_2O solution has been recorded in the 1-100 mM concentration range. The self-complementary dinucleoside monophosphates CpG and GpC and the complementary mixture GpU + ApC form intermolecular hydrogen-bonded complexes at low temperatures. The amino proton chemical shifts in the CpG and GpC spectra are consistent with the formation of a miniature double helical dimer in neutral aqueous solution at low temperatures (~2 °C). The complementary mixture of dinucleosides GpU + ApC formed much less stable complexes than either GpC or CpG, while UpA did not show any indication of the formation of intermolecular hydrogen-bonded complexes. This result is consistent with the well-known observation that the stability of a double helix is proportional to the percent of G·C base pairs present.

The use of oligonucleotides of defined length and sequence as models for the corresponding regions in nucleic acids has provided valuable information on the properties of the nucleic acids (e.g., Jaskunas et al., 1968; Gennis and Cantor, 1970; Martin et al., 1971; Day et al., 1973; Uhlenbeck et al., 1973; Gralla and Crothers, 1973a,b; Borer et al., 1974, 1975; Ts'o, 1974a,b; and references therein; D'Albis et al., 1975; Hingerty et al., 1975). Optical spectroscopies have generally been used to investigate the properties of oligonucleotides. However, in the past few years nuclear magnetic resonance spectra of nucleic acids and proteins dissolved in an H_2O solvent have provided useful geometrical, kinetic, and thermodynamic information (e.g., Glickson et al., 1969; Kearns et al., 1971; Crothers et al., 1973; Patel and Tonelli, 1974; Arter et al., 1974; Krugh and Young, 1975; Young and Krugh, 1975; Borer et al., 1975). The NH resonance from G·C and A·T base pairs is a convenient probe for studying small double helical nucleic acids because this resonance is 5-9 ppm downfield from the large solvent resonance. However, the NH resonances broaden into the baseline at a temperature far below the T_m of the double helix and, thus, the extraction of thermodynamic parameters is not straightforward (e.g., Crothers et al., 1973; Borer et al., 1975). Mooberry and Krugh (1975) have shown that it is relatively easy to obtain pulsed Fourier transform proton magnetic resonance spectra in H_2O solutions. We have used these techniques to monitor the amino resonances of several deoxydinucleotides and have shown that complementary deoxydinucleotides interact by the formation of hydrogen-bonded dimers (Krugh and Young, 1975; Young and Krugh, 1975; see these papers for a more complete introduction to this area). An analysis of the concentration dependence of the amino resonances illustrated the effect of nucleotide sequence on the stability of the miniature double helices formed by two complementary deoxydinucleotides. In this paper we will investigate the interaction of several of the *ribo*dinucleoside monophosphates in neutral aqueous solution and show that, at temperatures near 0 °C, both CpG and GpC self-associate by the formation of intermolecular hydrogen-bonded complexes.

Experimental Methods

The dinucleoside monophosphates used in this study were purchased from either Sigma Chemical Co., or P-L. Biochemicals, Inc., and were used without further purification. Aqueous solutions of the ribodinucleoside monophosphates were initially made as concentrated as possible within the limits of solubility (or availability) of the compounds. Approximately 5 μl of 10^{-2} M disodium ethylenediaminete-

† From the Department of Chemistry, University of Rochester, Rochester, New York, 14627. *Received July 22, 1975.* This investigation was supported by Public Health Research Grant No. CA-14103 from the National Cancer Institute.
‡ Present address: Xerox Corporation, Webster, New York.

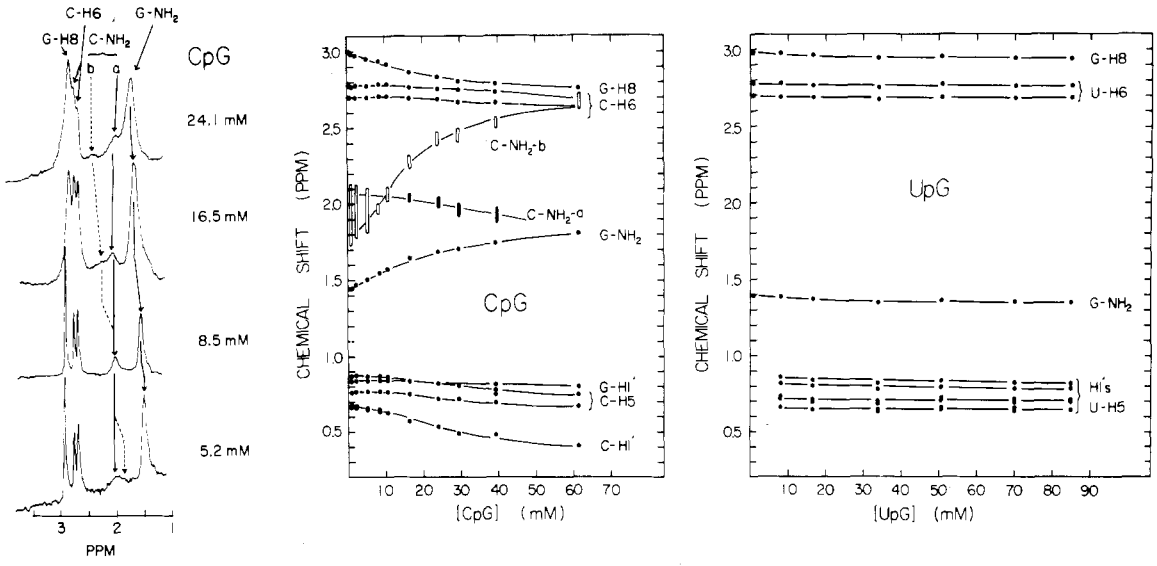


FIGURE 1: (Left) The 100-MHz spectra of CpG of varying concentration in H_2O illustrating the difficulty of monitoring the cytosine amino protons in the concentration range in which the resonances overlap. (Center) The concentration dependence of the chemical shifts of CpG in H_2O solution (4 °C). (Right) The concentration dependence of the chemical shifts of UpG in H_2O solution (1 °C).

traacetic acid was added to solutions of guanine-containing dinucleosides to eliminate metal ion broadening of the guanine resonances. A measured volume of solution was placed in a 5-mm NMR tube and the pH was adjusted to 7.05 ± 0.10 (22 °C). The spectrum was recorded and the solution was diluted by adding a calculated amount of filtered redistilled water. After each dilution the pH was adjusted, and the concentration of the dinucleoside monophosphate was determined by measuring the absorbance of a small amount of the sample solution diluted in 0.1 N HCl and using the extinction coefficients listed in No. 103 of P-L. Biochemicals, Inc.

The 100-MHz Fourier transform ^1H NMR spectra were recorded with a JEOL PFT-100 spectrometer interfaced with a JEOL EC-100 computer. The water peak was minimized with the 180°–τ(HSP[1])–90° pulse sequence (Mooberry and Krugh, 1975; Krugh and Schaefer, 1975). The sample temperature was maintained within ±0.5 °C during the spectral accumulation with a JEOL VT-3 temperature controller. The temperature was determined by measuring the separation of the methanol resonances (Van Geet, 1970). Chemical shifts were measured relative to the residual water resonance (except where noted otherwise) with an accuracy of ±0.005 ppm on narrow resonances. The chemical shifts of the water resonances relative to TSP[1] were determined at several temperatures in a separate experiment and found to be: 0 °C, 5.069; 1 °C, 5.058; 2 °C, 5.047; and 4 °C, 5.025.

Results

The proton chemical shifts of CpG in H_2O solution at 4 °C are shown as a function of concentration in Figure 1. The guanine amino resonance moves *downfield* as a function of increasing concentration of the self-complementary dinucleoside monophosphate. This large downfield shift is a result of intermolecular hydrogen-bond formation (e.g., Krugh and Young, 1975; Young and Krugh, 1975). The partial double-bond character of the cytidine C4–N bond restricts the rotation of the C-NH_2 group, which at low temperatures results in the observation of a separate broad resonance for each of the C-NH_2 protons in Figure 1 (e.g., Shoup et al., 1971; McConnell and Seawell, 1973; Raszka and Kaplan, 1972; Krugh and Young, 1975). The C-NH_2-b resonance moves downfield, while the C-NH_2-a resonance moves upfield as a function of increasing concentration of CpG (Figure 1), which is consistent with the formation of Watson–Crick base pairs. The magnitude of the downfield shift of the C-NH_2-b resonance is much larger than the downfield shift of the G-NH_2 resonance partly because the guanine amino group undergoes fast rotation on the NMR time scale. The observed G-NH_2 resonance position is an average of the two resonances and, thus, the observed deshielding due to hydrogen-bond formation is only one-half as large as the actual deshielding. The total intensity of the cytidine amino resonances appears to be less than one-half of the guanine resonance. This most likely results from the use of the 180°–τ–90° pulse sequence to eliminate the solvent water resonance since this sequence may lead to a loss of intensity of the resonance of protons that exchange with the solvent (Krugh and Schaefer, 1975; Pitner et al., 1974; and references therein).

The downfield region of the CpG spectra in the concentration range in which the cytidine amino resonances cross over is shown in Figure 1. The position of the cytidine amino resonances is much more difficult to determine than that of the guanine amino resonance, as characterized by the larger uncertainty bars in Figures 1 and 3.

The concentration dependence of the guanine amino resonance may be used to estimate the magnitude of the equilibrium constant for complex formation. Assuming that the self-complementary CpG dimerizes in aqueous solution to form a miniature double helix

$$2\,\overrightarrow{\text{CpG}} \underset{}{\overset{K}{\rightleftharpoons}} \begin{matrix}\overrightarrow{\text{CpG}} \\ \overleftarrow{\text{GpC}}\end{matrix} \qquad (1)$$

the appropriate expression for the calculation of the dimerization equilibrium constant is

[1] Abbreviations used: TSP, sodium 3-trimethylsilylpropionate-*2,2,3,3-d*$_4$; HSP, homogeneity spoiling pulse.

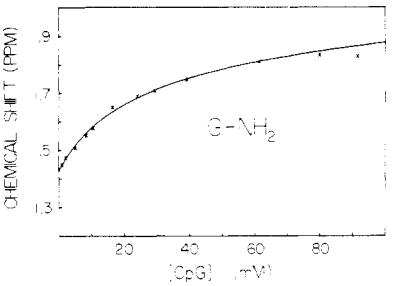


FIGURE 2: A comparison between the experimental G-NH_2 chemical shifts and the calculated dimerization curve using the parameters listed in the text. The chemical shift data of the two most concentrated solutions were not used in the least-squares analysis of the data.

$$\delta_{obsd} = \delta_M + (\delta_D - \delta_M)\left(\frac{4KN_0 + 1 - \sqrt{(8KN_0 + 1)}}{4KN_0}\right) \quad (2)$$

where δ_{obsd} is the observed chemical shift; δ_M is the chemical shift of the protons in the monomer state; δ_D is the chemical shift of the protons in the dimer state; and N_0 is the total concentration of CpG. A nonlinear least-squares regression analysis program was used to calculate the values of K, δ_M, δ_D, and associated standard deviations. The good agreement between the experimental and calculated data for the parameters (K = 14.0 ± 2.5 M^{-1}, δ_M = 1.42 ± 0.01 ppm, δ_D = 2.24 ± 0.06 ppm) is illustrated in Figure 2. We note, however, that the most concentrated solution shows significant deviation from the calculated curve (Figure 2). The upfield deviation is what we would expect to observe if the CpG monomers or dimers also formed nonspecific base stacked complexes. This point will be discussed in more detail in the following section. We did not include the two most concentrated data points in the calculations in order to minimize the influence of the nonhydrogen-bonded base stacking on the calculated parameters, and we thus caution that the calculated standard deviation listed above is not a true representation of the uncertainty in the chemical shift of the amino protons in the CpG dimer.

The dramatic difference in the concentration dependence of the guanine amino resonance of UpG when compared with CpG (Figure 1) clearly demonstrates the influence of complementarity (in the Watson-Crick sense) on the stability of hydrogen-bonded complexes. The small upfield shift of the guanine amino resonance of UpG is indicative of nonspecific base stacking. The concentration dependences of the chemical shifts of the other base protons of UpG are also quite different from those observed for CpG and indicate that UpG self-associates only by nonspecific base stacking.

The self-complementary dinucleoside monophosphate GpC does self-associate by the formation of hydrogen-bonded complexes as shown by the data in Figure 3. The guanine amino resonance and the C-NH_2-b resonance move downfield as a function of increasing GpC concentration, while the other base protons move upfield over the same concentration range. Unfortunately, the G-NH_2 and the C-NH_2-a resonances overlap at concentrations above 20 mM and we do not feel that the data would provide reliable estimates for a dimerization constant. In addition, the G-NH_2 chemical shift curve appears to be S shaped at low concentrations which indicates that species other than simple hydrogen-bonded dimers may be present in this case. The temperature dependence of the chemical shifts of the base protons of GpC is illustrated for a 37 mM solution in Figure 4. The fraction of GpC molecules existing as hydrogen-bonded complexes, and the equilibrium constant(s), are thus seen to be temperature sensitive.

A comparison of the concentration dependence of the chemical shifts of GpU (Figure 3) with the data for the self-complementary GpC once again shows the marked effect of complementarity on the formation of hydrogen-bonded complexes. That the upfield shifts of all the proton resonances of GpU (Figure 3) are small is consistent with the formation of nonspecific base stacked complexes.

The concentration dependence of the chemical shifts of an equimolar mixture of the complementary dinucleosides GpU + ApC is shown in Figure 3. The G-NH_2 resonance moves downfield as the concentration is increased, which is indicative of hydrogen-bond formation. On the other hand, the A-NH_2 resonance does not appear to move downfield. The chemical shifts of the ApC protons over this same concentration range (Figure 3) show clear evidence of intermolecular base stacking (Bangerter and Chan, 1969). The observed shifts of the A-NH_2 resonance in the GpU + ApC mixture will be a sum of both the shielding effects due to nonspecific intermolecular base stacking, as well as the deshielding due to base pair formation. The data for the mixture of GpU + ApC is thus not conclusive with respect to the formation of a miniature double helical complex, but the data (Figure 3) lead to the definite conclusion that the

$$\overrightarrow{\text{GpU}}$$
$$\underleftarrow{\text{CpA}}$$

miniature double helix is much less stable than either the

$$\overrightarrow{\text{GpC}} \quad \overrightarrow{\text{CpG}}$$
$$\text{or}$$
$$\underleftarrow{\text{CpG}} \quad \underleftarrow{\text{GpC}}$$

miniature double helices.

The concentration dependence of the chemical shifts of UpA are shown in Figure 5. The upfield shifts of all of the base protons are consistent with intermolecular base stacking, and no evidence is obtained for the formation of hydrogen-bonded complexes. We have also recorded the concentration dependence of the chemical shifts of CpA (not shown) and, as expected, we found only small upfield shifts of the nucleotide base resonances as the concentration was increased.

Discussion

In the GpC and CpG spectra, the large downfield shifts of the guanine amino resonance and the C-NH_2-b resonance as a function of increasing dinucleoside monophosphate concentration clearly show that the amino protons are involved in hydrogen-bond formation, presumably by the formation of a miniature double helix. GpC has previously been found to form double helical hydrogen bonded dimers in the solid state (Day et al., 1973; Hingerty et al., 1975) in agreement with the present results.

In principle the concentration dependence of the chemical shifts of the nonexchangeable base protons can also be analyzed using eq 2. For these protons, δ_M and δ_D are defined as the limiting shifts of each proton in the monomer (free dinucleoside) and dimer (two hydrogen-bonded dinucleosides) states, respectively. The term δ_M is determined by

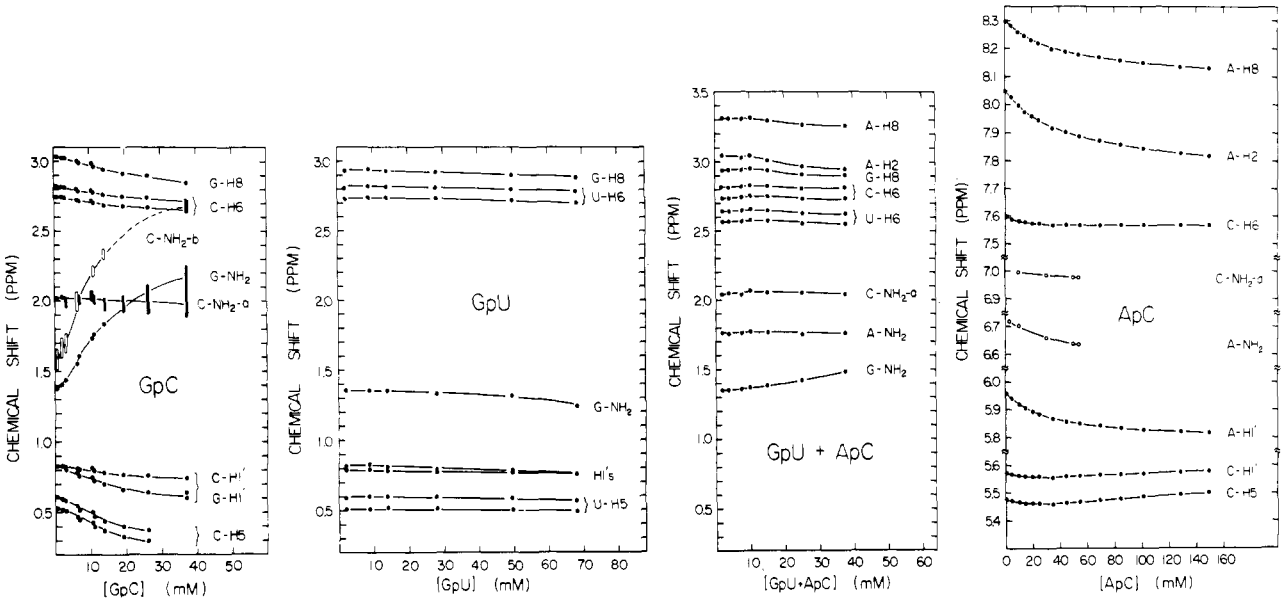


FIGURE 3: The concentration dependence of the chemical shifts of: GpC in H_2O solution (4 °C); GpU in H_2O (0 °C); GpU + ApC (1:1) in H_2O (1 °C); and ApC in D_2O (4 °C, solid circles). The ApC in D_2O chemical shifts are given relative to an *external* hexamethyldisiloxane capillary reference. Several ApC spectra were also recorded in H_2O solution (4 °C, open circles) in order to measure the concentration dependence of the amino resonances.

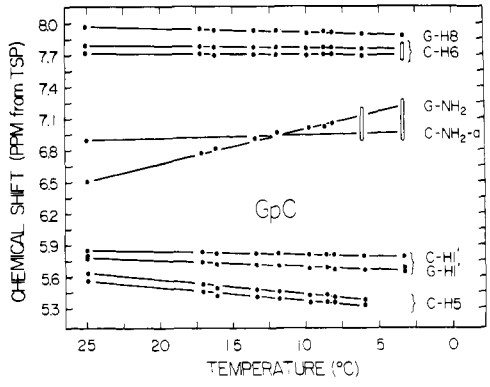


FIGURE 4: The temperature dependence of the chemical shifts of a 37 mM GpC solution in H_2O. These chemical shifts are given with respect to TSP (sodium 3-trimethylsilylpropionate-*2,2,3,3-d*$_4$).

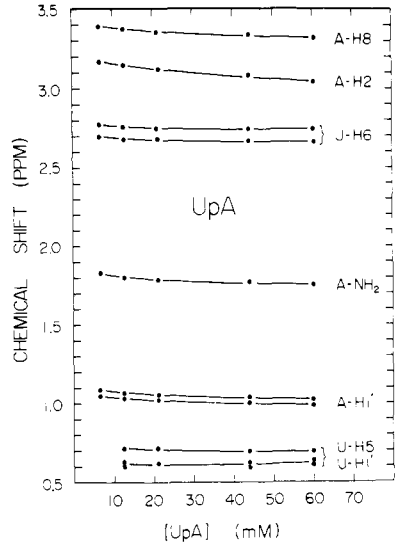
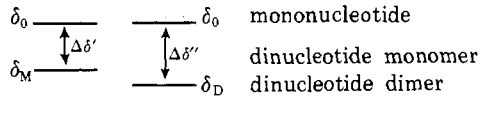


FIGURE 5: The concentration dependence of the chemical shifts of UpA in H_2O solution (2 °C).

extrapolating the chemical shift of the dinucleoside base resonance to zero concentration, while δ_D may be calculated (or estimated) by extrapolation of the chemical shifts to very high concentrations. The value of δ_D may also be estimated from the chemical shift of the nucleotide proton in an appropriate mononucleotide, δ_0, by adding the ring current shift (designated here as $\Delta\delta''$) predicted for that proton in possible double helical environments ($\delta_D = \delta_0 + \Delta\delta''$). The δ_0 and δ_M terms are not equivalent due to the ring current shift observed when two nucleosides are joined by a phosphodiester linkage (designated here as $\Delta\delta'$ and therefore $\delta_M = \delta_0 + \Delta\delta'$). The interrelationship of all of these terms is summarized in the diagram below.

downfield



upfield

The values of both $\Delta\delta'$ and $\Delta\delta''$ generally result in the resonances moving upfield with respect to the chemical shifts of the protons in the appropriate mononucleotide. In both the deoxy- and ribodinucleoside monophosphates and dinucleotides, we experimentally observe that all of the nonhydrogen-bonded proton resonances move upfield as the concentration is increased (Young and Krugh, 1975; this manuscript). The magnitudes of the changes in the chemical shifts of the C-H5 protons as a function of increasing concentration are significantly larger than predicted from the ring current shifts calculated from various double helical conformations of a miniature double helix composed of two dinucleoside monophosphate molecules. Thus there is no doubt that the ribodinucleoside monophosphates also self-associate by the formation of intermolecular base stacked complexes (e.g., see Figures 1c, 3b, 3d, and 5) as has been well documented by the work of Chan, Ts'o, and others (e.g., see the review by Ts'o, 1974a,b). The indication is, therefore, that the hydrogen-bonded complexes probably also associate to some degree by forming base

stacked complexes. Unfortunately, the interference of this intermolecular stacking on the chemical shifts of the base protons prevents the extraction of detailed geometric information concerning the conformation of these hydrogen-bonded complexes. It is interesting to note that considerably more stacking occurs with the hydrogen-bonded dimers than with the non-self-complementary ribo- and deoxyribodinucleosides, which suggests a cooperative interaction involving hydrogen bonding (to form dimers) and end stacking (of dimers). The *deshielding* of the amino resonances due to hydrogen-bond formation is much larger (~1.5 ppm) than the shielding effects due to nonspecific base stacking. The satisfactory fit of the CpG data by a dimerization curve (Figure 2) and the role of complementarity support the interpretation of the data in terms of the formation of a miniature double helix. The value of the equilibrium constant and the chemical shift of the amino protons in the dimer, δ_D, are definitely influenced by the form of the data analysis and the interference of intermolecular stacking and should be considered as only approximate.

The present data illustrate the well-known phenomenon that the stability of a double helix is proportional to the G + C percent composition since both CpG and GpC form much more stable complexes than the mixture GpU + ApC. In addition, we did not find any evidence for the formation of intermolecular hydrogen-bonded complexes in the UpA spectra.

Prescott et al. (1974) have also recently obtained evidence from Raman spectral data which indicate that GpC forms hydrogen-bonded base pairs in neutral aqueous solutions, but they found no evidence for base pairing in a CpG solution. On the other hand, the present proton magnetic resonance data are definitive in showing that *both* GpC and CpG form intermolecular hydrogen-bonded complexes.

The ^{1}H NMR data for the ribodinucleoside monophosphates are consistent with our results for the deoxydinucleotides (Young and Krugh, 1975) and continue to demonstrate the valuable information that may be obtained from monitoring the chemical shifts of the amino proton resonances in H_2O solutions. The dinucleoside monophosphates are the smallest structural units containing the 3′–5′-phosphodiester linkage and are thus the shortest oligonucleotides that can be used as models for RNA. The importance of chain length on the stability of the hydrogen-bonded complexes is clearly evident in a comparison of the present data with the results of Raszka and Kaplan (1972) for the mixtures of the mononucleotides in which small downfield shifts (<0.15 ppm) of the amino protons demonstrated the formation of hydrogen-bonded complexes. This comparison will also show the enhanced specificity obtained in the use of the dinucleoside monophosphates since we only find evidence for the formation of Watson–Crick base pairs while Raszka and Kaplan (1972) found evidence for several types of base pairs in the concentrated solutions required for their experiments. Increasing the strand length from a dinucleoside to the tetranucleotide CpCpGpG (Arter et al., 1974) or a hexanucleotide (Borer et al., 1975) substantially increases the stability of the helical complexes, as expected.

References

Arter, D. B., Walker, G. C., Uhlenbeck, O. C., and Schmidt, P. G. (1974), *Biochem. Biophys. Res. Commun. 61,* 1089.

Bangerter, B. W., and Chan, S. I. (1969), *J. Am. Chem. Soc. 91,* 3910.

Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974), *J. Mol. Biol. 86,* 843.

Borer, P. N., Kan, L. S., and Ts'o, P. O. P. (1975), *Biochemistry 14,* 4847.

Crothers, O. M., Hilbers, C. W., and Shulman, R. G. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 2899.

D'Albis, A., Wickens, M. P., and Gratzer, W. B. (1975), *Biopolymers 14,* 1423.

Danyluk, S. (1975), in Basic Principles in Nucleic Acid Chemistry III, Ts'o, P. O. P., Ed., New York, N.Y., Academic Press (in press).

Day, R. O., Seeman, N. C., Rosenberg, J. M., and Rich, A. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 849.

Gennis, R. B., and Cantor, C. R. (1970), *Biochemistry 9,* 4714.

Glickson, J. D., McDonald, C. C., and Phillips, W. D. (1969), *Biochem. Biophys. Res. Commun. 35,* 492.

Gralla, J., and Crothers, D. M. (1973a), *J. Mol. Biol. 73,* 497.

Gralla, J., and Crothers, D. M. (1973b), *J. Mol. Biol. 78,* 301.

Hingerty, B., Subramanian, E., Stellman, S. D., Broyde, S. B., Sato, T., and Langridge, R. (1975), *Biopolymers 14,* 227.

Jaskunas, S. R., Cantor, C. R., and Tinoco, I., Jr. (1968), *Biochemistry 7,* 3164.

Kearns, D. R., Patel, D. J., and Shulman, R. G. (1971), *Nature (London) 229,* 338.

Krugh, T. R., and Young, M. A. (1975), *Biochem. Biophys. Res. Commun. 62,* 1025.

Krugh, T. R., and Schaefer, W. (1975), *J. Magn. Reson. 19,* 99.

Martin, F. H., Uhlenbeck, O. C., and Doty, P. (1971), *J. Mol. Biol. 57,* 201.

McConnell, B., and Seawell, P. C. (1973), *Biochemistry 12,* 4426.

Mooberry, E. S., and Krugh, T. R. (1975), *J. Magn. Reson. 17,* 128.

Patel, D. J., and Tonelli, A. E. (1974), *Biopolymers 13,* 1943.

Pitner, T. P., Glickson, J. D., Dadok, J., and Marshall, G. R. (1974), *Nature (London) 250,* 582.

Prescott, B., Gamache, R., Livramento, J., and Thomas, G. J., Jr. (1974), *Biopolymers 13,* 1821.

Raszka, M., and Kaplan, N. O. (1972), *Proc. Natl. Acad. Sci. U.S.A. 69,* 2025.

Shoup, R. R., Becker, E. D., and Miles, H. T. (1971), *Biochem. Biophys. Res. Commun. 43,* 1350.

Ts'o, P. O. P. (1974a), in Basic Principles in Nucleic Acid Chemistry I, Ts'o, P. O. P., Ed., New York, N.Y., Academic Press, pp 457–585.

Ts'o, P. O. P. (1974b), in Basic Principles in Nucleic Acid Chemistry II, Ts'o, P. O. P., Ed., New York, N.Y., Academic Press, pp 306–471.

Uhlenbeck, O. C., Borer, P. N., Dengler, B., and Tinoco, I., Jr. (1973), *J. Mol. Biol. 73,* 483.

Van Geet, A. L. (1970), *Anal. Chem. 42,* 679.

Young, M. A., and Krugh, T. R. (1975), *Biochemistry 14,* 4841.