# $^{31}$P-NMR analysis of the B to Z transition in double-stranded $(dC\text{-}dG)_3$ and $(dC\text{-}dG)_4$ in high salt solution

Tadeusz A.Holak, Philip N.Borer, George C.Levy*, Jacques H.van Boom[1] and Andrew H.-J.Wang[2]

NIH Research Resource for Multi-Nuclei NMR & Data Processing, Dept. Chemistry, Syracuse University, Syracuse NY 13210, USA, [1]Dept. Organic Chemistry, University of Leiden, The Netherlands & [2]Dept. Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 26 March 1984; Accepted 9 May 1984

ABSTRACT

In 4M NaCl solutions $(dC\text{-}dG)_n$ (n=3,4; ~9mM) exist as a mixture of B and Z forms. The low and high field components of two $^{31}$P NMR resonances originating from internal phosphodiester groups are assigned to the GpC and CpG linkages, respectively. Low temperatures stabilize the Z-forms, which completely disappear above 50°C (n=3) and 65°C (n=4). $\Delta H$=-44 and -17 kJ/mol for B to Z transition in the hexamer and octamer duplexes, respectively. Temperature dependent changes (0-50°C range) in the spin-lattice relaxation times at 145.7 MHz are distinctly different for the $^{31}$P nuclei of GpC and CpG groups. The relaxation data can be explained by assuming that the GpC phosphodiester groups undergo more local internal motion than do the CpG groups.

## INTRODUCTION

$^{31}$P NMR spectroscopy is a valuable tool for monitoring conformational states of poly- and oligonucleotides in cases where separate resonances can be observed in the spectra. $^{31}$P NMR spectroscopy gives evidence for sequence-dependent conformations of right handed deoxynucleotides (1-9). Assignments have recently been made for phosphodiester groups in the sugar-phosphate backbones of short oligonucleotides using $^{31}$P-$^{1}$H decoupling experiments (10-12), two dimensional NMR $^{31}$P-$^{1}$H correlation spectroscopy (13), and from $^{17}$O labeled experiments (14). Also, $^{31}$P NMR spectra of the left handed Z DNA of polynucleotides show two characteristic signals with separation of ca. 0.5-1.5 ppm, which can be used to detect the presence of Z DNA in solution (3,6,8,9,15-23). It is usual to assume that the downfield and upfield components correspond to predominantly gt and gg conformations, respectively, of the $\xi$ and $\alpha$ torsion angles (24) in the O3'-P-O5' linkages (6,15-23). Such $^{31}$P studies would be even more valuable if one could unambigously assign the resonances of the Z form to the

particular type of phosphate group.

Recently the $^{31}P$ resonances of poly(dG-dC) (19) and poly-(dA-dT) (7) in high salt solutions were assigned by selective substitution of the phosphate group by phosphorothioate. The assignments we derive independently for $(dC-dG)_3$ and $(dC-dG)_4$ duplexes in high salt agree with those poly(dG-dC) assignments (19), although our spectra are much better resolved as a result of the lower molecular weight of the oligonucleotide. This correspondence occurs despite problems in the former study due to: (i) assumptions about the similarity in response of the thio- and oxo-phosphates to conformational change and (ii) problems associated with bulk magnetic susceptibility corrections necessary to relate chemical shifts measured in low and high salt with respect to an external standard.

The flexibility and molecular dynamics of deoxynucleic acids in solution have been subject to numerous investigations by NMR relaxation (for reviews see ref. 25 and 26 and references cited therein). So far, no relaxation studies of internal and overall motions in Z DNA have been reported. Although $^{31}P$ relaxation provides at best an incomplete picture of solution DNA dynamics (see below) the $^{31}P$ spin-lattice relaxation times for the Z-form of $(dC-dG)_3$ and $(dC-dG)_4$ do indicate different motional behavior for the CpG and GpC linkages in these oligonucleotides.

## RESULTS AND DISCUSSION

The $^{31}P$ spectra of $(dC-dG)_3$ and $(dC-dG)_4$ in high salt (4M NaCl) solutions at 30°C are shown in Figures 1 and 2. Both B and Z DNA signals which are present in the spectra move downfield ($\Delta\delta_{max}$=0.4ppm) on increasing temperature from ca. 0°C to ca. 50°C. These temperature dependent spectra together with the low salt spectra (*vide infra*) serve to distinguish between resonance signals from Z and B DNA (or the B-like conformation in high salt and at high temperatures). The dependence of the $^{31}P$ spectra of $(dC-dG)_4$ on temperature is shown in Fig. 3. The observed spectral changes are reversible with temperature. Similar spectra were obtained for $(dC-dG)_3$.

### Assignments

The curve fitting module (27) of our NMR data reduction

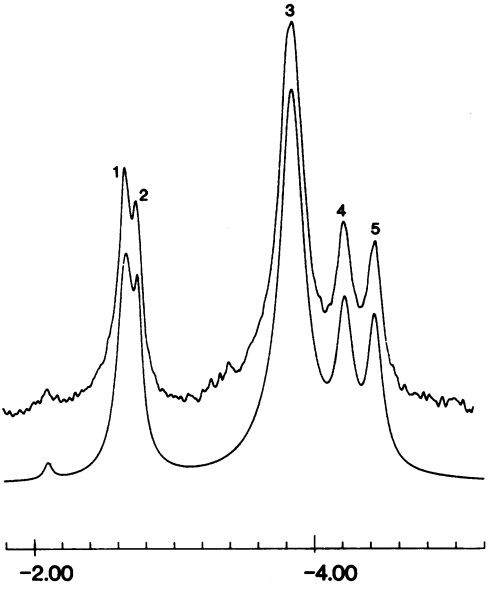


Figure 1. Experimental (upper) and simulated (lower) proton noise-decoupled 145.7 MHz $^{31}P$ NMR spectrum of $(dC\text{-}dG)_3$ in 4M NaCl at 31°C.

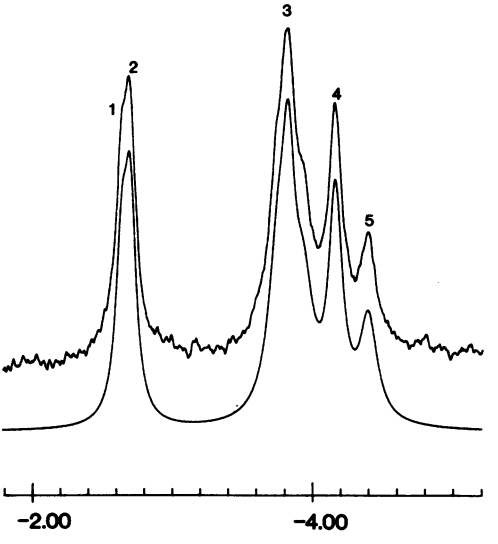


Figure 2. Experimental (upper) and simulated (lower) proton noise decoupled 145.7 MHz $^{31}P$ NMR spectrum of $(dC\text{-}dG)_4$ in 4M NaCl solution at 31°C.

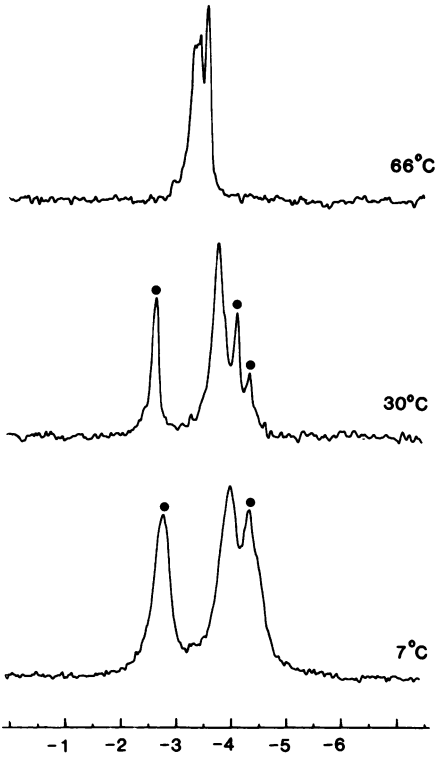


Figure 3. Proton noise-decoupled 145.7 MHz $^{31}P$ NMR spectra of $(dC-dG)_4$ as a function of temperature. The dots represent $^{31}P$ resonances characteristic of left-handed Z-DNA.

software system, NMR1, was used to obtain quantitative measures of the parameters associated with overlapping peaks. Along with experimental spectra, Figures 1 and 2 show simulated spectra obtained from the fitted peaks for $(dC-dG)_3$ and $(dC-dG)_4$. Spin--lattice relaxation times, $T_1$, and nuclear Overhauser enhancement factors (NOEF=NOE-1) measured for each temperature are reported in Table 1. The degree of fit of the simulated to the experimental spectra was evaluated with the DIS discrepancy parameter (27), which was in the range of 0.3-0.5% for these simulations. The analyses were performed on the fully relaxed spectra. Spectra with and without NOE gave nearly identical results since the NOE factors are small at the magnetic fields used, Table 1.

Table 1. $^{31}P$ $T_1$ and NOEF (in parentheses) for $(dC\text{-}dG)_3$ and $(dC\text{-}dG)_4$ at 145.7 MHz.

| Temp.(°C) | $T_1$(sec) of peak[a] | | | | %Z |
|---|---|---|---|---|---|
| | 1+2 GpC | 3 CpG+B | 4 CpG | 5 CpG | |
| $(dC\text{-}dG)_3$ | | | | | |
| 4.5 | 1.37(0.20) | 1.60(0.20) | 1.95(0.40) | 1.89(0.40) | 79 |
| 18 | 1.46(0.10) | 1.34(0.10) | 1.61(0.10) | 1.66(0.20) | 64 |
| 30 | 1.49(0.09) | 1.20(0.08) | 1.32(0.11) | 1.41(0.10) | 52 |
| 40 | 1.40(0.00) | 1.08(0.05) | 1.15(0.05) | 1.55(0.05) | 30 |
| 51 | - | 1.10(0.05) | - | - | 0 |
| $(dC\text{-}dG)_4$ | | | | | |
| -5 | 1.62(0.18) | 2.11(0.25) | 2.42(0.25) | 2.44(0.25) | 73 |
| 7 | 1.46(0.20) | 1.69(0.14) | 2.07(0.25) | 2.07(0.25) | 70 |
| 17.5 | 1.39(0.08) | 1.61(0.10) | 1.91(0.20) | 2.08(0.14) | 66 |
| 30 | 1.43(0.01) | 1.22(0.10) | 1.40(0.00) | 1.32(0.12) | 65 |
| 45 | 1.30(0.03) | 1.10(0.08) | 1.12(0.10) | 1.05(0.10) | 48 |
| 66 | - | 0.90(0.08) | - | - | 0 |
| 31[b] | 1.75(0.17) | 1.77(0.08) | 1.96(0.20) | 1.91(0.15) | 61 |
| 45[b] | 1.69(0.08) | 1.46(0.08) | 1.52(0.10) | 1.56(0.10) | 49 |

[a]Numbering as in Fig. 1 and 2. For $T_1$'s and NOEF's two figures shown although second decimal place not always significant. [b]101.7MHz.

There are 7 distinguishable phosphorous atoms in the octamer duplex:

```
          a b c b''c'b'a'
(5')CpGpCpGpCpGpCpG(3')
     • • • • • • • •
(3')GpCpGpCpGpCpGpC(5')
          a'b'c'b''c b a
```

with atoms b, b', and b'' in GpC linkages, whereas a, a', c, and c' are in CpG linkages. The b and c type atoms are the best models for internal phosphates in a long DNA chain since a and a' are near the chain termini. The internal phosphates are in the ratio 2:1, GpC:CpG, for the hexamer, and 3:2 for the octamer duplex [and 1:1 in poly(dG-dC) (15-17, 19)]. The ratio of peaks (1+2) to 4 to 5 is 2:1:1 for the hexamer and 3:2:1 for the octamer

(see Figures 1 and 2; accuracy of peak areas within 5%), so we attribute peaks (1+2) to the GpC phosphates, peak 4 to the internal CpG phosphates, and peak 5 to one of the CpG phospates near the chain ends. This analysis is the simplest interpretation of the intensity ratios, and does not constitue an unequivocal assignment. The remaining CpG phosphate (a or a') must resonate under peak 3 which also contains the B-form signals. The fraction of the Z-form present can be calculated by subtracting the contribution or a single Z-form phosphate atom from peak 3; the remaining area represents the B-form contribution. The data are listed in Table 1 for some temperatures, the others are represented in Fig. 4 (see below). The percentages at 30°C are supported by integration of $^{13}C$ NMR peaks that originate from the B and Z-forms in 4M NaCl (28).

Temperature dependence

The temperature dependence in the spectra demonstrates that the slow equilibrium between right-handed B-DNA and left-handed Z-DNA conformations in 4M NaCl shifts dramatically toward the latter with decreasing temperature. This is in agreement with the results of temperature measurements of the circular dichroism

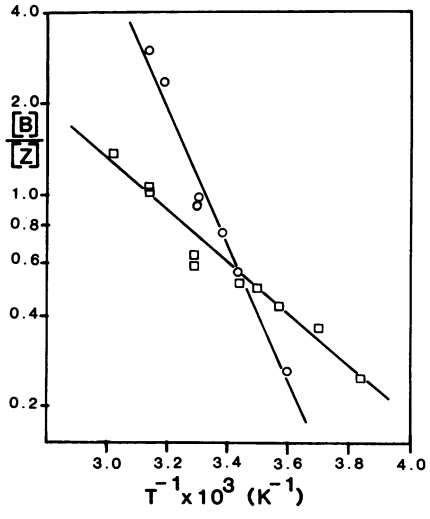


Figure 4. Plot of ln([B]/[Z]) vs. [1/T(°K)]x10$^3$, where [B] and [Z] are the fractons of B and Z DNA. Circles, $(dC-dG)_3$; squares, $(dC-dG)_4$.

of $(dC\text{-}dG)_3$ in 2.8M NaCl (29), and is in contrast to the trend for poly(dG-dC) nucleotides, for which elevated temperature shifts the B-Z equilibrium toward the Z-form (17). It is worth mentioning that signals from the Z form of $(dC\text{-}dG)_4$ disappear completely at 66 °C (the melting temperature should be much higher since it is 90°C in 0.1M NaCl; 28), and that the transition is entirely reversible. For $(dC\text{-}dG)_3$, this temperature is 51 °C on warming, cooling the sample brings back the Z form peaks at 45 °C. The temperature dependence of the equilibrium constant K=[B]/[Z], where [B] and [Z] are the populations of the B and Z forms, allows estimation of the thermodynamic parameters $\Delta H$ and $\Delta S$ for the Z to B transition. Van't Hoff plots of lnK *vs.* inverse temperature give straight lines for $(dC\text{-}dG)_3$ and $(dC\text{-}dG)_4$ (Fig. 4). The values of enthalpies for $(dC\text{-}dG)_3$ and $(dC\text{-}dG)_4$ are 44 and 17 kJ/mol, respectively, and the entropy terms are 146 and 56 J/mol·K, respectively. These data indicate that in these systems the enthalpy term favors the Z-form, while the entropy term favors the B-DNA conformation.

Finally, at the DNA concentrations studied here, both duplexes do not undergo B to Z transformation, but precipitate from solution, after addition of up to 3.5mM of $Co(NH_3)_6Cl_3$ in the temperature range from 0°C to 60°C. In 6M LiCl solution only ca. 7-10% of the octamer exists in the Z-form. Here also lower temperature stablilizes the Z-form [e.g. $(dC\text{-}dG)_4$ Z-form 7% at 0°C, 0% at 50°C].

Relaxation

Tne $^{31}P$ relaxation data are collected in Table 1. The $T_1$ data show an increase in relaxation rates ($1/T_1$) observed at higher temperature and at higher field (145.7 vs. 101.7 MHz). The $T_1$ data are not consistent with interpretation of the results in terms of a single correlation time in the fast motional region either for dipolar (DD) or chemical shift anisotropy (CSA) relaxation [CSA and DD relaxation follow similar expressions for the spectral densities (25,26, c.f. discussion in ref. 30)]. We are currently expanding our relaxation measurements in an attempt to separate various relaxation mechanism contributions (CSA and DD) and to find optimal models to fit the relaxation data. One should be aware, however, of considerable complications

in the analysis and the interpretation of $^{31}P$ relaxation data of nucleic acids in solution (31-34), especially where different mechanisms contribute to the relaxation (34). However, it appears that regardless of the details of the model for motions of DNA [which may be best probed by multi-field $^{13}C$ measurements (31)], NMR relaxation experiments have produced a relatively homogenous picture of the magnitudes of significant motional frequencies, if not a consistent picture of the forms of these motions.

Both B and Z forms of the hexanucleotide double helix should be roughly spherical, according to the data from x-ray diffraction: B-DNA, 2 nm diameter and 2.04 nm length (35); Z-DNA, 1.8 nm diameter and 2.2 nm length (36). This agrees with the analysis of the $^{13}C$ relaxation data in low salt solution using our dynamics modeling program MOLDYN (37); the molecules tumble essentially isotropically with an overall correlation time ( $\tau_R$) near 5 ns (28). Base and sugar carbons appear to undergo restricted internal motions with the correlation time $\tau_G \leq 1$ ns; with sugar carbons experiencing higher amplitude motions. The $^{31}P$ relaxation data are also consistent with this interpretation. The temperature dependence of rotational correlation times ( $\tau_{eff}$, containing contributions from $\tau_G$+ $\tau_R$), obtained by measuring $T_1$ at varying temperature, can be used to estimate apparent activation energies for molecular motion using the Arrhenius equation. For both duplexes, plots of relaxation rate *vs.* inverse absolute temperature give straight lines in the temperature ranges measured. By assuming $T_1$ increases with increasing $\tau_{eff}$ ( $\tau_{eff}$ is just to the right of the $^{31}P$ $T_1$ minimum), the apparent activation energies calculated from the plots for GpC linkages are 0 kJ/mol [$(dC\text{-}dG)_3$] and 3.8 kJ/mol [$(dC\text{-}dG)_4$]; and for CpG phosphates $E_a$=9.2 kJ/mol [$(dC\text{-}dG)_3$] and 10.0 kJ/mol [$(dC\text{-}dG)_4$]. Since $\tau_R$ is the same for both types of phosphorus atoms these results may be taken to indicate that the GpC phosphodiester groups show more internal motions than the CpG linkages. The x-ray diffraction data for $(dC\text{-}dG)_3$ in the solid state also indicate that the GpC links exist in 2 distinct conformations, $Z_I$ and $Z_{II}$, whereas CpG has only one conformation (36). It is possible that the extra motion in GpC links, indicated from

Table 2. $^{31}P$ Chemical Shifts of CG oligonucleotides at 30°C.

| | |
|---|---|
| $(dC\text{-}dG)_3$ | -4.01,-4.15,-4.11,-4.13,-4.23 |
| $(dC\text{-}dG)_4$ | -4.03,-4.12,-4.13,-4.16,-4.18,-4.24,-4.26 |

the $^{31}P$ relaxation data, is related to rapid jumps between conformations similar to the x-ray structures.

Low Salt $^{31}P$ Spectra

Although compared to the high salt solutions the range of chemical shift of the B-form is small ( $\Delta\delta$ =0.23ppm), in low salt solutions the unbroadened (~7 Hz) B-form $^{31}P$ resonances of phosphodiester groups in $(dC\text{-}dG)_3$ and $(dC\text{-}dG)_4$ show sufficient dispersion in chemical shift to permit detection of all individual resonances (Table 2). Recently, the assignment of $^{31}P$ resonances in $(dC\text{-}dG)_2$ has been reported using selective labeling by $^{17}O$ of the nonesterified phosphoryl oxygen of sugar-phosphorous backbone (14). These data, however, do not permit assignment of the phosphorus resonances in $(dC\text{-}dG)_3$ and $(dC\text{-}dG)_4$. Several assignments for $(dC\text{-}dG)_2$ and $(dC\text{-}dG)_3$ have also been made using the heteronuclear decoupling method (11).

EXPERIMENTAL

Deoxy $(dC\text{-}dG)_3$ and $(dC\text{-}dG)_4$ were synthesized by the phosphate triester method in solution (38). High purity NaCl purchased from Johnson Matthey Chemicals Limited was used in the preparations of the NMR solutions. NMR spectra were recorded with proton broadband decoupling on a Bruker WM-360 spectrometer with a widebore 8.46 Tesla magnet and are referenced relative to an internal standard or trimethyl phosphate (TMP) at 30°C. A standard inversion-recovery (IRFT) method was used for the $T_1$ determination. Both $T_1$'s and NOE's were measured using a gated two-level decoupling procedure and cycling through a list of delays several times during the measurements in order to alleviate heating problems (c.f. ref.6). The temperature was directly measured in the probe on a sample of TMP in appropiate salt solutions. The temperature was controlled within ±1 °C. Samples were run in 10 mm tubes, concentrations: 11mM for $(dC\text{-}dG)_3$ and 8mM for $(dC\text{-}dG)_4$ (in single strands); 25 mM cacodylate, 1 mM EDTA, 25% $D_2O$, 0.1 or 4M NaCL;

pH 7.1; 1.6 ml total volume. The sample of $(dC\text{-}dG)_3$ in 4M NaCl began to precipitate within several hours after preparation, so it was redissolved by heating above 60°C prior to the high-salt experiments.

ACKNOWLEDGMENTS

We thank Anil Kumar for discussion and assistance with the MOLDYN calculations, Dr. M. Petersheim for supplying unpublished data and Dr. L. S. Kan for a preprint. This work was supported by NIH Grant GM32691 (to PNB), NIH Grants GM29778, RR01317, and NSF Grant 81-05109 (to GCL), NIH Grant CA29753 (AH-JW), and American Cancer Society Grant NP 3348 (to AH-JW and A. Rich).

*To whom correspondence should be addressed

REFERENCES

1. Shindo,H., Simpson,R.T., & Cohen,J.S. (1979) J. Biol. Chem. 254, 8125-8128.
2. Patel,D.J. & Canuel,L.L. (1979) Eur. J. Biochem. 96, 267-276.
3. Cohen,J.S., Wooten,J.B., & Chatterjee,C.L. (1981) Biochemistry 20, 3049-3055.
4. Kypr,J., Vorlickova,M., Budesinsky,M., & Sklenar,V. (1981) Biochem. Biophys. Res. Commun. 99, 1254-1264.
5. Gorenstein,D.G. (1981) Annu. Rev. Biophys. Bioeng. 10, 355-386.
6. Gorenstein,D.G., Luxon,B.A., Goldfield,E.M., Lai,K., & Vegeais, D. (1982) Biochemistry 21, 580-589.
7. Eckstein,F. & Jovin,T.M. (1983) Biochemistry 22, 4546-4550.
8. Chen,C.-W. & Cohen,J.S. (1983) Biopolymers 22, 879-893.
9. Saenger,W. (1984) Principles of Nucleic Acid Structure, ch. 11 & 12, Springer-Verlag, New York.
10. Cheng,D.M., Kan,L.-S., Miller,P.S., Leutzinger,E.E., & Ts'o, P.O.P. (1982) Biopolymers 21, 697-701.
11. Cheng,D.M, Kan,L.-S., Frechet,D., Ts'o,P.O.P., Uesugi,S., Shida,T., & Ikehara,M. (1984) Biopolymers 23, (in press).
12. Cheng,D.M., Kan,L.-S., Iuorno,V.L., & Ts'o,P.O.P. (1984) Biopolymers 23, (in press).
13. Pardi,A., Walker,R., Rapoport,H., Wider,G., & Wuethrich,K.J. (1983) J. Am. Chem. Soc. 105, 1652.
14. Petersheim,M., Mehdi,S., & Gerlt,J.A. (1984) J. Am. Chem. Soc. (submitted).
15. Patel,D.J., Canuel,L.L., & Pohl,F.M. (1979) Proc. Natl. Acad. Sci. USA 76, 2508-2511.
16. Simpson,R.T., & Shindo,H. (1980) Nucleic Acids Res. 8, 2093-2103.
17. Patel,D.J., Kozlowski,S.A., Nordheim,A., & Rich,A. (1982) Proc. Natl. Acad. Sci. USA 79, 1413-1417.
18. Chen,C.-W., Cohen,J.S., & Behe,M. (1983) Biochemistry 22, 2136-2142.
19. Jovin,T.M., van de Sande,J.H., Zarling,D.A., Arndt-Jovin,D.J., Eckstein,F., Fueldner,H.H., Greider,C., Greiger,I., Hamori,E.,

Kalisch,B., McIntosh,L.P., & Robert-Nicoud,M. (1983) Cold Spring Harbor Symp. Quant. Biol. 47, 143-154.
20. Hartmann,B., Thuong,N.T., Pouyet,J., Ptak,M., & Leng,M. (1983) Nucleic Acids Res. 11, 4453-4466.
21. McIntosh,L.P., Grieger,I., Eckstein,F., Zarling,D.A., van de Sande,J.H., & Jovin,T.M. (1983) Nature 304, 83-86.
22. Feigon,J., Wang,A.H.-J., van der Marel,G.A., van Boom,J.H., & Rich,A. (1984) Nucleic Acids Res. 12, 1243-1263.
23. Moeller,A., Nordheim,A., Kozlowski,S.A., Patel,D.J., & Rich,A. (1984) Biochemistry 23, 54-62.
24. IUPAC-IUB Joint Commission on Biochemical Nomenclature (1983) Eur. J. Biochem. 131, 9-15.
25. James,T.L. (1983) Bulletin of Magn. Reson. 4, 119-159.
26. Shindo,H. (1983) in Magnetic Resonance in Biology, Cohen,J.S., Ed., Wiley, New York, vol. 2, 95-129.
27. Kumar,A., Sotak,C.H., Dumoulin,C.L., & Levy,G.C. (1983) Computer Enhanced. Spectr. 1, 107-114.
28. Borer,P.N., Zanatta,N., Holak,T.A., Levy,G.C., van Boom,J.H., & Wang,A.H.-J. (1984) J. Biomol. Structure & Dynamics (in press)
29. Quadrifoglio,F., Manzini,G., Vasser,M., Dinkelspiel,K., & Crea,R. (1981) Nucleic Acids Res. 9, 2195-2206.
30. Hayashi,F., Akasaka,K., & Hatano,H. (1977) Biopolymers 16, 655-667.
31. Levy,G.C., Craik,D.J., Kumar,A., & London,R.E. (1983) Biopolymers 22, 2703-2726.
32. Levy,G.C., Marchetti,P.S., Ejchart,A., Levy,L.F., Kumar,A., Hilliard,P.R., & Rill,R.L. (1983) J. Biomol. Struct. Dynamics 1, 795-808.
33. Rill,R.L., Hilliard,P.R., & Levy,G.C. (1983) J. Biol. Chem. 258, 250 256.
34. Allison,S.A., Shibata,J.H., Wilcoxon,J., & Schurr,J.M. (1982) Biopolymers 21, 729-762.
35. Arnott,S., & Hukins,D.W.L. (1972) Bichem. Biophys. Res. Commun. 47, 1504-1509.
36. Wang,A.H.-J., Quigley,G.J., Kolpak,F.J., van der Marel,G., van Boom,J.H., & Rich,A. (1980) Science 211, 171-176.
37. Craik,D.J., Kumar,A., Levy,G.C. (1983) J. Chem. Inf. Comput. Sci. 23, 30-38.
38. van der Marel,G., van Boeckel,C.A.A., Wille,G., & van Boom, J.H. (1981) Tetrahedron Letters 22, 3887-3890.