Volume 16 Number 5 1988 Nucleic Acids Research

Hydrogen-bonding effects and ^{13}C-NMR of the DNA double helix

Philip N.Borer*, Steven R.LaPlante, Nilo Zanatta+ and George C.Levy

NIH Research Resource, Bowne Hall, Syracuse University, Syracuse, NY 13244-1200, USA

Received September 3, 1987; Revised and Accepted January 25, 1988

ABSTRACT

^{13}C-nmr chemical shifts of the nucleotides in DNA are sensitive to hydrogen bonding, especially for three of the carbons immediately bonded to exocyclic oxygen or nitrogen atoms acting as H-bond acceptors or donors. GuoC2, GuoC6 and ThdC4 are strongly deshielded (about 1 ppm) upon Watson-Crick pairing in oligodeoxynucleotide duplexes, regardless of the base sequence. Deshielding at these sites may be useful to distinguish bases involved in Watson-Crick pairs from unpaired bases.

INTRODUCTION

Both nucleic acid self-association and ligand binding are mediated by a combination of hydrogen bonding, van der Waals, electrostatic and various solute-water interactions. Determination of the details of these interactions would assist in designing drugs and proteins with altered specificities. X-ray diffraction analyses have been published for complexes between DNA fragments and the drugs netropsin (1,2), daunomycin (3), triostin A (4, 5) and echinomycin (6), for the restriction endonuclease EcoR1 (7), as well as for a number of self-complexes of DNA and RNA (8 - 16). These studies suggest that a lattice of hydrogen bonds within and between the molecules is the most important factor in organizing the complexes in their native three-dimensional structures.

The work presented here shows that ^{13}C-nmr spectra may be useful in locating the sites of H-bond formation in the nucleic acid bases. Specific enrichment with ^{13}C should allow future application of the method in high-molecular weight complexes where three-dimensional structure determination by ^{1}H-NMR (17 - 22) is not practical.

This report describes some aspects of the ^{13}C-nmr spectra of

several DNA oligonucleotide duplexes giving particular attention to the changes in chemical shift (δ) of the base carbons upon formation of hydrogen-bonded, base-stacked structures. Base stacking tends to shield base carbons (smaller δ) consistent with ring-current and steric shielding arguments (23). Hydrogen bond formation usually deshields carbons near certain of the H-bond donor and acceptor heteroatoms (see below). Shielding and deshielding effects cancel each other at some locations, but a substantial net deshielding effect at a particular carbon nucleus appears to be diagnostic of a hydrogen bond at a nearby site.

MATERIALS AND METHODS

Spectra were acquired at 90.56 MHz (Bruker WM-360) or 125.8 MHz (General Electric GN-500). Conditions for the WM-360: 2 ml samples at 10-20mM in single strands (50-100mg) in 10 mm diameter sample tubes, 10,000-50,000 transients were averaged with 4kx2 (complex) data tables, 90° pulses with "WALTZ" decoupling; for the GN-500: 0.4 ml 10-20mM samples (10-20 mg) in 5 mm tubes, 5000-20,000 transients, 8kx2 data tables, 90° pulses with "MLEV" decoupling. Repetition rates were 1-2 s and decoupler heating was negligible. Duplex (1) $[d(TAGCGCTA)]_2$, was synthesized on a Biosearch model 8600 DNA synthesizer (20 umol scale) using the phosphoramidite protocol and purified on a Nucleogen DEAE 60-7 HPLC column (Machery-Nagel, CH_3CN/H_2O (20:80) + 20 mM phosphate, 0-1M LiCl gradient); (2), $[d(GGTATACC)]_2$, was synthesized from the phosphoramidites using the manual syringe method (25) and purified by C_8 reverse-phase HPLC using tetrabutylammonium acetate as an ion-pairing buffer in CH_3CN/H_2O gradients (the DEAE purification method gives superior results); duplex (3) $[d(CG)_3]_2$, was the kind gift of Prof. A.H.-J. Wang.

RESULTS

^{13}C Spectra and Class Assignments. ^{13}C-nmr offers a wide variety of atomic sites for observing nucleic acids, particularly in the bases. For instance, in guanine, only H8 and H1 (see Fig. 1a) are easily monitored in ^1H-spectra, whereas there are five easily studied ^{13}C-nuclei. The low-field base region of the 125.8 MHz ^{13}C-spectrum for $[d(TAGCGCTA)]_2$ is shown in Fig. 2, where it is seen that 17 lines can be distinguished for the 20

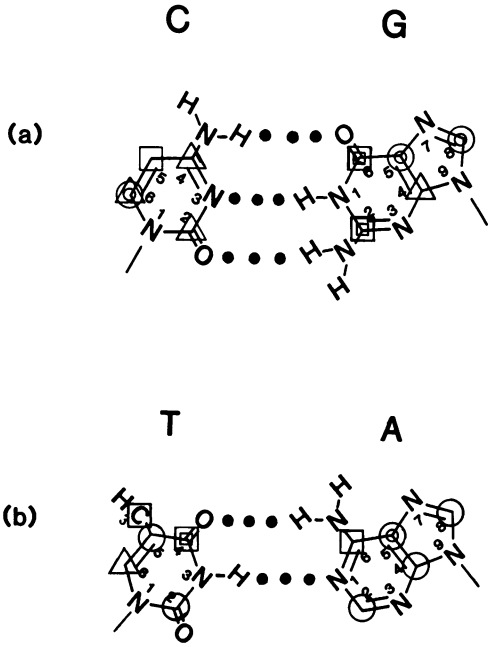

Figure 1. Base pairing schemes in (a) C-G and (b) T-A. $\Delta\delta_{av}$ values are indicated by symbols on each carbon: double square ≤ -0.8 ppm, single square ≤ -0.2 ppm, triangle ≤ 0 ppm, single circle $0 < \Delta\delta_{av} < 0.7$ ppm, double circle $\Delta\delta_{av} \geq 0.8$ ppm.

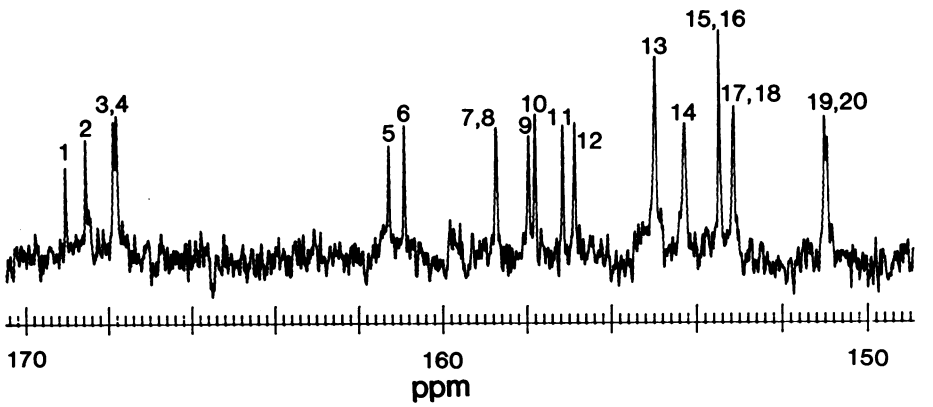

Figure 2. A section of the 125.8 MHz ^{13}C-nmr spectrum of $[d(TAGCGCTA)]_2$. 0.4 ml sample was 7mM in single strands, 0.15M NaCl, 0.01M cacodylate pH 7.2, 1 mM EDTA, 20% D_2O, 27°C, 15,700 acquisitions, 5.2 hr, 23 kHz spectral width, 16K real data points, 3 Hz exponential line broadening; δ is in ppm from sodium 3-(trimethylsilyl)-1-propanesulfonate (DSS) referenced through cacodylate at 20.11 ppm. Peak assignments: 1,2 TC4; 3,4 CC4; 5,6 GC6; 7,8 CC2; 9,10 AC6; 11,12 GC2; 13,14 AC2; 15-18 TC2, GC4; 19,20 AC4.

carbons in this partial spectrum. The 36 unique base carbons resonate over a 70 ppm range, in contrast with the ^{1}H-spectrum (not shown) where the C-H protons are dispersed over a 3 ppm range. The magnitude of sequence and "remote" structural effects is of the order of 1 to 2 ppm in both kinds of spectra, so, in general, assignment is much simpler for the ^{13}C-spectra. In fact, most carbon classes can be distinguished by comparisons between monomer spectra and duplexes with differing numbers of the bases. For example, $[d(CG)_3]_2$ contains only G and C carbons, so it is easy to distinguish these signals from the A and T classes (26, 27, P.N.B., S.R.L., N.Z., G.C.L., unpublished; here "class" means a single base position, e.g., GC8 designates the guanine carbon 8 class). It should be emphasized that the conclusions of this paper depend only on the proper assignment of carbons according to class, although in many cases using our methodology it is also possible to make unequivocal assignments of resonances to individual carbons (Laplante et al., submitted).

An additional level of confidence in the class assignments is obtained by comparing the chemical shift vs. temperature profiles for the resonances. The profiles for the G carbons of duplexes 1, 2, and 3 are shown in Fig. 3, where it can be seen that the curves in the C6 and C2 regions of chemical shift always have negative slopes, the profiles in the C4 region have very shallow slopes, while the C8 and C5 curves exhibit pronounced positive slopes. The similarity of the profiles within a class is readily apparent. Profiles for the A, C and T bases (not shown) also exhibit strong similarities within each class. It is apparent that the helix-coil transition strongly affects carbon chemical shifts and that the effects are more closely related to carbon class than to sequence. The similarities in δ vs. T profiles within a class holds for each of the duplexes examined thus far (LaPlante et al. submitted; Borer et al., unpublished).

DISCUSSION

Base Stacking and H-bonding Effects on ^{13}C Chemical Shifts.

Most non-exchangeable base protons in oligonucleotide duplexes exhibit δ vs. T profiles with positive slopes, similar to those displayed by GC8 and GC5 in Fig. 3 (24, 28, 29). Such profiles

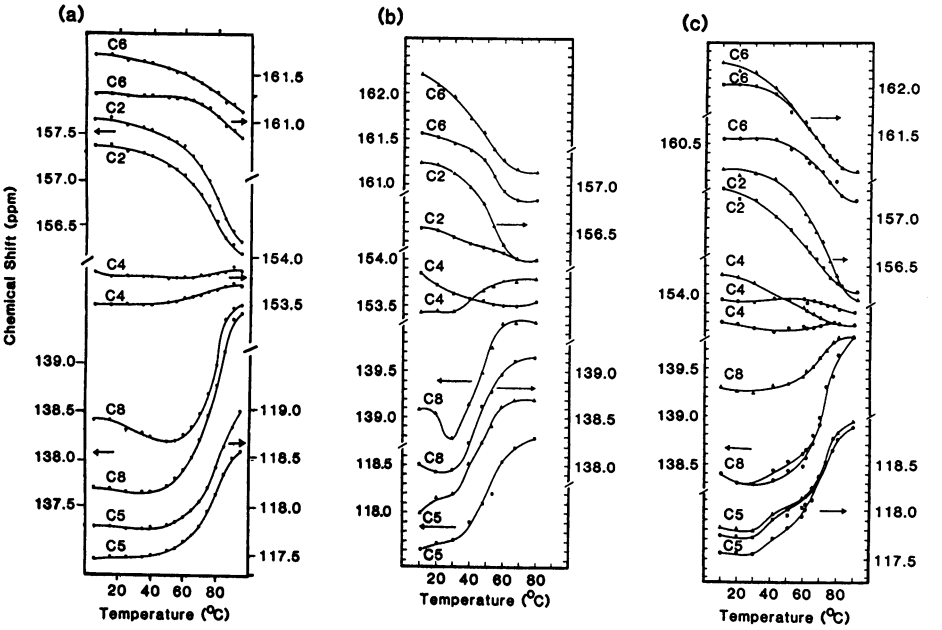

Figure 3. Chemical shift vs. temperature profiles for the guanine carbons of three oligonucleotide duplexes; shielding increases toward the bottom of each panel.

are consistent with a decrease in ring current and steric shielding upon melting the double helix (23, 30). Some of the ^{13}C profiles, by contrast, have large negative slopes (e.g., GC6 and GC2 in Fig. 3), which we attribute to breaking H-bonds upon strand dissociation. Shallow profiles such as those observed for the GC4 (Fig. 3) apparently have nearly equal contributions from the two effects. Crossovers in the melting profiles were resolved by determining the melting curves in the absence of added salt. This reduces the T_m but the profiles for a given carbon remain very similar.

It is interesting to know which of the carbons are most sensitive to hydrogen bonding. Figure 1 illustrates the variation in $\Delta\delta$ in a graphic format. Close examination of the figure shows that, except for AC2 and TC2, carbons adjacent to heteroatoms engaged in H-bonds show negative (squares) or occasionally negative (triangles) $\Delta\delta$ values. Most of the

Table 1. Changes in chemical shift (Δδ, ppm) for base carbons upon thermally induced base-pair disruption.

	Guo		Cyd		Ado		Thd	
	Av[a]	Calc[b]	Av[a]	Calc[b]	Av[a]	Calc[b]	Av[a]	Calc[b]
C2	-1.1	-3.1	0.4[c]	-2.0	0.7	-1.0	0.6	-0.6
C4	0.0[c]	-1.9	0.3[c]	-1.8	0.4	-0.4	-0.9	-2.5
C5	1.2	1.4	-0.4[c]	-1.5	0.8	-0.3	0.4	-0.4
C6	-0.8	-3.2	0.9[c]	-0.5	-0.2[c]	-1.7	0.3[c]	-1.5
C8	1.3	0.9	--	--	0.2	0.3	--	--
CH_3	--	--	--	--	--	--	-0.6	nd[d]

[a] Averages of $\Delta\delta = \delta_{high\ temp} - \delta_{low\ temp}$ for duplexes 1, 2, and 3, which involve 4 to 7 carbon observations for each entry (only two duplexes have A and T carbons). Standard deviations are 0.1 to 0.7 ppm and reflect sequence and end effects.

[b] Calculated difference $\delta_{monomer} - \delta_{dimer}$; these are differences in magnetic shielding constants calculated by a self-consistent perturbation method using a minimal basis set of gauge-invariant atomic orbitals, with compensation for basis set extension errors (33).

[c] Some measured Δδ values are negative, some are positive.

[d] not determined

positive (circles) Δδ values occur at carbons distant from the H-bond sites.

Table 1 collects averages of the measured Δδ values (δ at high temperature where the strands are dissociated minus δ in the duplex at low temperature) for each of the carbon classes in the three duplexes; in the case of duplex 1, due to its high T_m, the high temperature δ values were obtained from the profiles obtained in the absence of added salt. Table 1 compares the average Δδ with values calculated from *ab initio* quantum mechanical calculations for the disruption of isolated Watson-Crick base pairs (31). These calculations suggest that breaking H-bonds should shield (negative Δδ) the ^{13}C nuclei of the bases, with only a few exceptions. Stacking contributions range up to +2 ppm as judged by results on stacked, single-stranded trinucleotides (23). The stacking effects depend strongly on the mutual orientation of the bases and result from some combination of polarization, ring-current and local magnetic anisotropies as well as steric contact between the van der Waals surfaces of the bases (23, 32). These stacking effects should oppose the negative H-bonding contributions.

The quantum mechanical calculations of H-bonding include effects on a given base from its Watson-Crick partner due to: (i) a "geometric" factor that includes ring current and local magnetic anisotropies (which include effects due to carbonyls and other localized groups), (ii) polarization, (iii) charge transfer plus exchange, and (iv) a counterpoise correction factor that compensates for the use of a limited basis set of atomic orbitals. The sum of the terms is reported in Table 1 and is dominated by the polarization and charge transfer terms. Another way of saying this is that changes in electron density due to H-bonding make the largest changes in the magnetic shielding of the ^{13}C-nuclei in the bases. The geometric term averages only -0.2 ppm (31), so through-space effects from the opposite strand are small in comparison to the primary anisotropic effects; thus contributions to base carbon shifts from propeller twist and local helix distortions are likely to be small.

There is fair agreement with the measurements on the duplexes and the quantum mechanical calculations, given that the latter use an incomplete set of basis orbitals and make no corrections for solvation effects, and that positive contributions are expected for stacking. The most negative calculated values are for GC2, GC6, and TC4, exactly those nuclei which are the most shielded upon H-bond disruption in the duplexes. The same nuclei are directly bonded to H-bond donors or acceptors (see Fig. 2) and are probably the most reliable indicators of Watson-Crick hydrogen bonds. TCH_3, CC5 and AC6 usually show negative δ *vs.* T profiles upon duplex melting, positive profiles occurring only when the base has an A neighbor or is at the end of a chain, where one might expect H-bonds to be partially frayed.

The carbons attached to exocyclic H-bonding sites in the cytidine base (CC2 and CC4) are unusual in that they are calculated to have sizable negative $\Delta\delta$ values, yet the measurements average to small positive values. Corresponding sites on the other bases (GC2, GC6, TC4, and AC6; see Fig. 1) all average to negative $\Delta\delta$ values. In our limited sample of sequences, negative $\Delta\delta$ values are observed for CC2 and CC4 only at the 5'-terminal residue of $[d(CG)_3]_2$ (-0.18 and -0.30 ppm, respectively); this terminal C-base is weakly stacked in the standard B-DNA model. We conclude that the sign of the H-bonding effect is the same at

CC2 and CC4 as it is for the others, and suggest that it has a smaller magnitude than predicted by the quantum mechanical calculations. The H-bonding effect at CC2, CC4 and probably also AC6 can be easily counteracted by stacking effects.

The six carbon sites indicated by boldface type in Table 1 should be the most useful ^{13}C-markers for Watson-Crick hydrogen bonding; caution dictates that TCH_3, CC5 and AC6 should be used only as supporting evidence, not as an unequivocal indication of Watson-Crick H-bonding. We are currently engaged in further experiments to test the validity of using these shifts in identifying H-bonding sites: (i) with duplexes composed of distinguishable strands, where complex formation can be studied at room temperature; (ii) with RNA duplexes, which should have very different stacking geometries; and (iii) with specific drug molecules which are known to associate with the DNA bases by H-bonding to the bases. It is noteworthy that TCH_3 and CC5 are not bonded to heteroatoms involved in H-bonds, yet exhibit shielding upon duplex melting. Thus it appears that H-bonding can exert effects through the aromatic electron systems of the bases at some considerable distance from the atoms that are directly involved in the interaction.

Unique Assignments. Recently, two-dimensional NMR methods have been developed that use polarization transfer (PT) from the proton manifold to observe ^{13}C-nuclei (33-36). Unequivocal ^{1}H-assignments have been made for 1 (S.R.L., G.C.L. & P.N.B., unpublished) and 2 (37); the ^{1}H assignments have been transferred to their attached carbons for the AC2, AC8, GC8, CC5, CC6, TC6 and TCH_3 (S.R.L., J. Ashcroft, D. Cowburn, G.C.L. & P.N.B., submitted). There are 14 of these carbons on each of the two duplexes, so there are potentially 28 tests of the accuracy of the assignment methods outlined above. In every case, the class assignments made by visual inspection of the 1-D spectra agree completely with the PT assignments. It should be obvious that careful comparisons of base carbon spectra and δ *vs.* T profiles can produce accurate assignments; this arises from the additional dispersion in ^{13}C-spectra in contrast to ^{1}H-spectra.

CONCLUSIONS

We conclude that six ^{13}C nuclei: GC2, GC6, TC4, TCH_3, CC5,

and AC6 are deshielded (chemical shifts become larger) in response to formation of Watson-Crick hydrogen bonds. The most useful markers for G-C pairing are the GC2 and GC6 signals, while TC4 is the most reliable indicator of A-T pairing. It may be possible to use characteristic changes in ^{13}C chemical shift to distinguish some of the sites involved in hydrogen bonding interactions in nucleic acid self-complexes, and in complexes with proteins, drugs, and various other ligands. Given sufficient local mobility, it should be possible to extend the method to complexes in the molecular weight range of 100,000 Daltons. For such large complexes, use of ^{13}C-enriched nucleotides would increase the sensitivity of chemical shift measurement by a factor of 100, greatly simplify the spectra, and clarify the assignments.

This work was supported by National Institutes of Health Grants GM32691 (to P.N.B.) and GM35069 & RR01317 (to G.C.L. and P.N.B.); N.Z. held a fellowship from the Conselho Nacional de Desenvolvimento Cientifico e Techologico (CNPq) of Brazil. Some of the nmr measurements were performed on an instrument purchased through grants from the Division of Research Resources of the NIH and from the Biophysics Program of the National Science Foundation. We thank Prof. Andy Wang (M.I.T.) for supplying one of the DNA duplexes.

*To whom reprint requests should be sent

+Current address: Universidade Federal de Santa Maria, Departamento de Quimica, 97100-Santa Maria-RS, Brazil

REFERENCES

1. Kopka, M.L., Yoon, C., Goodsell, D., Pjura, P. & Dickerson, R.E. (1985) Proc. Natl. Acad. Sci. USA, 82, 1376-1380.
2. Kopka, M.L., Yoon, C., Goodsell, D., Pjura, P. & Dickerson, R.E. (1985) J. Mol. Biol., 183, 553-563.
3. Quigley, G.J., Wang, A.H.-J., Ughetto, G., van der Marel, G., van Boom, J.H. & Rich, A. (1980) Proc. Natl. Acad. Sci. USA, 77, 7204-7208.
4. Wang, A.H.-J., Ughetto, G., Quigley, G.J., Hakoshima, K., van der Marel, G.A., van Boom, J.H. & Rich, A. (1984) Science, 225, 1115-1121.
5. Quigley, G.J., Ughetto, G., van der Marel, G.A., van Boom, J.H., Wang, A.H.-J. & Rich, A. (1986) Science, 232, 1255-1258.
6. Ughetto, G., Wang, A.H.-J., Quigley, G.J., van der Marel, G.A., van Boom & Rich, A. (1985) Nuceic Acids Res., 13, 2305-2323.
7. Frederick, C.A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobson, L., Wang, B.-C., Greene, P., Boyer, H.W. & Rosenberg, J.M. (1984) Nature (London), 309, 327-331.
8. Dickerson, R.E.., Drew, H.R., Conner, B.N., Wing, R.M.,

Fratini, A.V. & Kopka, M.L. (1982) Science, 216, 475-485.
9. Drew, H., Takano, T., Tanaka, S., Itakura, K. & Dickerson, R.E. Nature (London), 286, 567-573.
10. Wang, A.H.-J., Quigley, G.J., Kolpak, F.J., van Der Marel, G., van Boom, J.H. & Rich A. Science, 211, 171-176.
11. Wang, A.H.-J., Fujii, S., van Boom, J.H. & Rich, A. Proc. Natl. Acad. Sci. USA, 79, 3968-3972.
12. Shakked, Z., Rabinovich, D. Cruse, W.B.T., Egert, E., Kennard, O., Sala, G., Salisbury, S.A. & Viswamitra, M.A. Proc. R. Soc. Lond. Ser. B, 213, 479-487.
13. Wang, A.H.-J., Fujii, S., van Boom, J.H., van der Marel, G.A., van Boeckel, S.A.A. & Rich A. (1982) Nature (London), 299, 601-604.
14. Holbrook, S.R. Sussman, J.L., Warrant, R.W. & Kim, S.-H. (1978) J. Mol. Biol., 123, 631-660.
15. Hingerty, B., Brown, R.S. & Jack, A. (1978) J. Mol. Biol., 124, 523-534.
16. Saenger, W. (1984) Principles of Nucleic Acids Structure, (Springer-Verlag, New York)
17. Gronenborn, A.M. & Clore, G.M. (1985) Prog. Nuclear Magnetic Resonance Spectroscopy, 17,. 1-32.
18. Wemmer, D.E., Chou, S.-C., Hare, D.R. & Reid, B.R. (1984) Biochemistry, 23, 2262-2268.
19. Scheek, R.M., Boelens, R., Russo, N., van Boom, J.H. & Kaptein, R. (1984) Biochemistry, 23, 1371-1376.
20. Kumar, A., Wagner, G., Ernst, R.R. & Wuthrich, K. (1981) J. Am. Chem. Soc., 103, 3654-3658.
21. Williamson, M.P., Marion, D. & Wuthrich, K. (1984) J. Mol. Biol., 173, 341-359.
22. Jamin, N., James, T.L. & Zon, G. (1985) Eur. J. Biochem., 152, 157-166.
23. Stone, M.P., Winkle, S.A. & Borer, P.N. (1984) J. Biomol. Sruct. Dyn., 3, 767-781.
24. Patel, D.J. (1979) Eur. J. Biochem., 99, 369-378.
25. Tanaka, T. & Letsinger, R.L. (1982) Nucleic Acids Res., 10, 3249-3260.
26. Zanatta, N.(1986) Ph.D. Thesis, Syracuse University.
27. Borer, P.N., Zanatta, N., Holak, T.A., Levy, G.C., van Boom, J.H. & Wang, A.H.-J (1984) J. Biomol. Struct. Dyn., 1, 1373-1386.
28. Cruz, P., Bubienko, E. & Borer, P.N. (1982) Nature (London), 298, 198-200.
29. Borer, P.N., Kan, L.S. & Ts'o, P.O.P. (1975) Biochemistry, 14, 4847-4863.
30. Giessner-Prettre, C. & Pullman, B. (1976) Biochem. Biophys. Res. Commun., 70, 578-581.
31. Giessner-Prettre, C. (1984) J. Biomol. Struct. Dyn., 2, 233-248.
32. Giessner-Prettre, C. (1986) J. Biomol. Struct. Dyn., 4, 99-110.
33. Bax, A., Griffey, R.H. & Hawkins, B.L. (1983) J. Am. Chem. Soc., 105, 7188-7190.
34. Live, D.H., Davis, D.G., Agosta, W.C. & Cowburn, D. (1984) J. Am. Chem. Soc., 106, 6104-6105.
35. Leupin, W., Wagner, G., Denny, W.A. & Wuthrich, K. (1987) Nucleic Acids Res., 15, 267.
36. Sklenar, V. & Bax, A. (1987) J. Am. Chem. Soc., 109, 2221.
37. Patel, D.J., Shapiro, L. & Hare, D. (1986) Biopolymers, 25, 693-706.