Volume 12 Number 21 1984 Nucleic Acids Research

Helix opening in deoxyribonucleic acid from a proton nuclear magnetic resonance study of imino and amino protons in $d(CG)_3$.

G.Victor Fazakerley[1]*, Gijs A.van der Marel[2], Jacques H.van Boom[2] and Wilhelm Guschlbauer[1]

[1]Service de Biochimie, Bat. 142, Département de Biologie, Centre d'Etudes Nucléaires de Saclay, F-91191 Gif-sur-Yvette Cédex, France, and [2]Gorlaeus Laboratories, State University of Leiden, P.O. Box 9502, NL-2300RA Leiden, The Netherlands

Received 31 July 1984; Revised and Accepted 9 October 1984

Abstract: All exchangeable protons in a short DNA helix, $d(CG)_3$ sodium salt, have been studied by proton nuclear magnetic resonance. The cytidine and guanosine amino protons have been assigned for the first time. As a function of temperature the cytidine amino protons and the imino protons behave very similarly, their relaxation is dominated by exchange with solvent above 30°C. The guanosine amino protons, however, show that helix opening can only be described by a multistate model. The most rapid process observed is probably a twist about the helix axis which lengthens or breaks the guanosine amino hydrogen bond and allows rotation of the amino group. The second fastest process is a scissor opening into the major groove which gives rise to solvent exchange with the imino and cytidine amino protons. The slowest process observed is the complete base pair opening in which the guanosine amino protons also exchange with solvent. For the ammonium salt of the oligonucleotide, a specific ammonium ion complex is observed which at low temperature may catalyze exchange of the guanosine amino protons with the protons of the ammonium ion, but retards exchange with solvent. The complex appears to be specific for the sequence d(CpG).

INTRODUCTION

Double stranded nucleic acids of defined sequence have been the object of many studies by nuclear magnetic resonance (NMR) with a view of establishing their conformational states and the factors governing the opening of a specific base pair or complete strand separation (1). These latter kinetic processes are primarily obtained from the interpretation of relaxation data of the hydrogen bonded imino protons. Opening of the double strand would also be expected to facilitate exchange of the amino protons. These protons have been observed in t-RNA-Phe (2), oligonucleotides (3) and $d(CG)_3$ (4), although no assignment or kinetic studies were made.

Tritium or deuterium exchange experiments on polymers (5-7) showed different rates of exchange for different types of exchangeable protons and also the effect of catalysts upon these rates. Thus under certain conditions exchange is observed to take place each time a base pair opens - the open limited pathway. However, assignment by comparison of the rates of

different polymers can only be tentative and this study reports results on all the exchangeable protons in $d(CG)_3$.

The lifetimes of imino protons in short helical DNA fragments have been interpreted in terms of processes involving opening of individual base pairs or of complete strand separation to give solvent access to these protons (8-13). A third mechanism which would permit solvent access of the imino protons exists, that of a scissor opening into either the major or minor groove in which the base pairs would still be held together by a hydrogen bond involving one amino proton. In the presence of a catalyst where the exchange rates of the imino and amino protons are limited by the accessibility to the solvent, we observe that the rates for the imino protons and the cytosine amino protons are approximately equal and that relaxation is dominated by exchange at temperatures above 30°C. The guanine amino protons, however, exchange more slowly, their relaxation being dominated by exchange only above 60°C. As this is not due to an inherently slower exchange rate of guanosine compared to cytidine amino protons, we interpret these observations in terms of a lower accessibility to solvent and a probable scissor opening of the base pairs into the major groove, retaining the guanosine amino hydrogen bond.

In the crystal structure of the dodecamer d(CGCGAATTCGCG) containing a molecule of spermine (14) the amine group bridges the phosphate groups across the major groove. Hydrogen bonds are formed to both guanosine and cytidine. We observe an interaction of $d(CG)_3$ with ammonium ion which may be specific for the sequence d(CpG).

MATERIAL AND METHODS

$d(CG)_3$ was synthesized by the phosphate triester method in solution (15).

NMR spectra at 500 MHz were recorded on a Bruker WM-500 spectrometer. The sequence θ_x-t-θ_{-x} (16) was used to suppress the water resonance. The delay between the two pulses was 350 usec. In addition to the minimum at the water resonance this produces a null at ca. 5.6 ppm on either side of the water resonance. The imino protons are excited in the second enveloppe. Thus an appoximate 90° pulse can be applied both over the aromatic protons and the imino protons which facilitates NOE measurements in particular. Spin-spin relaxation rates were calculated either from the half width ($\pi\Delta\nu_{1/2} = T_2^{-1}$) or using the sequence $\pi/4(x)$-t_1-$\pi/4(-x)$-t_2-$\pi/2(x)$-t_1-$\pi/2(-x)$-t_2.

A preirradiation time of 500 msec. was employed. The power level was just sufficient to saturate the resonance.

The oligonucleotide was 3 mM in strand dissolved in 10 mM phosphate buffer (pH 7.2) (unless otherwise stated), 150 mM NaCl and 0.5 mM EDTA in 90% H_2O/ 10% D_2O. Chemical shifts are reported relative to sodium 4,4-dimethyl-4-silapentane-1-sulphonate, DSS).

RESULTS

(1) Sodium form: At 0°C nine resonances of exchangeable protons are observed (Figure 1, peaks labelled a - i). Peak b corresponds to two protons and all the others to single protons. The apparently different integrated intensities arise from the different effective flip angles as a function of frequency (see Material and Methods). Thus of the 15 exchangeable (3 imino and 12 amino) protons only 10 are observed. The chemical shifts of the nonexchangeable protons are very similar to those in a recent study of $d(CG)_3$ (19).

On raising the temperature progressively to 30°C the imino proton, peak a, and three amino resonances, peaks e, f and i, broaden and then almost disappear from the spectrum. The early broadening and disappearance of the imino resonance, peak a, corresponds to the previously observed fraying of the terminal base pair (17,18) and thus peak b corresponds to the imino protons of the internal base pairs (2,5) and (3,4).

Assignment of the amino protons was made by a series of nuclear Overhauser effect (NOE) experiments at 20°C (Figure 2). The resolution in our spectra is apparently poorer because of a larger exponential multiplication factor introduced before Fourier transformation. At 20°C saturation of the two imino resonances gives large negative NOEs on peaks c and d, and smaller ones on peaks g and h. This indicates that c and d are the closest protons. Saturation of peak g produces a large NOE on peak c and another at 5.477 ppm (Figure 2), but not on the imino resonances. Due to resonance overlap peak h was partially irradiated, giving an NOE on peak d (see below). Similarly, saturation of peak h gives a large NOE on peak d, another one at 5.447 ppm, but nothing on the imino protons. The resonances at 5.477 and 5.447 ppm were assigned to the cytidine H^5 protons of base pairs (2,5) and (3,4), respectively (19). Saturation of the imino protons gives much larger NOEs on peaks c and d than on peaks g and h. Saturation of peak c or d gives only a very small (less than 5%) NOE on the cytidine H^5 resonances. Thus we can assign peaks c and g to the hydrogen bonded

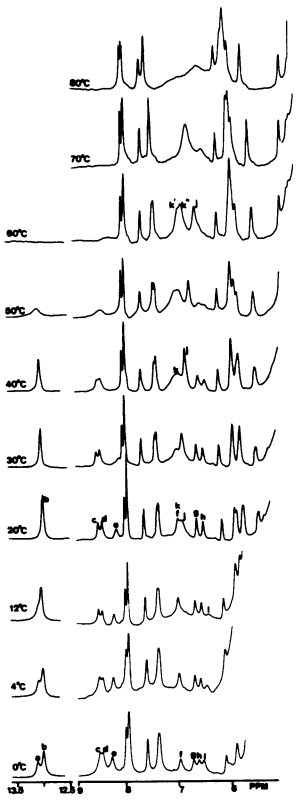

Figure 1: 500 MHz proton NMR spectrum of $d(CG)_3$, sodium salt, at pH 7.2 in 90% H_2O/10% D_2O at various temperatures.

and non hydrogen bonded cytidine amino protons of base pair (2,5), respectively. Similarly, peaks d and h correspond to those of base pair (3,4). Saturation of peak f gives a large NOE on peak e and a smaller one at 5.86 ppm, the H^5 resonance of cytidine of the terminal base pair (1,6). Correspondingly, saturation of peak e gives a large NOE on peak f and a very small one at 5.86 ppm. Thus peaks e and f are the hydrogen bonded and non hydrogen bonded cytidine amino protons of base pair (1,6), respectively. At 20°C the only resonance remaining unassigned is peak j which integrates for two protons.

Raising the temperature from 20°C to 60°C results in the progressive broadening and then disappearance of first the resonances assigned to the terminal cytidine amino group, followed by the imino protons and those of the cytidine amino protons of the interior base pairs (Figure 1). At the same time peak j becomes narrower and a new resonance, peak K, appears which integrates for four protons and is poorly resolved into two

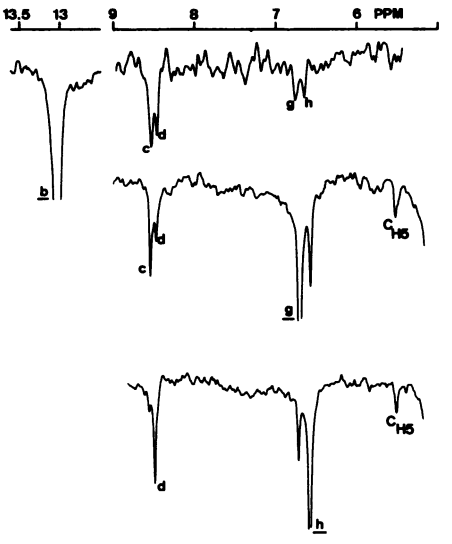

Figure 2: Negative NOE's observed at 20°C after preirradiation for 0.5 sec. The irradiated resonance is underlined. Due to resonance overlap partial irradiation of adjacent resonance is observed for experiments on peaks g and h (see text).

resonances at 50 - 60°C. Peaks j and k correspond to six protons in total and these must therefore be the guanosine amino protons. Peak j is tentatively assigned to the protons of the terminal base pair (see below). Above 60°C, first peak j and subsequently peak k broaden and disappear completely at 85°C. Below 20°C peak j broadens and disappears, while peak i appears. This latter peak corresponds to a single proton and must be a guanosine amino proton, the other five of which are missing.

(2) Ammonium form. When the ammonium form of $d(CG)_3$ is dissolved in the same solvent (10 mM phosphate, 0.15 M NaCl), the spectra show some important differences relative to the sodium form: at 0°C three very large resonances (peaks l, m, n) are observed (Figure 3). Intergration of the region between 7.0 and 7.4 ppm relative to the terminal H^6 of cytidine gives 19 ± 0.5 protons. The two internal cytidine H^6 protons must lie in this region (they are resolved at higher temperatures, Figure 3) and the remaining intensity must arise from exchangeable protons. As spectra were recorded in 90% H_2O, each exchangeable proton will be reduced by 10% in its intensity by deuteration. Thus this region contains 18.7 ± 0.5 exchangeable protons, one of which will be proton f of Figure 1. The

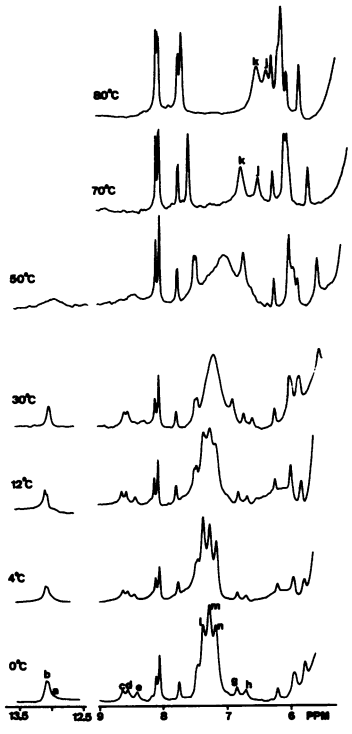

Figure 3: 500 MHz protons NMR spectrum of $d(CG)_3$, ammonium salt, at pH 7.2 in 90% H_2O/10% D_2O at various temperatures.

separation between the adjacent resonances in the enveloppe l, m, n is 52 Hz, very close to the coupling constant for the ammonium ion. The protons of the free ammonium ion, however, exchange rapidly with water at pH 7.2, they are only observed at pH<6. It should be noted that peak i (Figure 1) is absent in the spectra of the ammonium form (Figure 3). The only other differences observed in the spectrum at 0°C relative to that of the sodium form are minor chemical shift differences, except that the terminal imino resonance a is observed to high field of the internal imino protons, peak b. The assignments were made by NOE measurements in the same way as for the sodium form.

DISCUSSION

Cytidine amino and guanosine imino protons, sodium form: At 0° C all these protons are observed in the spectrum. On raising the temperature to 10°C, the imino proton of the terminal base pair exchanges rapidly with solvent, whereas the relaxation rate of the corresponding amino protons decreases. Base catalysis would seem to be more efficient for the imino

protons in this temperature range than for the amino protons. This assumes that transient base pair opening is a prerequisite for exchange to occur at all. It is, however, impossible to exclude a mechanism whereby the catalyst interacts with the accessible base pairs and proton exchange of the imino protons occurs without opening of the base pair. The longer rates observed for the amino protons relative to those of dCMP above 10°C suggest that exchange without base pair opening is an inefficient pathway for these protons.

The rates of the four cytidine amino and the two imino protons of the internal base pairs as a function of temperature are almost the same, indicating a common rate limiting pathway for all these protons. Therefore, the non hydrogen bonded amino protons - although much more accessible to solvent - cannot exchange independently of the hydrogen bonded protons. The absence of any influence of base catalyst concentration above 10 mM phosphate is consistent with an open limited pathway for exchange under the conditions of our experiments.

Guanosine amino protons, sodium form: In the range 50-60°C we observe three resonances k', k" and j, each of which integrates for two protons. The two protons of each guanosine amino group are magnetically equivalent which could be either accidental or due to rapid rotation of the amino group on an NMR time scale. Rapid rotation of the C^2-amino bond has been observed on the monomer level (20-23) even at 0°C and we find similarly only a single amino resonance for dGTP.

At 60°C base pair opening is a relatively frequent event as shown by the almost total disappearance of all the other exchangeable protons and this would facilitate rotation. Only above 60° C does solvent exchange have an influence upon the relaxation rates and this contrasts very strongly with dGTP which shows a line width of ca. 150 Hz at 0°C and is unobservable at 25°C (22,23). Thus the catalytic mechanism for exchange of the imino protons is inefficient for the guanosine amino protons.

Below 50°C, first peak k and then j broaden and disappear. The exchange mechanism which induces this broadening cannot be due to interaction with solvent but must be an intramolecular process. Below 8°C, we observe a peak i, a single proton, corresponding to one of the guanosine amino groups. We have not been successful in finding the other proton of this amino group which may still be too broad for detection. However, the observation of a single proton demonstrates that for this amino group rotation about the C^2-amino bond is now slow on an NMR time scale.

Thus the broadening of the amino resonances below 50°C is attributed to a progressive slowing down of rotation as has been proposed for C-G base pairs in a tetranucleotide (24). The coalescence temperature for the amino group for which peak i corresponds to one of the protons is ca. 18°C. Since this amino group has the slowest rate of rotation, it should be found as the last one to appear as the temperature is raised. Due to resonance overlap we cannot differentiate between peaks k′ and k", but we can exclude peak j, for in the range 8 to 20° both peaks j and i are observed. We can predict that the proton geminal to that giving peak i would have a chemical shift of ca. 7.4 ppm which would put it under the two overlapping cytidine H^6 protons. The lifetime, t, in each state can be estimated for this amino group at 18°C from $t = \sqrt{2}/(2\pi(\nu_A-\nu_B))$ and is found to be about 0.5 msec. At high temperature the relaxation of the imino and cytidine amino protons is dominated by exchange and is open limited. Extrapolating back to 18°C gives an exchange rate of ca. 50 msec.. While such a procedure cannot be regarded as very accurate, there is approximately a difference of two orders of magnitude between the predicted rate of base pair opening and exchange between these two amino protons. For the two other base pairs the rate of intramolecular amino proton exchange is even faster. Peak i belongs to an amino group of one of the internal base pairs and most probably to base pair (3,4). We would expect the terminal guanosine group to exchange with solvent more rapidly than the internal amino groups. At 70°C, peak j is much more broadened by exchange than peak k, thus peak j can tentatively be assigned to base pair (1,6). Coalescence of peak i gives either peak k′ or k". The observation that at 0°C only one amino group is in slow intramolecular exchange argues against a mechanism of the protons flipping without any distortion of the base pair, for in that case the rate should be the same for all the amino groups. The slowing down of exchange is probably due to greater helix rigidity for this base pair which would indicate it as belonging to base pair (3,4). With the available data we cannot determine the rate limiting step for this exchange. As dGTP shows rapid rotation of its amino group on an NMR time scale, then hydrogen bond formation quite reasonably slows it down. In-plane twisting of the bases relative to each other with consequent shortening of one amino hydrogen bond and lengthening of the other one could weaken sufficiently the hydrogen bond to allow rotation without base pair opening.

The behaviour of the guanosine amino protons at high temperature is

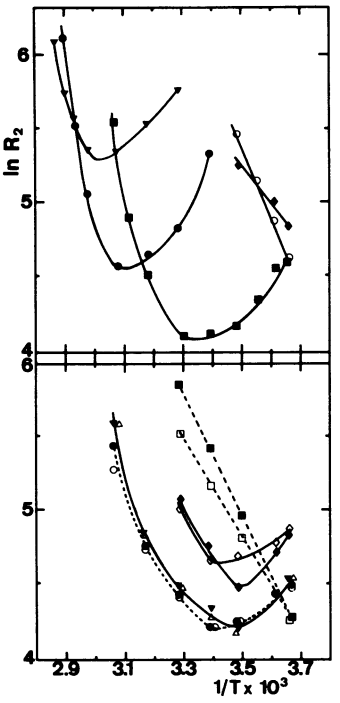

Figure 4: Plot of the natural logarithm of the spin-spin relaxation rates of exchangeable protons of $d(CG)_3$ from line width measurements vs. the reciprocal of the absolute temperature. Top: Guanosine protons: ◆ peak a, imino, base pair 1,6; ■ peak b, imino, base pairs 2,5 and 3,4 (average); o peak i; ● peak j, amino base pair 1,6; ▼ peak k (k′, k″) amino, base pairs 2,5 and 3,4. Bottom: Cytidine protons: ◇ peak f, non hydrogen bonded (nWC) amino, base pair 1,6; ◆ peak e, hydrogen bonded (WC) amino, base pair 1,6; ▼ peak d, WC amino, base pair 3,4; △ peak c, WC amino, base pair 2,5; ● peak h, nWC amino base pair 3,4; o peak g, nWC amino, base pair 2,5. dCMP: ■ upfield resonance; □ downfield resonance (same buffer solution as for $d(CG)_3$.

also somewhat unexpected. That they are still observed at temperatures 50 - 60° higher than for dGTP shows that only close to the Tm does solvent have access to these groups. This, however, has to be reconciled with the observation that the imino proton exchanges much more rapidly. From Figure 4 the rate of exchange with solvent of the guanosine amino protons relative to the cytidine amino protons is a factor of 8 to 10 slower. If hydrogen bonding is the most probable mechanism for retarding the exchange of these protons then the data can be explained by a scissor opening of the G-C pairs into the major groove rather than complete base pair opening. In the scissor action the guanosine amino hydrogen bond is retained, yet solvent and catalyst would have access and promote exchange of the other protons. We do not claim that complete base pair opening does not occur in the temperature range below 60°C, but this scissor movement would account for the major mechanism of imino proton exchange.

Ammonium form. Some chemical shift differences are observed for the ammonium salt relative to the sodium salt. In particular, the terminal imino proton is observed at high field of the internal imino protons

(Figure 3). From ring current considerations it has been proposed (8) that the terminal imino proton should resonate to low field of the imino proton of the second base pair in a 12 base pair DNA fragment beginning with d(CCGC-). While the sequence is not the same we note that the counterion effect in $d(CG)_3$ appears to be of greater importance than ring current considerations.

The major difference observed is the absence of peak i and the appearance of three resonances around 7.2 ppm integrating for 18 protons (peaks l, m and n, after substracting peak f which is assumed to be under this enveloppe) with the typical ammonium ion triplet structure. While free ammonium ion protons exchange rapidly with solvent, the slow exchange observed here must be due to complex formation with DNA. As a function of temperature the cytidine amino and the imino protons show only small differences compared with the sodium salt. On the other hand peaks l, m and n broaden above 4°C and reduce to a six proton spectrum at 70°C. These two resonances have been labelled k and j in Figure 3. Since they must be the protons of the internal and terminal amino groups, respectively and even at 70°C there is little or no exchange broadening. Thus, although the ammonium ion protons now exchange rapidly with solvent, the ammonium ions retard exchange of the amino protons which are still clearly observed at 80°C and only disappear above the Tm (76°C) which is, however, slightly lower than for the sodium form (79°C). The formation of this ammonium ion complex may be specific for the sequence d(CpG), as we have not observed the same phenomenon in d(GGATCC) (25). If complexation occurred interstrand (between the same base pair), we would not expect a sequence effect which we have observed.

Our data do not explain unambiguously the absence of peak i at 0°C. The guanosine amino group may be rotating faster than in the sodium salt or these amino protons may be exchanging with the ammonium protons. We have proposed above that a twist of a base pair about the helix axis is necessary to allow the guanosine amino group to rotate. Proton exchange with the ammonium ion may occur at the same time. The rapid proton exchange may simply be due to the presence of exchangeable protons held in the vicinity, but it may also be possible that in hydrogen bonding to the DNA the ammonium ion catalyzes this exchange by a mechanism which is not clear from our data. The retardation of amino proton - solvent exchange at high temperature suggests that ammonium ion - DNA interaction still occurs at 80°C and may sterically hinder solvent access.

REFERENCES:

1) Feigon, J., Wright, J.M., Denny, W.A., Leupin, W. & Kearns, D.R. (1982) Cold Spring Harbour Symp. Quant. Biol. 47, 207-217
2) Wong, Y.P., Kearns, D.R., Reid, B.R. & Shulman, R.G. (1972) J. Mol. Biol. 72, 725-740
3) Clore, G.M. & Gronenborn, A.M. (1984) Eur. J. Biochem. 141, 119-129
4) Patel, D.J. (1976) Biopolymers 15, 533-558
5) Teitelbaum, H. & Englander, S.W. (1975) J. Mol. Biol. 92, 55-78
6) Teitelbaum, H. & Englander, S.W. (1975) J. Mol. Biol. 92, 79-92
7) Mandal, C., Kallenbach, N.R. & Englander, S.W. (1979) J. Mol. Biol. 135, 391-411
8) Early, T.A., Kearns, D.R., Hillen, W. & Wells, R.D. (1981) Biochemistry 20, 3756-3764
9) Early, T.A., Kearns, D.R., Hillen, W. & Wells, R.D. (1981) Biochemistry 20, 3764-3769
10) Pardi, A. & Tinoco, Jr.,I. (1982) Biochemistry 20, 4686-4693
11) Pardi, A., Morden, K.M., Patel, D.J. & Tinoco, Jr.,I. (1982) Biochemistry 21, 6567-6574
12) Feigon, J., Denny, W.A., Leupin, W. & Kearns, D.R. (1983) Biochemistry 22, 5930-5942
13) Mirau, P.A. & Kearns, D.R. (1984) J. Mol. Biol. in press
14) Kopka, M.L., Fratini, A.V., Drew, H.R. & Dickerson, R.E. (1983) J. Mol. Biol. 163, 129-146
15) van der Marel, G., van Boeckel, C.A.A., Wille, G. & van Boom, J.H. (1981) Tetrahedron Lett. 22, 3887-3890
16) Plateau, P. & Guéron, M. (1982) J.Amer. Chem. Soc. 104, 7310-7311
17) Patel, D.J. & Hilbers, C.W. (1975) Biochemistry 14, 2651-2660
18) Kan, L.-S., Borer, P.N. & Ts'o, P.O.P. (1975) Biochemistry 14, 4864-
19) Fréchet, D., Cheng, D.M., Kan, L.-S. & Ts'o, P.O.P. (1983) Biochemistry 22, 5194-5200
20) McConnell, B. & Seavell, P.C. (1972) Biochemistry 11, 4382-4392
21) McConnell, B. & Seawell, P.C. (1973) Biochemistry 12, 4426-4434
22) Raszka, M. (1974) Biochemistry 13, 4616-1622
23) Raszka, M. & Kaplan, N.O. (1972) Proc. Natl. Acad. Sc. U.S. 69, 2025-2029
24) Patel, D.J.(1977) Biopolymers 16, 1635-1656
25) Fazakerley, G.V., Guy, A., Téoule, R. & Guschlbauer, W. (1984) unpublished results
26) McConnell, B. (1984) J. Biomol. Structure Dynamics 1, 1407-1421

Note: After termination of this manuscript our attention was drawn to an article (26) in which the amino protons of $d(CG)_2$ were assigned by comparison with the exchange rates of monomers with solvent. This method could, however, not be used to identify protons from a specific base pair in a longer DNA fragment.