Volume 331, number 3, 217–222 FEBS 13026 October 1993

Determination of distances involving exchangeable protons from NOESY spectra of two DNA dodecamers

Ryszard Michalczyk, Irina M. Russu*

Department of Molecular Biology and Biochemistry, Wesleyan University, Middletown, CT 06459, USA

Received 8 July 1993

Two-dimensional proton nuclear Overhauser effect (NOESY) spectra were obtained as a function of mixing time for two DNA dodecamers, 5′-CGCGAATTCGCG-3′ and 5′-CGCAAATTTGCG-3′, in aqueous solutions. The time evolution of cross-peak volumes was quantitatively analyzed based on the approximate solution of the relaxation/exchange equation over a range of mixing times from 30 to 150 ms. Inter-proton distances, involving exchangeable protons in key locations of the structures, were calculated and compared to the distances predicted by the crystal structures. These NMR-derived distances enhance the number of constraints, and their accuracy, for determination of solution structures of the two DNA dodecamers.

2D NMR; NOE; DNA oligonucleotide; Inter-proton distance constraint; Exchangeable proton

1. INTRODUCTION

Nuclear magnetic resonance (NMR) spectroscopy is the method of choice for determination of three-dimensional structure of nucleic acid oligonucleotides in the solution state. The method relies upon 2D nuclear Overhauser effect spectroscopy (NOESY) to obtain inter-proton distances, and these are subsequently used as constraints in distance geometry or molecular dynamics algorithms to generate the structures (for a review see [1]). Inter-proton distances are calculated from a two-spin approximation or from a full relaxation rate matrix analysis of NOESY cross-peak volumes [1–6]. Distance information is often supplemented by restraints on torsion angles obtained from 2D correlation spectra [1].

The majority of inter-proton distances used in structure determination of nucleic acids has involved non-exchangeable protons of bases and sugars. The use of exchangeable protons has been limited by technical complexities associated with their quantitative observation and by their exchange with solvent protons. In some cases, NOESY cross-peaks of exchangeable protons at a single mixing time, usually longer than 100 ms, have been used to obtain approximate lower and upper bounds for the distances [7–13]. In other cases, cross-peaks of DNA exchangeable protons have been treated qualitatively, by dividing them into weak, medium and strong categories, in a manner similar to that used in proteins [14].

*Corresponding author. Fax: (1) (203) 344- 7952.

Abbreviations: NMR, nuclear magnetic resonance; 1D, one-dimensional; 2D, two-dimensional; NOESY, nuclear Overhauser effect spectroscopy.

The resolution of nucleic acid NMR structures can be improved by using a greater number of constraints and by increasing the accuracy of these constraints. Distance constraints on exchangeable protons are of special interest due to the key locations of these protons in double-helical DNA molecules. In the present work, we have used NOESY spectra, and their time-dependence, to obtain reliable estimates for inter-proton distances involving exchangeable protons in two DNA dodecamers in aqueous solution. The DNA dodecamers chosen (Fig. 1) have been previously investigated by X-ray crystallography and by NMR in deuterated solvents [15–19]. Imino proton exchange and base-pair opening in these dodecamers have also been characterized by this laboratory [20,21].

2. MATERIALS AND METHODS

The DNA dodecamers were synthesized using the solid-support phosphoramidite method and purified using reverse-phase HPLC on a Hamilton PRP-1 semi-preparative column. Samples were dialyzed against 10 mM phosphate buffer, pH 7.0, containing 100 mM NaCl and 2 mM EDTA. The DNA concentration (duplex) was 3.9 mM for A_2T_2 dodecamer and 1.9 mM for A_3T_3 dodecamer.

The NMR experiments were carried out at 5°C on a Varian VXR-400 spectrometer operating at a proton frequency of 400 MHz. The spectra were referenced to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) and the water resonance was at 4.98 ppm. 1D spectra were recorded using the jump-and-return pulse sequence [22]. The excitation maxima were set at 7.80 ppm and 13.45 ppm thus allowing for nearly optimal observation of both the imino and amino protons in the same experiment. Non-selective proton spin-lattice relaxation times (T_1) were measured by the inversion–recovery method. Selective T_1's were measured by the saturation–recovery method in which each resonance was saturated individually. NOESY spectra were obtained using the regular pulse sequence:

[Preparation, P_1 – Evolution (t_1), P_2 – Mixing (τ_m), P_3 – Detection (t_2)]

Each of the pulses, P_1, P_2 and P_3, was replaced by a jump-and-return pulse [23], and phase cycling was as described by States et al. [24]. Seven NOESY spectra were obtained for mixing times, τ_m ranging from 30 to 150 ms. The preparation delay was 2.06 s which allows for at least 80% recovery of magnetization of all protons (under the same experimental conditions, the non-selective proton T_1 values in the two dodecamers range from 0.4 s to 1.3 s). 512 complex points were acquired with a sweep width of 9,000 Hz and 32 scans for each t_1 value. The data were apodized with a gaussian (15.5 Hz) and a 90°-shifted sine-bell function (12.5 Hz) and transformed as a 4K × 4K matrix.

The 2D excitation profile for the NOESY pulse sequence was determined on a sample of 99.8% D_2O. The transmitter offset was incremented in 0.3 ppm steps away from the frequency of the residual HDO resonance. The excitation of the HDO resonance was found to be within 5% of the values predicted by

$$A(\omega_1,\omega_2) = \sin^2(\omega_1 \tau) \cdot \sin(\omega_2 \tau) \quad (1)$$

where ω_1 and ω_2 denote the precession frequencies relative to water signal during t_1 and t_2, respectively [23].

Cross-peak volumes were measured by integration of the traces in both dimensions as described [25]. For each symmetrically related pair of cross-peaks, volumes were measured separately on each side of the diagonal, corrected for excitation according to eqn. 1 and then averaged. The errors in cross-peak volume measurements were estimated at ± 15%. Cross-peak volumes were normalized using the volume of a single proton diagonal peak, namely H2 of adenine in position 6 (resonances at 7.61 ppm in the A_2T_2 dodecamer, and at 7.49 ppm in the A_3T_3 dodecamer). The intensity of the diagonal peak was measured as a function of mixing time, and the resulting relaxation curve was extrapolated to a mixing time of zero to obtain the intensity used for normalization [26].

NOESY build-up curves were analyzed based on the equation:

$$\frac{dA(\tau_m)}{d\tau_m} = -L \cdot A(\tau_m) \quad (2)$$

$A(\tau_m)$ is the matrix of diagonal and cross-peak volumes and $L = R - K$ where R is the symmetrical relaxation rate matrix and K is the kinetic matrix describing the exchange between DNA exchangeable protons and water protons [27]. Solution of eqn. 2 for cross-peak volumes was approximated in the range of mixing times used as [27]:

$$A_{ij}(\tau_m) = -L_{ij}\tau_m + \frac{1}{2}\sum_k L_{ik} L_{kj} \tau_m^2 - \frac{1}{6}\sum_k\sum_p L_{ik} L_{kp} L_{pj} \tau_m^3 \ldots \quad (3)$$

Experimental cross-peak volumes were fitted to eqn. 3 using a non-linear least-squares algorithm. For most cross-peaks, the NOESY build-up curves were described satisfactorily by the second-order polynomial in τ_m. In some cases (such as cross-peaks between cytosine amino protons in the same base and cross-peaks between imino protons in successive base pairs), the third-order polynomial gave a significantly better fit of the data.

Atomic coordinates for the crystal structures of A_2T_2 [28] and A_3T_3 [17] dodecamers were obtained from the Nucleic Acid Database [29]. Proton coordinates were generated using the program Quanta on an Iris Silicon Graphics workstation.

3. RESULTS

The NOESY spectra and resonance assignments are illustrated in Fig. 2 for the A_3T_3 dodecamer. All cross-peaks between thymine H3 imino and adenine H2 protons, and between guanine H1 imino and cytosine H4 amino protons are observed. The cross-peaks between thymine H3 imino and adenine H6 amino protons are broad or poorly resolved except for the cross-peak between the imino and non-hydrogen-bonded amino proton in base pair A6 · T7 (f_1 = 5.98 ppm and f_2 = 13.73 ppm). Similarly, the cross-peaks between guanine H1 imino and H2 amino protons are not observed because, at 5°C, the resonances of these amino protons are broadened, probably due to rotation about the C–N bond.

Representative examples of NOESY build-up curves for the two dodecamers are shown in Figs. 3–6.

Inter-proton distances were calculated from cross-relaxation rates determined from the term linear in τ_m in eqn. 2. The reference for cross-relaxation rate was that between thymine H6 and CH_3 protons within the same base (Fig. 3). Alternative reference cross-peaks (e.g. cytosine H5–H6 or deoxyribose H2′–H2″) were not chosen for the following reasons. The volumes of most cytosine H5–H6 cross-peaks are too low to be quantitated because H5 protons resonate close to water (5.4– 5.6 ppm) and their excitation in the pulse sequence is low (eqn. 1). For the terminal cytosines, the volumes of the H5–H6 cross-peaks are higher (e.g. cross-peak at 5.81 ppm/7.61 ppm in Fig. 2B) but are also affected by fraying at the ends of the duplex. Deoxyribose H2′–H2″ cross-peaks are not resolved sufficiently well for accurate integration of their volumes. The correlation times for overall DNA motion were calculated at 5°C using rotational diffusion coefficients obtained from the hydrodynamic theory of Tirado and Garcia de la Torre [30], and they are: τ_{II} = 4.4 ns and $\tau_{\perp}$ = 55.2 ns. The effect of internal rotation upon the cross-relaxation rate between thymine H6 and CH_3 protons was included in the analysis using the model developed by Tropp [31] for cylindrical molecules with internal rotation described as three-state jumps. A correlation time of 0.1 ns was assumed for internal rotation of the methyl group. With these values, the observed cross-relaxation rate corresponds to a pseudoatom replacing the methyl group and situated at r_0 = 2.89 Å from H6, with the inter-nuclear vector perpendicular to the helix axis. This distance was used as a reference in calculations of all other distances based on: $\sigma_{ij}/\sigma_0 = (r_0/r_{ij})^{1/6}$, where σ_0 is the reference cross-relaxation rate and σ_{ij} and r_{ij} are the cross-relaxation rate and distance, respectively, of the proton pair of interest. For the latter, the dependence

Fig. 1. Base sequences of the DNA dodecamers used in this study and their abbreviations.

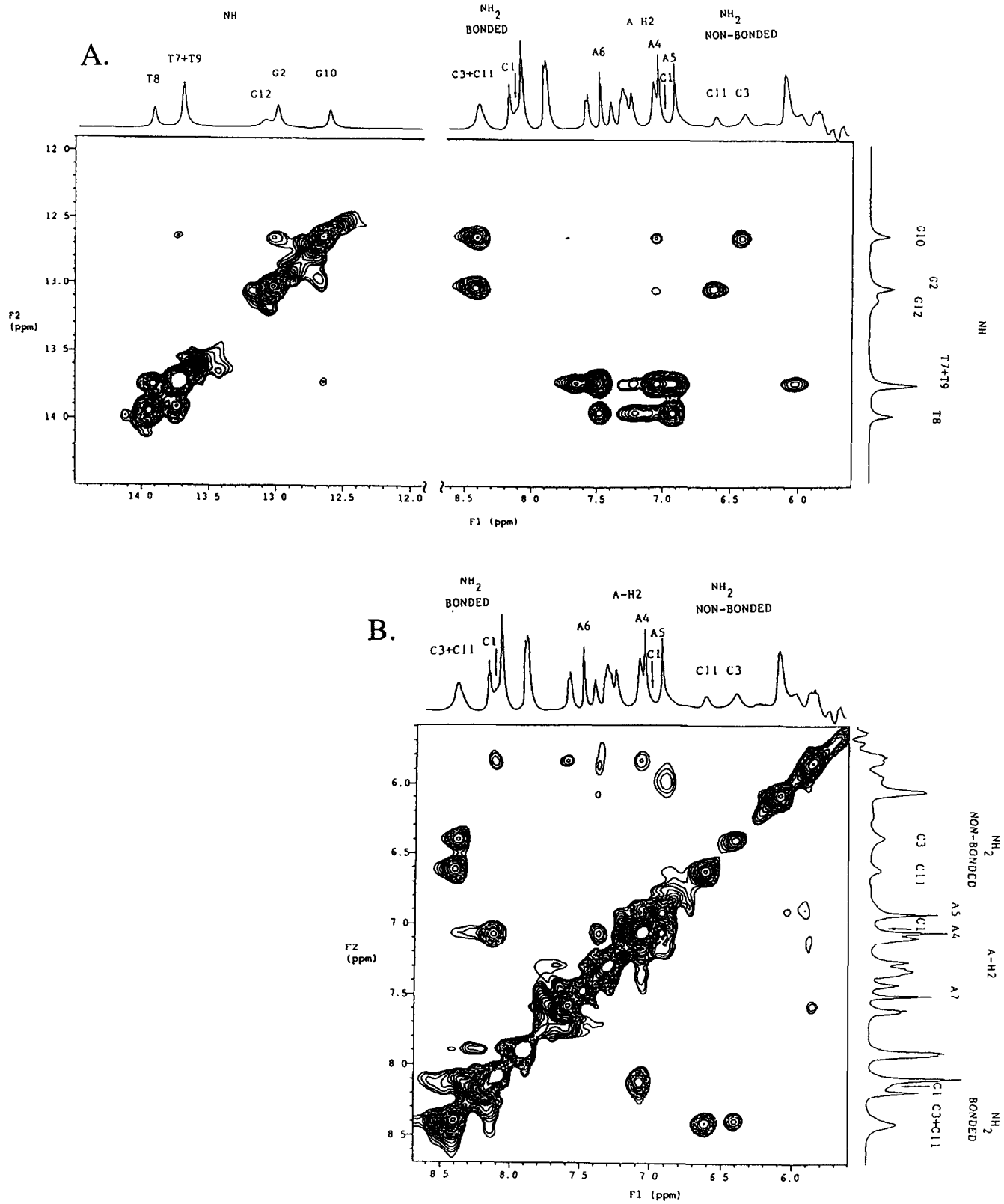

Fig. 2. Expanded regions of the NOESY spectrum of A_3T_3 dodecamer (100 ms mixing time). (A) Cross-peaks involving imino protons; (B) cross-peaks involving amino and non-exchangeable protons.

of the cross-relaxation rate upon the orientation of the inter nuclear vector was included in the calculations using Woessner's model [32]. The angles between the inter-nuclear vector and the helix axis were calculated

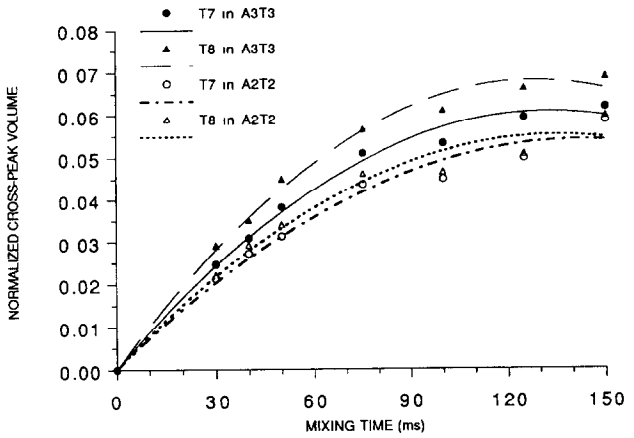

Fig. 3. Dependence of NOESY volumes on mixing time for cross-peaks between thymine H6 and CH_3 protons in A_2T_2 and A_3T_3 dodecamers. The curves represent fits to a second-order polynomial in τ_m (eqn. 3).

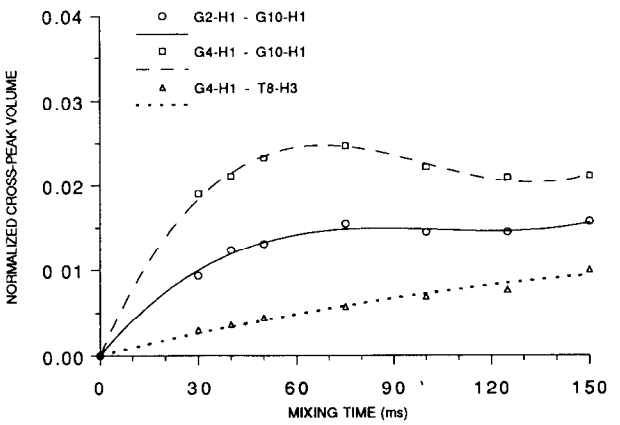

Fig. 5. Dependence of NOESY volumes on mixing time for cross-peaks between imino protons in A_2T_2 dodecamer. The curves represent fits to a third-order polynomial in τ_m (eqn. 3).

based on the crystal structure coordinates. Representative inter-proton distances for the two dodecamers are shown in Tables I and II. The tables also include distances derived from the crystal structures for each oligonucleotide [17,28].

4. DISCUSSION

The results obtained in the present work demonstrate that reliable NOESY build-up curves involving exchangeable protons in DNA oligonucleotides can be obtained by using the NOESY pulse sequence with three jump-and-return pulses [23]. The reliability of the method can be tested by comparing build-up curves for cross-peaks between protons for which the distance is fixed. For example, the build-up curves for thymine H6–CH_3 cross-peaks in base pairs A5 · T8 and A6 · T7 of both dodecamers are almost identical (Fig. 3). Similar results are observed for cross-peaks between hydrogen bonded and non-hydrogen-bonded protons in amino groups of inner cytosines of both dodecamers (Fig. 6).

Evaluation of inter-proton distances from NOESY build-up curves requires solving the full set of differential equations in eqn. 2. One limitation in this approach has been the lack of experimental volumes for overlapped and poorly resolved cross-peaks. In aqueous solutions, the number of NOESY cross-peaks that are unaccessible experimentally is larger than in deuterated solvents due to the large solvent resonance, the pulse scheme used for its supression, and the exchange-broadening of some resonances. Moreover, under these experimental conditions, the matrix, $\boldsymbol{K}$, describing the exchange between DNA protons and water protons, must be known. Again, this matrix is incomplete since, for protons that are exchange-broadened, exchange rates cannot be measured (for example, amino protons in the guanines and in some of the adenines in the two dode-

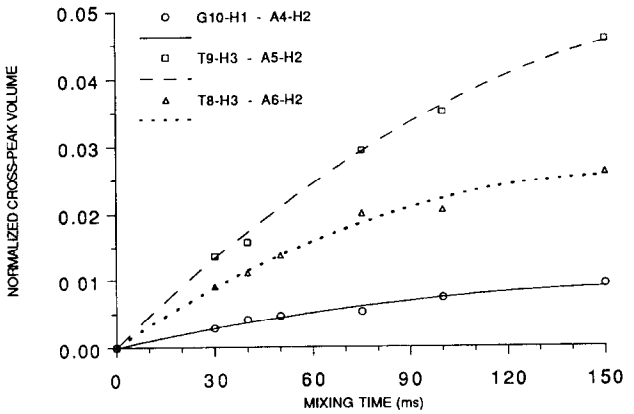

Fig. 4. Dependence of NOESY volumes on mixing time for cross-peaks between imino and adenine H2 protons in successive base pairs in A_3T_3 dodecamer. The curves represent fits to a second-order polynomial in τ_m (eqn. 3).

Table I

Comparison of selected NMR-derived distances (Å) involving exchangeable protons with distances predicted by the crystal structure for A_2T_2 dodecamer

Proton pair	NMR	Crystal
G2-H1–G10-H1	3.2	3.5
G4-H1–G10-H1	2.5	3.1
G4-H1–T8-H3	3.7	3.9
G10-H1–C3-H4A	2.4	2.7
G4-H1–C9-H4A	2.4	2.7
C3-H4A–C3-H4B	1.8	1.8
C9-H4A–C9-H4B	1.8	1.8
T7-H3–A6-H2	2.4	2.8
T8-H3–A5-H2	2.4	2.7
T8-H3–A6-H2	2.9	3.4
G4-H1–A5-H2	3.4	3.7

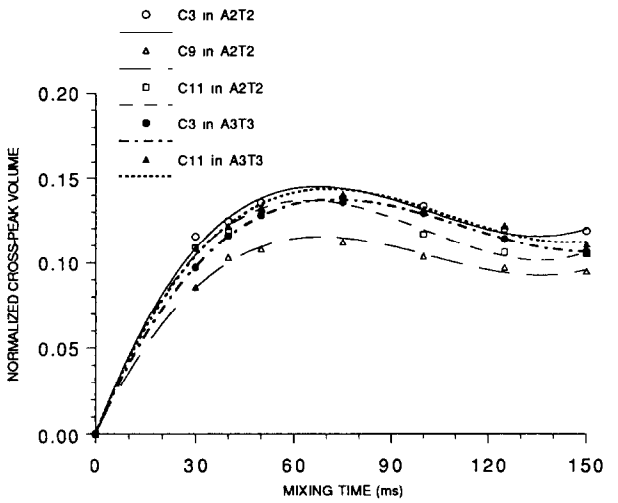

Fig. 6. Dependence of NOESY volumes on mixing time for cross-peaks between hydrogen-bonded (H4A) and non-hydrogen-bonded (H4B) amino protons in inner cytosines in A_2T_2 and A_3T_3 dodecamers. The curves represent fits to a third-order polynomial in τ_m (eqn. 3).

camers studied here). For these reasons, in the present work, we have opted for a range of mixing times in which the approximated solution (eqn. 3) for multi-spin relaxation and exchange can be used (Figs. 3–6). The coefficient, L_{ij}, in the term linear in τ_m in eqn. 3 gives directly the cross-relaxation rate and is not affected by exchange with water. Exchange with water affects, however, the coefficients of the terms of second- and higher-order in τ_m. For example, it can be readily shown that the coefficient of the term quadratic in τ_m depends on exchange as follows:

$$\sum_k L_{ik} L_{kj} = \sum_k R'_{ik} R'_{kj} + k_{wi} \cdot k_{jw}$$

where R'_{ij} is related to the elements of the relaxation rate matrix by: $R'_{ij} = R_{ij}$ for off-diagonal and $R'_{ii} = R_{ii} + k_{iw}$ for diagonal elements. The rate constants for exchange of a DNA proton, i, with water protons are defined as:

$$i \underset{k_{wi}}{\overset{k_{iw}}{\rightleftarrows}} W$$

At equilibrium: $f_i \cdot k_{iw} = f_w \cdot k_{wi}$, where f_i and f_w are the molar fractions of spin i and water spins, respectively. Since $f_w \cong 1$, $k_{wi} = f_i \cdot k_{iw}$ (e.g. $k_{wi} = k_{iw} \cdot 10^{-5}$ for a 1.1 mM DNA solution). Hence, the term $k_{wi} \cdot k_{jw}$ is negligible and exchange affects only the spin-lattice relaxation rates (R'_{ii}) of the protons of interest. Larger spin-lattice relaxation rates should clearly decrease the observed NOE's.

In the present work, we have chosen experimental conditions that should minimize the exchange between DNA exchangeable protons and water protons, namely, lower temperature and a buffer (i.e. phosphate) which is an inefficient catalyst for imino proton exchange. The exchange rates were measured from NOESY experiments in which only the observation pulse (P_3) was a jump-and-return pulse. They were calculated, as described [33], from the volumes of exchange cross-peaks between DNA and water protons and the selective spin-lattice relaxation rates of DNA exchangeable protons at 5°C (2.9–4.5 s^{-1}). The exchange rates were found to range from 0.1 to 0.3 s^{-1}. These values are 10% or less of the spin-lattice relaxation rates (R'_{ii}). Hence, the contribution of exchange to the term quadratic in τ_m in eqn. 3 should be small and should affect to a minimal extent the NOESY build-up curves over the range of mixing times employed here.

Comparison of distances obtained from NMR to those derived from crystal structures suggests that the overall structures of both dodecamers are similar in solution and in the crystal (Tables I and II). From a total of 80 distances determined in the present work, 55% of them are within ± 10% of those predicted by crystal structures. Differences larger than 0.6 Å are observed for approximately 20% of the distances. The detailed comparison of inter-proton distances awaits determination of the complete solution structures.

Inter-proton distances obtained from NOESY spectra in water provide additional accurate constraints for the refinement of the structures of DNA oligonucleotides in solution state. They complement similar measurements in deuterated solvents and, when used together, these distance constraints should improve significantly the determination of solution structure of DNA. We are currently using this approach to determine the structures of the two dodecamers discussed in this paper.

Table II

Comparison of selected NMR-derived distances (Å) involving exchangeable protons with distances predicted by the crystal structure for A_3T_3 dodecamer

Proton pair	NMR	Crystal
G2-H1–G10-H1	2.8	3.7
T9-H3–G10-H1	3.6	3.5
G2-H1–C11-H4A	2.5	2.8
G10-H1–C3-H4A	2.5	2.5
C3-H4A–C3-H4B	1.9	1.8
T7-H3–A6-H2	2.5	2.8
T8-H3–A5-H2	2.6	3.1
T9-H3–A4-H2	2.6	3.1
T8-H3–A6-H2	3.1	3.2
T9-H3–A5-H2	2.9	3.1
G10-H1–A4-H2	3.7	3.9

Acknowledgements: This work was supported by a grant from the National Science Foundation (88-17589).

REFERENCES

[1] Clore, G.M. and Gronenborn, A. (1989) CRC Crit. Rev. Biochem. Mol. Biol. 24, 479–564.
[2] Borgias, B.A. and James, T.L. (1988) J. Magn. Reson. 79, 493–512.
[3] Borgias, B.A. and James, T.L. (1990) J. Magn. Reson. 87, 475–487.
[4] Boelens, R., Koning, T.M.G., Van der Marel, G.A., Van Boom, J.H. and Kaptein, R. (1989) J. Magn. Reson. 82, 290–308.
[5] Post, C.P., Meadows, R.P. and Gorenstein, D.G. (1990) J. Am. Chem. Soc. 112, 6796–6803.
[6] Kaluarachchi, K., Meadows, R.P. and Gorenstein, D.G. (1991) Biochemistry 30, 8785–8797.
[7] Clore, G.M., Oschkinat, H., McLaughlin, L.W., Benseler, F., Happ, C.S., Happ, E. and Gronenborn, A.M. (1988) Biochemistry 27, 4185–4197.
[8] Katahira, M., Sugeta, H., Kyogoku, Y. and Fujii, S. (1990) Biochemistry 29, 7214–7222.
[9] Zhang, X. and Patel, D.J. (1990) Biochemistry 29, 9451–9466.
[10] Gopalakrishnan, S., Liu, X. and Patel, D.J. (1992) Biochemistry 31, 10790–10801.
[11] Cosman, M., Santos, de los C., Fiala, R., Hingerty, B.E., Ibanez, V., Luna, E., Harvey, R., Geacintov, N.E., Broyde, S. and Patel, D.J. (1993) Biochemistry 32, 4145–4155.
[12] Weisz, K., Shafer, R.H., Egan, W. and James, T.L. (1992) Biochemistry 31, 7477–7487.
[13] Mujeeb, A., Kerwin, S.M., Egan, W., Kenyon, G.L. and James, T.L. (1992) Biochemistry 31, 9325–9338.
[14] Banks, K.M., Hare, D.R. and Reid, B.R. (1989) Biochemistry 28, 6996–7010.
[15] Dickerson, R.E. and Drew, H.R. (1981) J. Mol. Biol. 149, 761–786.
[16] Coll, M., Frederick, C.A., Wang, A.H.-J. and Rich, A. (1987) Proc. Natl. Acad. Sci. USA 84, 8385–8389.
[17] Edwards, K.J., Brown, D.G., Spink, N., Skelly, J.V. and Neidle, S. (1992) J. Mol. Biol. 226, 1161–1173.
[18] Nerdal, W., Hare, D.R. and Reid, B.R. (1989) Biochemistry 28, 10008–10021.
[19] Lane, A.N., Jenkins, T.C., Brown, T. and Neidle, S. (1991) Biochemistry 30, 1372–1385.
[20] Moe, J.G. and Russu, I.M. (1990) Nucleic Acids Res. 18, 821–827.
[21] Moe, J.G. and Russu, I.M. (1992) Biochemistry 31, 8421–8428.
[22] Plateau, P. and Gueron, M. (1982) J. Am. Chem. Soc. 104, 7310–7311.
[23] Otting, G., Grutter, R., Leupin, W., Minganti, C., Ganesh, K.N., Sproat, B.S., Gait, M.J. and Wuthrich, K. (1987) Eur. J. Biochem. 166, 215–220.
[24] States, D.J., Haberkorn, R.A. and Ruben, D.J. (1982) J. Magn. Reson. 48, 286–292.
[25] Chary, K.V.R., Hosur, R.V., Govil, G., Chen, C.-q. and Miles, H.T. (1988) Biochemistry 27, 3858–3867.
[26] Mirau, P.A. (1988) J. Magn. Reson. 80, 439–447.
[27] Ernst, R.R., Bodenhausen, G. and Wokaun, A. (1987) in: Principles of Nuclear Magnetic Resonance in One and Two Dimensions, pp. 490–501, Clarendon Press, Oxford.
[28] Holbrook, S.R., Dickerson, R.E. and Kim, S.H. (1985) Acta Crystalogr. Ser. B. 41, 255–262.
[29] Berman, H.M., Olson, W.K., Beveridge, D.L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.-H., Srinivasan, A.R. and Schneider, B. (1992) Biophys. J. 63, 751–759.
[30] Tirado, M.M. and Torre, de la J.G. (1980) J. Chem. Phys. 73, 1986–1993.
[31] Tropp, J. (1980) J. Chem. Phys. 72, 6035–6043.
[32] Woessner, D.E. (1962) J. Chem. Phys. 37, 647–654.
[33] Dobson, C.M., Lian, L.-Y., Redfield, C. and Topping, K.D. (1986) J. Magn. Reson. 69, 201–209.