2D NOE FULL RELAXATION MATRIX ANALYSIS AND MOLECULAR MECHANICS CALCULATIONS FOR STRUCTURE DETERMINATION IN SOLUTION: DNA FRAGMENTS

T.L. JAMES, E.-I. SUZUKI, N. PATTABIRAMAN and G. ZON

Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143 (USA)

INTRODUCTION

The DNA double helix proposed by Watson and Crick (1) has the following general features: DNA is double-stranded with a) antiparallel sugar-phosphate chains and b) Watson-Crick G-C and A-T base pairs. However, during the last decade, experimental and theoretical studies have shown that there are structural variations for DNA, such as left-handed B DNA (2) and Z DNA (3) as well as more subtle sequence-dependent variations. There have even been proposals for Hoogsteen and reversed Watson-Crick base-pairs (4). Ionic interactions play a dominant role in stabilizing as well as transition between structures such as B DNA and Z DNA etc. (5). Here, we present experimental and theoretical 2D NOE spectra and molecular mechanics calculations on the self-complementary decamer $[d\text{-}(5'ATATATATAT3')]_2$. The results from the combination of theoretical and experimental techniques suggest that there are interchain sugar-phosphate interactions in the minor groove for this right-handed decamer duplex, different from Z DNA, which serves to form a hydration tunnel in the minor groove. There has been some controversy about the solution structure of poly(dA-dT)·poly(dA-dT) with limited NMR data being cited for it as left-handed B DNA (6) and other NMR data as regular B DNA (7). Under varying conditions, structures for the alternating d-(AT) polymer from X-ray diffraction have included A DNA (8), B DNA (8), alternating B DNA (9), C DNA (10), D DNA (11) and wrinkled D DNA (12, referred to as WD DNA).

THEORY

The theory for 2D NOE was developed in Ernst's lab (13). Subsequently, we (14) described a matrix method relating all interproton distances in a molecule to peak intensities in a theoretical 2D NOE spectrum which can be iteratively compared with the experimental pure absorption 2D NOE spectrum. The 2D NOE pulse sequence $[(\pi/2) - t_1 - (\pi/2) - \tau_m - (\pi/2) - t_2 -]$ yields a 2D spectrum following Fourier transforms over t_2 and t_1. We have used a modification of the phase cycling route of States et al. (15) to obtain pure absorption phase spectra.

The intensity of an auto- (diagonal) or cross-peak for any value of the experimental parameter τ_m, the mixing time, may be calculated from

$$\mathbf{a}(\tau_m) = \mathbf{\chi} \exp(-\mathbf{\lambda}\tau_m)\ \mathbf{\chi}^{-1} \qquad [1]$$

where element a_{ij} of matrix **a** gives the 2D NOE cross-peak intensity for nuclei i and j, $\mathbf{\chi}$ is the matrix of eigenvectors of the relaxation rate matrix **R**, and $\mathbf{\lambda}$ is a diagonal matrix of eigenvalues, i.e., the solution to the system of equations comprising the rate matrix. Relaxation rate matrix **R** has diagonal elements

$$R_{ii} = \sum_{j} (W_o^{ij} + 2W_1^{ij} + W_2^{ij}) + R_{1i} \quad [2]$$

representing the direct relaxation contributions from all sources to spin i, and off-diagonal elements

$$R_{ij} = W_2^{ij} - W_o^{ij} \quad [3]$$

representing the cross-relaxation rate between spins i and j. R_{1i} comprises all sources of relaxation other than proton-proton dipolar interactions. The zero-, single-, and double-quantum transition probabilities W_o^{ij}, W_1^{ij}, and W_2^{ij} are proportional to $J(\omega)/r_{ij}^6$ where $J(\omega)$ is the spectral density for the molecular motion modulating the ij interaction and r_{ij} is the distance between i and j.

A pure absorption 2D NOE spectrum will depend on choice of the mixing time τ_m used, molecular motions as manifest in the spectral densities, and internuclear distances between all protons, not just the two giving rise to a cross-peak. Macura and Ernst (13) showed that a pair of spins i and j can be considered as isolated if τ_M is sufficiently short, thereby enabling use of only a couple terms in a series expansion. In practice it is difficult to isolate spins i and j completely from spin k if $r_{ik} \lesssim r_{ij}$. Although an ij cross-peak can still be observed and used qualitatively, the peak intensity will be modified, thus preventing quantitative internuclear distance determinations.

Our theoretical investigation (14) and rigid molecule test study (16) of the potential for 2D NOE spectral determinations of time-averaged internuclear distances lead to the following conclusions: (a) several pure absorption 2D NOE spectra at a series of mixing times should be obtained; (b) a practical upper limit on determination of distance is ~ 5Å; (c) distances should be determined to an accuracy of ~ 10% if a good fit to all data (non-overlapping peaks assumed) is obtained; and (d) isotropic motion with a single effective correlation time can be assumed for 10% distance accuracy even though the actual motion may be more complicated.

METHODS

The self-complementary d-(AT) decamer was synthesized as described previously (17). All ^{1}H NMR spectra were obtained at 500MHz on the GE/Nicolet GN500 spectrometer. Pure absorption phase proton 2D NOE spectra in deuterium oxide solution [7 mM d-(AT) decamer, 180 mM NaCl, 100 mM phosphate buffer, pH 7.0] at 15°C were acquired using the (delay time - 90° -t_1 - 90° - τ_m - 90° - t_2)$_n$ pulse sequence with alternate block accumulation and phase cycling to eliminate artifacts arising from pulse imperfections and single and double quantum coherences (15). Mixing times τ_m used were 50, 100, 250, and 400 ms, and the delay time was 12 sec. A small (10%) variation in τ_m was used to minimize zero quantum coherence.

RESULTS AND DISCUSSION

Our intent is to obtain detailed structures in solution for DNA fragments via internuclear distances from the 2D NOE spectra augmented by torsion angle information from J coupling constants. To accomplish this in detail, it is necessary to resolve most 2D NMR cross-peaks and to assign most proton resonances.

Resonance assignments

Figure 1 contains the proton pure absorption 2D NOE spectrum (500 MHz) of [d-(ATATATATAT)]$_2$ in D_2O solution. As previously described (17-24), a sequential assignment procedure can be used to identify the non-exchangeable proton resonances assuming that the DNA duplex structure is a right-handed helix. Any problems in this assumption are immediately apparent. Most of the assignments are quite secure; as most nuclei have several cross-peaks to other

nuclei, there are cross-checks on the assignments. The 100 proton resonances of the d-(AT) decamer duplex have been assigned, with some ambiguity remaining for about 10 resonances (24).

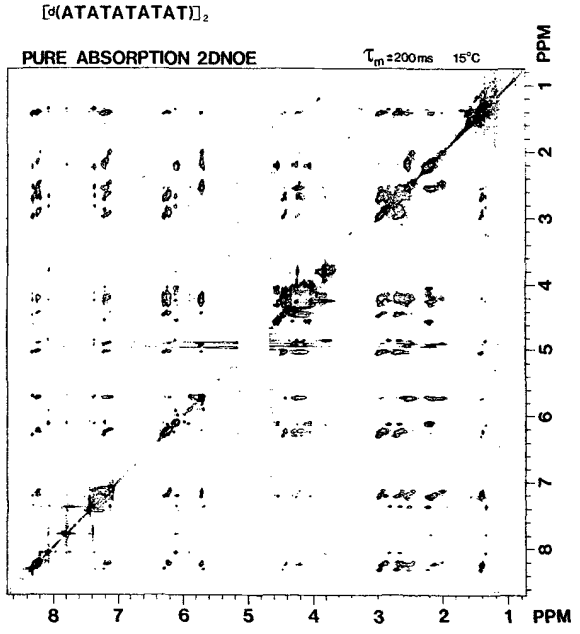

Fig. 1. 500 MHz ^{1}H 2D NOE spectrum (pure absorption phase) of [d-(5'ATATATATAT3')]$_2$. Sample conditions: 15°C; 7 mM, 180 mM NaCl, 100 mM phosphate buffer, pH 7.0 in deuterium oxide. The HDO signal was used as chemical shift reference 4.89 ppm downfield of sodium 3-trimethylsilyl-(2,2,3,3-$^2H_4$) proprionate.

Analysis of 2D NOE spectra

The experimental 2D NOE spectra at four mixing times were compared with theoretical spectra calculated with the complete relaxation analysis method (14) using X-ray diffraction-determined atomic coordinates of Z, A, B, alternating B, left-handed B, C, D, and wrinkled D forms of DNA. Although not required for the relaxation matrix approach, isotropic motion of the decamer and dominance of proton-proton dipolar relaxation were assumed here to simplify the model calculations. From non-selective T_1 and Hahn spin-echo T_2 relaxation time measurements of the relatively well-resolved A(9)-H2 proton (3.6 s and 0.011 s, respectively) and TH1' proton (2.3 s and 0.0085 s, respectively) isotropic correlation time τ_c values of 7.0 ns and 6.0 ns were calculated. A value of 7 ns was used for τ_c in the 2D NOE calculations; A-H1' diagonal peak decay curves were calculated for several DNA structures. All of the calculated curves effectively fit the experimental observations. The value of 7 ns as the isotropic correlation time was used in subsequent calculations of the 2D NOE spectra reported here, although a limited number of calculations using 3 ns exhibited only slight variations from the 7 ns spectra and did not influence the choice of molecular model which best fits the experimental data.

From a detailed comparison of the theoretical and experimental 2D NOE spectra, it was recognized that Z, A, C, left-handed B, and alternating B models can be clearly excluded, that the WD form is the best one, and that regular B and D form models are less likely possibilities. A comparison between the B and WD structures is shown in Figure 2 as stereodiagrams.

Energy-minimized structures

Structures of B DNA, D DNA and WD DNA were displayed using the program MIDAS (25) on an Evans & Sutherland PS2 color graphics terminal. From the display, it appeared that there is a possible interchain sugar-phosphate interaction in the minor groove of WD DNA. In addition there are hydrophobic patches due to CH_2-2' and CH_2-5' moieties of interchain sugars forming a zig-zag pattern from one strand to the opposing one along the minor groove, and hydrophilic patches are formed by interchain phosphate groups. It is interesting to note that the minor groove space between the sugar-phosphate chains and bases is inaccessible to

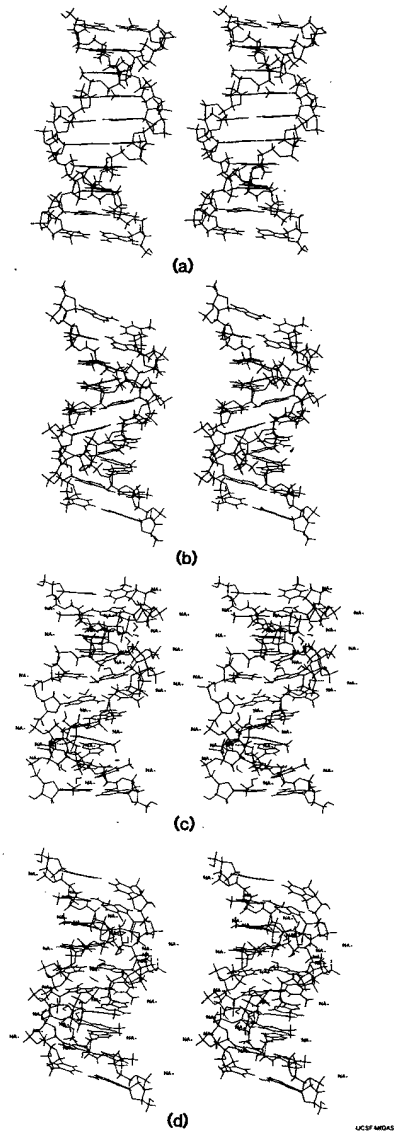

Fig. 2. Stereodiagrams for B and WD models: (a) original B, (b) original WD, (c) energy-minimized B, (d) energy-minimized WD models. Figures (c) and (d) include sodium ion and water molecules.

solvent molecules from either the minor or major groove side; in effect, a tunnel is formed. The size of the tunnel is such that one can place water molecules bridging between N3 atoms of adenines on complementary strands in neighboring base-pairs and also between O2 atoms of thymines in neighboring base-pairs. Apart from the two water molecules bridging the complementary strands, there is sufficient space to place a third water molecule bridging the first two water molecules. With these three types of hydration water molecules, the entire tunnel can be filled completely with a good match of the molecular surface of DNA inside the tunnel to the water molecules' volume. The minor groove is too shallow in B DNA for such interstrand water bridging, although there is an earlier report on a hydration backbone exposed to bulk water (26). The minor groove in D DNA is too small, and its inner cavity is not continuous. Hence, only WD DNA has an ideal hydration backbone in a tunnel.

To understand the relative stability of B and WD DNA, right-handed double helical structures for the d-(AT) decamer in regular B and wrinkled D forms were generated from the monomer or dimer coordinates obtained by X-ray fiber-diffraction studies (8,12). Nine water molecules were placed between adenine and thymine bases and eight water molecules to bridge these nine water molecules as described above. Sodium ions were distributed by iteratively neutralizing the electrostatic potential around the decamer as described previously (27). Molecular mechanics calculations were carried out on the double helical decamer using the program AMBER, Assisted Model Building with Energy Refinement (28). The calculations indicated that WD DNA is more stable than B DNA for alternating AT DNA. Even though the intramolecular energy of DNA and counterion-counterion interactions favor the B DNA structure, inclusion of counterion-DNA interactions, water-water interactions and water-DNA interactions

favor WD DNA. The tunnel formation and interchain sugar-sugar and sugar-phosphate interactions also favor the WD form for structures with alternating d-(AT) sequences.

The 2D NOE spectra calculated for the energy-minimized WD-DNA are almost the same as for the original structure, reflecting the small changes upon energy refinement. In contrast, energy minimization of B DNA and D DNA caused a more substantial change in structure yielding structures and, consequently, theoretical 2D NOE spectra similar to that of energy-minimized WD DNA, which, in turn, was similar to the initial WD DNA. The theoretical spectra of the energy-minimized forms of B-, D- and WD DNA all fit the experimental spectra quite well.

REFERENCES AND NOTES

1 J.D. Watson and F.H.C. Crick, Nature, 171 (1953) 737.

2 V. Sasisekharan and N. Pattabiraman, Nature, 275 (1978) 159.

3 A.H-J. Wang, S. Fujii, G.J. Quigley, F.J. Kolpak, J.L. Crawford, J.H. van Boom, G.A. van der Marel and A. Rich, Nature, 282 (1979) 680; H. Drew, T. Takano, S. Tanaka, K. Itakura and R.E. Dickerson, Nature, 286 (1980) 567.

4 J. Donohue and K.M. Trueblood, J. Molec. Biol., 2 (1960) 363.

5 F.M. Pohl and T.M. Jovin, J. Molec. Biol., 67 (1972) 375; M. Behe, G. Felsenfeld, Proc. Natl. Acad. Sci., 78 (1981) 1619.

6 G. Gupta, M.H. Sarma, M.M. Dhingra, R.H. Sarma, M. Rajagopalan and V. Sasisekharan, J. Biomol. Struc. Dyn., 1 (1983) 395.

7 N. Assa-Munt and D.R. Kearns, Biochemistry, 23 (1984), 791; B. Borah, J.S. Cohen and A. Bax, Biopolymers, 24 (1985) 747.

8 S. Arnott and D.W.L. Hukins, Biochem. Biophys. Res. Comm., 47 (1972) 1504.

9 A. Klug, A. Jack, M.A. Viswamitra, O. Kennard, Z. Shakked and T.A. Steitz, J. Mol. Biol., 131 (1979) 669.

10 D.A. Marvin, M. Spencer, M.H.F. Wilkins and L.D. Hamilton, J. Mol. Biol., 3 (1961) 547.

11 S. Arnott and E. Selsing, J. Mol. Biol., 88 (1974) 509.

12 S. Arnott, R. Chandrasekaran, L.C. Puigjaner, J.K. Walker, I.H. Hall, D.L. Birdsdall and R.L. Ratliff, Nucl. Acids Res., 11 (1983) 1457.

13 S. Macura and R.R. Ernst, Mol. Phys., 41 (1980) 95.

14 J.W. Keepers and T.L. James, J. Magn. Reson, 57 (1984), 404.

15 D.J. States, R.A. Haberkorn and D.J. Ruben, J. Magn. Reson., 48 (1982) 286.

16 G.B. Young and T.L. James,J. Am. Chem. Soc., 106 (1984) 7986.

17 M.S. Broido, G. Zon and T.L. James, Biochem. Biophys. Res. Commun., 119 (1984) 663.

18 R.M. Scheek, R. Boelens, N. Russo, J.H. van Boom and R. Kaptein, Biochemistry, 23 (1984) 1371.

19 M.S. Broido, T.L. James, G. Zon, J.W. Keepers, Eur. J. Biochem., 150 (1984) 117.

20 N. Jamin, T.L. James and G. Zon, Eur. J. Biochem., 152 (1985) 157.

21 D.E. Wemmer, S.H. Chou, D.R. Hart and B.R. Reid, Nucl. Acids Res., 13 (1985) 3755.

22 C.W. Hilbers, C.A.G. Haasnoot, S.H. de Bruin, J.J.M. Joordens, G.A. van der Marel and J.H. van Boom, Biochimie, 67 (1985) 685.

23 J.A. Cailles, J.-M. Neumann, S. Tran-Dinh, T. Huyuh-Dinh, B.L. d'Estaintot and J. Igolen, Eur. J. Biochem., 147 (1985) 183.

24 E. Suzuki, N. Pattabiraman, G. Zon and T.L. James, Biochemistry (1986) in press.

25 L. Gallo, C. Huang, T.E. Ferrin, R. Langridge, User's Manual-Molecular Interactive Display and Simulation (MIDAS), Computer Graphics Laboratory, University of California, San Francisco, 1985; R.

Langridge, T.E. Ferrin, I.D. Kuntz and M.L. Connolly, Science, 211 (1981) 661.

26 H.R. Drew and R.E. Dickerson, J. Mol. Biol., 151 (1981) 5335.

27 N. Pattabiraman, R. Langridge and P.A. Kollman, J. Biomol. Struc. Dyn., 1 (1984) 1525; S. Devarajan and N. Pattabiraman, a FORTRAN program to calculate iteratively ionic positions around DNA, 1986.

28 P.K. Weiner and P.A. Kollman, J. Comp. Chem., 2 (1981) 287; S. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona and P.K. Weiner, J. Am. Chem. Soc., 106 (1984) 765.