NMR Relaxation Matrix Refinement of a DNA Octamer Solution Structure

Lawrence Shapiro,[a] Michael Nilges[b] and Magdalena Eriksson[c,]*

[a] Department of Biochemistry and Molecular Biophysics, College of Physicians and Surgeons, Columbia University, New York, NY 10032 USA, [b] EMBL Heidelberg, D-6900 Heidelberg, Federal Republic of Germany and [c] Department of Medical Biochemistry and Biophysics, University of Umeå, S-90187 Umeå, Sweden

The data used in a typical NMR structure determination generally comprise a set of interproton distances estimated based on an approximation of the nuclear Overhauser effect as a two-spin phenomenon. The nuclear Overhauser effect is, however, a multi-spin phenomenon. Herein we show that this isolated spin pair approximation can lead to very distorted sets of distances for NMR structure determinations of DNA. Strong spin diffusion pathways, which we have visualized by three-dimensional nuclear Overhauser spectroscopy, correlate with systematic underestimation of interproton distances in a model DNA oligomer. These difficulties can be substantially alleviated by abandoning the isolated spin pair approximation in favor of the more robust full relaxation matrix approach.

The nuclear Overhauser effect is a spin–spin relaxation phenomenon which arises from the interaction of an ensemble of spins.[1] However, for practical purposes, it has often proved necessary to approximate the nuclear Overhauser effect as an interaction of only two spins. In this two-spin approximation or isolated spin pair approximation (ISPA), interproton distances r_{ij} are simply related to the build-up rates, η_{ij}, of NOE crosspeaks with respect to mixing time, eqn. (1). As a practical matter, this quantity

$$\frac{r_{ij}^6}{r_{kl}^6} = \frac{\eta_{kl}}{\eta_{ij}} \tag{1}$$

is quite simple to measure. NOESY spectra are acquired at several mixing times; the intensities of the crosspeaks are measured at each of these mixing times τ_m, and η_{ij} is approximated as the slope of the increase of intensity with respect to τ_m. η_{kl} and r_{kl} correspond to the build-up rate and known interproton distance of a reference peak. In DNA this is typically a resolved cytidine H5–H6 crosspeak with a fixed separation of 2.45 Å. A number of studies have incorporated the NMR-based distance constraints defined by lower and upper bounds with either distance geometry[2–4] or molecular dynamics[5–8] in attempts to define structural features of DNA and RNA oligomers [reviewed in Refs. 9 and 10].

The NOE build-up rates are generally non-linear as a function of mixing time in the NOESY experiment. Spin diffusion effects in macromolecules[11] limit the use of distance information by the two-spin approximation as a satisfactory means of structure determination. These

difficulties have been pointed out specifically for the case of DNA by a number of authors (see, e.g., the review in Ref. 12). To alleviate these difficulties and analyze correctly the NOE data one may utilize an approach based on the NMR relaxation matrix[13] which takes into account contributions from indirect magnetization transfer. Such full relaxation matrix calculations[14] and related analysis of the Bloch equations[15] have been invoked to calculate the two-dimensional NOE intensities of the structure being refined and minimization of the difference between observed and calculated intensities. The early approaches focused on hand-wise refinement procedures where, after each cycle of refinement, the constraining distances were altered manually after inspection of observed and calculated NOESY spectra.[16–19]

Recently, relaxation matrix refinement schemes have been introduced which incorporate the difference between the observed and calculated two-dimensional NOE intensities directly into the molecular modeling potential. These methods include an angle space least-squares minimization technique, COMATOSE[20] and an iterative relaxation matrix technique, IRMA.[21] IRMA is a hybrid of distance and relaxation matrix approaches in which experimental NOEs are combined with calculated NOEs to back-calculate a relaxation matrix. Distance estimates are obtained from the relaxation matrix, and these distances are used in restrained molecular dynamics calculations. As the structure evolves in dynamics, the updated relaxation matrix yields new distances to drive the dynamics in the appropriate direction. The iteration continues until the experimental NOE intensities are satisfied. This technique has been applied to determine structural features of several DNA oligomers in solution.[21–28] A similar refinement procedure, MARDIGRAS, which

---

refines the relaxation matrix alone, largely independently of the model, has been developed by James and coworkers.[22-24]

More recently an NOE-based refinement technique has been outlined in which the comparison between observed and calculated NOE intensities is used as an energy potential in energy-minimization calculations.[29] The 'driving force' here is the difference $I_{obs} - I_{calc}$; there is no intermediate step of translation of NOE intensities into distances. To drive an energy minimization calculation by this potential, one must calculate the force generated by the potential on each coordinate of each hydrogen atom under refinement. The force generated by the calculated intensity I^c of the ijth crosspeak on the coordinates of an atom μ is given by the gradient of the relaxation pseudoenergy term given in eqn. (2). This derivative must

$$F_\mu = \nabla_\mu I^c_{ij} \tag{2}$$

be calculable to implement a molecular dynamics or minimization refinement scheme based on $I_{obs} - I_{calc}$. This gradient was obtained numerically through a finite difference calculation[29] and this approach has been used to define the solution conformation of a series of DNA oligomer duplexes using relaxation matrix-restrained minimization.[30, 31] The finite difference approximation utilized by Baleja, *et al.*[29-31] was too computationally intensive to allow the use of more powerful molecular dynamics refinement techniques.

In a recent paper, Yip and Case[32] presented the gradient of the mixing coefficients matrix in closed analytical form. This work has enabled us to calculate the forces generated by the relaxation pseudoenergy required in the implementation of molecular dynamics calculations in a more computationally tractable form. Nilges *et al.*[33] introduced an algebraically equivalent simplification of this expression, which further reduces the computational requirements. In this study, we have applied these algorithms to the study of a DNA octamer.

The self-complementary d(G-G-T-A-T-A-C-C) duplex has previously been investigated by X-ray and NMR techniques. The X-ray structure was solved to 1.8 Å resolution with d(G-G-T-A-T-A-C-C) adopting an A-DNA-like helix in the crystalline state.[34] Sequence-dependent local conformational variations were detected in this structure solved at atomic resolution. The NMR spectral characteristics of the d(G-G-T-A-T-A-C-C) duplex in solution have been analyzed by one-dimensional[35] and two-dimensional[36-37] through-bond and through-space experiments. An attempt was made to define the solution conformation of this octanucleotide duplex based on a comparison of the observed NOESY spectrum with calculated NOESY spectra based on a complete relaxation matrix analysis for a limited range of DNA conformations.[38] The authors concluded that the best fit between observed and calculated NOESY plots corresponds to the conformation where the central

d(T-A-T-A) segment adopts a wrinkled D-DNA structure.

The present study has used complete relaxation matrix potentials and forces incorporated in the molecular dynamics program X-PLOR.[33, 39] We clearly see the effects of spin diffusion through comparison of distance-refined and relaxation matrix-refined structures. Significant structural changes take place during relaxation matrix-restrained molecular dynamics. The average conformational parameters of the deoxyoctanucleotide including helical twist, glycosidic torsion angle, and sugar pucker, all move nearer to B-DNA values as a result of the iterative relaxation matrix refinement. These results illustrate the importance of relaxation matrix based refinement techniques and the limitations of the isolated spin pair approximation in the determination of DNA structure from NOE-based NMR data.

Theory

Cross relaxation during the mixing time of a two-dimensional NOE experiment can be described by eqn. (3)

$$\frac{\partial M}{\partial t} = - RM \tag{3}$$

where M is the magnetization vector and R is the relaxation matrix which includes all proton–proton relaxation terms. This equation has the solution shown in eqn. (4).

$$M(\tau_m) = A(\tau_m)\, M(0) = \exp(-R\tau_m)\, M(0) \tag{4}$$

Here $M(0)$ is the magnetization vector at the start of the mixing time which, assuming complete relaxation during the repetition delay, is easily calculated. The matrix A contains the mixing coefficients which are proportional to the two-dimensional NOE intensities. Simple direct analytical computation of A from this equation is precluded by the lack of a general means of matrix exponentiation. This matrix can, however, be represented as a Taylor series, eqn. (5). This equation has provided

$$A(\tau_m) = \exp(-R\tau_m) = 1 - R\tau_m + \frac{1}{2}R^2\tau_m^2 - \frac{1}{6}R^3\tau_m^3 \cdots$$

$$+ \frac{(-1)^i}{i!}R^i\tau_m^i + \cdots \tag{5}$$

the primary impetus for the isolated spin pair approximation. The amplitude of an individual 2D-NOE crosspeak, a_{ij}, assuming constant interproton distances and isotropic molecular tumbling, can be calculated according to eqn. (5) as eqn. (6)

$$a_{ij} = -\sigma_{ij}\tau_m + \frac{1}{2}\sum_k (\sigma_{ik}\sigma_{kj})\tau_m^3 \cdots \tag{6}$$

where

$$\sigma_{ij} = \frac{\hbar^2\gamma^4}{10r_{ij}^6}\left(\frac{6\tau_c}{1 + 4\omega^2\tau_c^2} - \tau_c\right) \tag{7}$$

Here, $\hbar$ is Planck's constant divided by 2π, γ is the proton magnetogyric ratio, ω is the Larmor frequency, and τ_c is the correlation time of molecular tumbling. If we make the severe approximation of considering only the first term of this sum, we find that a_{ij} is proportional to r_{ij}^{-6}, which is the isolated spin pair approximation.

The first term in the sum of eqn. (6) represents 'direct' NOE relaxation, whereas all subsequent terms represent 'spin diffusion' effects. It is clear (because the next largest term is positive) that distance calculation by the ISPA method will systematically tend to underestimate distances. Furthermore, it is apparent that spin diffusion effects will be less pronounced at lower mixing times, as the spin diffusion terms are scaled by powers of τ_m.

Spin diffusion effects in macromolecules are often so severe as to preclude the use of distance estimation by the two-spin approximation as a satisfactory means of structure determination. To obviate these difficulties, several investigators have undertaken methods which involve calculation of the two-dimensional NOE intensities of the structure being refined and iterative minimization of the difference between observed and calculated intensities. These intensities can be found exactly by solution for the mixing time coefficients matrix, A, of eqn. (4). The relaxation matrix R is symmetric, and hence diagonable. Eqn. (5) collapses for diagonal matrices R, allowing simple exact solution for A.

The relaxation matrix R is a function of spectral densities and dipolar coupling strengths. The off-diagonal elements of R are given by eqn. (8) and the diagonal elements are given by eqn. (9).

$$R_{ij} = d_{ij}[6J(2\omega) - J(0)] \tag{8}$$

$$R_{ii} = \sum_{i \neq j} d_{ij}[J(0) + 3J(\omega) + 6J(2\omega)] \tag{9}$$

The dipolar coupling strengths d_{ij} depend on interproton distance, eqn. (10). Our calculations are based on

$$d_{ij} = \frac{\gamma^4 \hbar^2}{10 r_{ij}^6} \tag{10}$$

a model which assumes that the interproton distances r_{ij} remain constant. We have used an analogous approximation in calculating the spectral densities. The spectral density function $J(\omega)$ can be calculated as a Fourier transform of the motional correlation function. We have used a single overall isotropic correlation time to describe molecular motions. In this limiting approximation, the spectral densities are the same for each crosspeak, and are given by eqn. (11).

$$J(\omega) = \frac{\tau_c}{1 + \omega^2 \tau_c^2} \tag{11}$$

A recent study has shown that this approach can underestimate distances involving methyl protons, for which motional effects should be most pronounced, by as much

as 0.3 Å.[40] In the current study we have included methyl groups as sets of three magnetically equivalent protons. Thus, spin diffusion effects contributed by the methyl groups are included in the calculation of intensities of all other crosspeaks. We have not, however, attempted to fit the volumes of crosspeaks involving methyl groups.

Materials and methods

Sample preparation. The d(G–G–T–A–T–A–C–C) duplex was synthesized by solution-phase phosphotriester chemistry as reported previously.[38] The NMR sample was made by dissolving 200 A_{260} units of the octamer in 0.4 ml of 10 mM phosphate buffer, pD 7.4, containing 0.1 M NaCl and 1 mM ethylenediaminetetraacetic acid (EDTA).

NMR data collection and processing. All two dimensional phase-sensitive NMR spectra were collected on a Bruker AM 500 spectrometer. Two-dimensional phase-sensitive NOESY spectra of the duplex in D_2O solution were recorded at 25°C with a repetition delay of 2.5 s for mixing time values 30, 60, 90, 120, 200 and 300 ms. The carrier was placed on the HDO resonance which was suppressed by a low power presaturation, and 1024 complex data points were acquired in the t_2 dimension with a spectral width of 4000 Hz. For each NOESY experiment, real and imaginary data were collected for 512 t_1 values using the method of States, et al.[41] The data in the t_1 dimension were apodized with a 90° phase-shifted sine-bell. This window function was used to avoid truncation effects in t_1 without distorting the relative intensities of resonances having different line widths. After apodization, t_1 data were Fourier-transformed and a small phase correction was applied.

The phase-sensitive three-dimensional NOESY–NOESY spectrum was acquired on a Bruker AM 400 spectrometer by the method of Boelens et al.[42] For this experiment, 256 points were collected for each FID in the t_3 (real time) dimension. 128 points were collected in each of the t_1 and t_2 dimensions. Time-proportional phase incrementation was applied independently for the t_1 and t_2 domains. Both mixing times, τ_{m1} and τ_{m2} were set equal to 120 ms. The relaxation delay was set to 2.0 s, and a spectral width of 4000 Hz was acquired in each dimension. The 3D NOESY–NOESY spectrum was Fourier-transformed using the FELIX NMR processing software (Hare Research, Inc., Woodinville, WA). Each domain was apodized with a 90° sine-bell function. Both the t_1 and t_2 domains were zero-filled to 256 points prior to apodization.

Volume integration and distance estimation. Volume integrals for each crosspeak at each of the six mixing times were obtained using the FTNMR software package (Hare Research, Inc., Woodinville, WA). We used a subroutine which first fits an ellipse to the shape of the two-dimensional crosspeak, and then calculates the volume within

that ellipse by summing the intensities at all contained points.

To obtain a set of distance estimates, these two-dimensional NOE volumes were plotted on a computer against a mixing time axis. Intensities from longer mixing times which clearly, from visual inspection, diverged from the linearity of the initial build-up rates of these plots were deleted. A first-order polynomial line fit was then performed on the remaining peaks, and the slope of the fitted line was used as η_{ij} in eqn. (1) to calculate the distances r_{ij}.

Dynamics target function. All molecular dynamics calculations were carried out using the program X-PLOR.[39] The empirical part of the target function is based on the potential developed by Nilsson and Karplus[5] which has been used in a number of restrained molecular dynamics studies of DNA and RNA.[6-8, 43]

The total potential energy of the target function can be described by the sum of the empirical and experimental portions of the force field, eqn. (12).

$$E_{\text{Total}} = E_{\text{Empirical}} + E_{\text{Distance}} + E_{\text{Relaxation}} \tag{12}$$

The empirical and NOE-distance based components of the potential have been described previously.[6-8, 43] The relaxation energy is a function of the difference between observed and calculated NOE intensities, eqn. (13), where

$$E_{\text{Relaxation}} = k_R \sum_{S} \sum_{i=1}^{N} w_i \, deviation(I_i^c, k_s I_i^o, \delta_i)^2 \tag{13}$$

k_R is the energy constant for the relaxation term, I_i^c and I_i^o are the calculated and observed intensities, respectively, δ_i is the experimental error estimate for I_i^o, w_i is a weight factor, N is the number of crosspeaks in a given spectrum at a particular mixing time, k_s is a scale factor relating observed and calculated intensities at a given mixing time, S is the number of spectra with different mixing times. $deviation(a, b, \delta)$ is defined as the absolute value of the difference between a and b where b has an error estimate of δ, eqn. (14).

$$deviation(a, b, \delta) \equiv \begin{cases} (b - \delta) - a & \text{if } a < (b - \delta) \\ 0 & \text{if } (b - \delta) < a < (b + \delta) \\ a - (b + \delta) & \text{if } a > (b + \delta) \end{cases} \tag{14}$$

In this study we have used error estimates for each crosspeak of $\delta = 0.1 \cdot I_i^o$.

As a measure of the fit of the refined structure to the NOE data, we used a generalized R-factor, eqn. (15).

$$R = \frac{\sum_{i=1}^{N} w_i \cdot deviation(I_i^c, k_s I_i^o, \delta_i)}{\sum_{i=1}^{N} w_i (k_s I_i^o)} \tag{15}$$

The value of the scale factor k_s is evaluated for each spectrum as given in eqn. (16) where the sum runs over

$$k_s = \frac{\sum_{i=1}^{N'} I_i^c}{\sum_{i=1}^{N'} I_i^o} \tag{16}$$

all N' crosspeaks which are 'well-determined.' For a peak to be considered 'well-determined' and included in the calculation of the scale, it must be at least three times as large as the corresponding error estimate. The weighting factors were set as shown in eqn. (17), which is,

$$w_i = \frac{1}{\sqrt{I_i^o}} \tag{17}$$

incidentally, the weighting generally used in protein crystallography. The weights are scaled such that the minimum weight is 1. The maximum weight used is determined by the smallest peak which is well determined. For the present data, the maximum weight was 144.

Refinement protocol. A schematic representation of the refinement protocol is depicted in Scheme 1. The NOE-derived distance constraints are slowly introduced during the first 8 ps of dynamics at 400 K. Constraints on the backbone α, β, ϵ, and ζ dihedral angles are slowly introduced here, and are kept constant throughout the remainder of refinement. Bond distances involving hydrogens are kept constant using the SHAKE algorithm.[44] At the point labeled a in Scheme 1, the velocities are rescaled to a temperature of 300 K, and the system is allowed to equilibrate for 12 ps. The time step throughout the entire protocol is 1 fs. During the time period represented by the shaded region, 12–20 ps, coordinates are saved every 0.1 ps. These coordinates are then averaged, and the averaged coordinate set is used to continue the dynamics protocol at point b. The velocities are now rescaled to a temperature of 200 K, and the relaxation potential is slowly
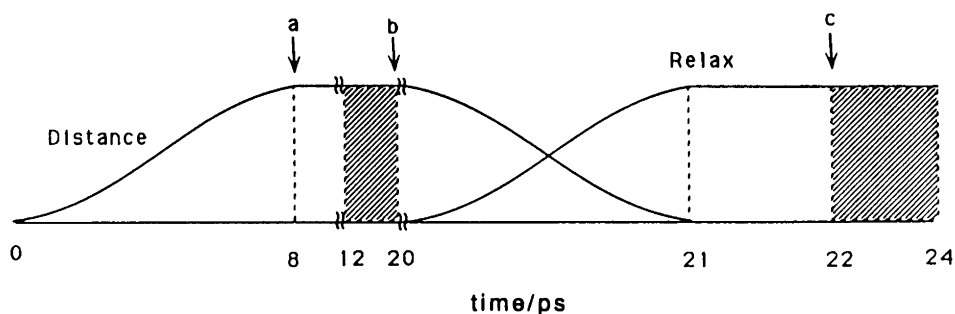
Scheme 1.

introduced over the time period from 20–21 ps. During the time period represented by the shaded region, 22–24 ps, the coordinates are saved every 0.1 ps. These 20 coordinate sets are averaged and subjected to 400 cycles of conjugate gradient minimization resulting in 'final' structures.

The symmetry between the two strands of the octamer was maintained by the addition of a symmetry penalty function to the refinement potential. Each coordinate x is restrained to the symmetry-average value $\langle x \rangle$ computed from the coordinates of both symmetry-related segments, eqn. (18). The weighting factor w_{Sym} was set to 0.1. This

$$E_{Sym} = w_{Sym}(x - \langle x \rangle)^2 \qquad (18)$$

value was maintained throughout the refinement process. Perfect symmetry is not maintained with this w_{Sym} value, however, using significantly higher values unduly restricts motion in dynamics; this value represents an adequate compromise.

The NOE-derived distance potential, $E_{Distance}$ has the form of a harmonic square well, given by eqn. (19)

$$E_{Distance} \equiv \begin{cases} 0 & \text{if} \quad (r_{ij}^{o} - \delta) < r_{ij} < (r_{ij}^{o} + \delta) \\ k_{Distance}(r_{ij} - r_{ij}^{o})^2 & \text{else} \end{cases} \qquad (19)$$

where

$$k_{Distance} = \frac{k_B TS}{2\delta^2} \qquad (20)$$

Here k_B is the Boltzmann constant, T is the absolute temperature, δ is the experimental distance error estimate, and S is a scale factor. The error estimate, δ was set to ± 0.2 Å. We have tethered the Watson–Crick hydrogen bonds with constraints set to ± 0.1 Å of their exact values. Following Happ et al.,[8] we scaled up the distance potential during the first 8 ps of dynamics by increasing the scale factor S from a value of 0.3 to a maximum value of 0.8 by multiplying it by a factor of 1.58 every ps. Similarly, the dihedral angles were included as a harmonic square well potential, eqn. (21). The dihedral angle error

$$E_{Dihedral} \equiv \begin{cases} 0 & \text{if} \quad (\phi_{ij}^{o} - \delta) < \phi_{ij} < (\phi_{ij}^{o} + \delta) \\ k_{Dihedral}(\phi_{ij} - \phi_{ij}^{o})^2 & \text{else} \end{cases} \qquad (21)$$

estimate, δ, was set to 40°, and the constant $k_{Dihedral}$ was scaled from an initial value of 1.0 to a maximum value of 40.0 during the first 8 ps of dynamics by multiplying its value by a factor of 1.58 every 0.05 ps.

The relaxation potential, which is described in eqn. (13) was introduced by increasing the value of k_R from an initial value of 10.0 to a maximum value of 80.0 with successive multiplications by a factor of 1.58 every 0.05 ps during the 20–21 ps time period. The distance constraints, with the exception of the Watson–Crick hydrogen bond tethers, were concomitantly eliminated by successive division by the same factor.

Results

The computer time requirements of NMR relaxation matrix refinement are quite considerable. Each picosecond of dynamics in our calculations, using 1 fs timesteps with six mixing times, takes about 10 h of Convex C220 CPU-time. In coping with this requirement, we have taken the approach of using a set of approximate distances for an initial cycle of dynamics to bring the structure more nearly to agreement with experimental data, and subsequently reducing these distance potentials to zero weight in the refinement while concomitantly introducing the relaxation matrix potential.

Approximate distance measurements. Earlier studies have assigned the crosspeaks in the NOESY spectrum of the d(G–G–T–A–T–A–C–C) duplex at ambient temperature.[36, 37] The NOE build-up data set used in this
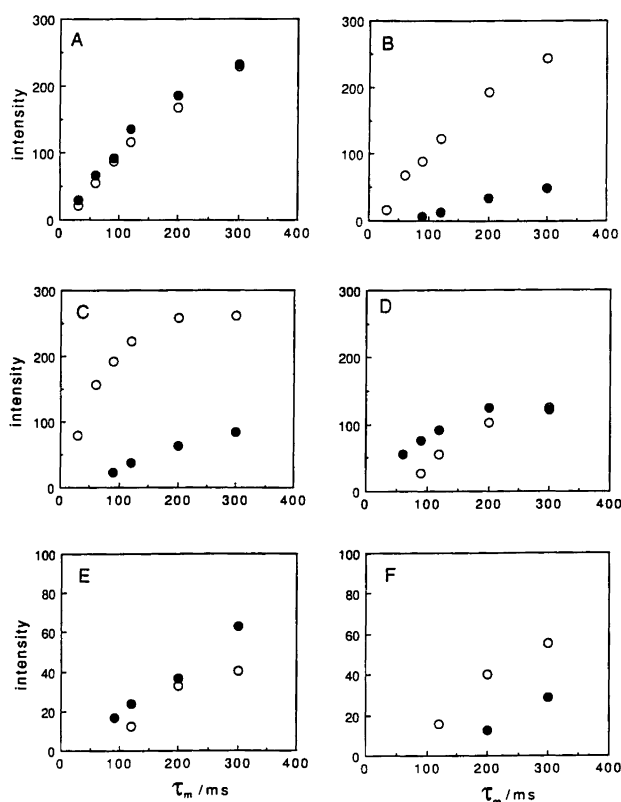

Fig. 1. Build-up curves for the NOE cross peak volume integrals in NOESY spectra of d(G–G–T–A–T–A–C–C) duplex in 0.1 M NaCl, D_2O buffer, 25°C recorded at mixing times of 30, 60, 90, 120, 200 and 300 ms. The build-up for the fixed reference distances are plotted in (A) while build-ups involving protons within residue T3 (open circles) and between residues G2 and T3 (filled circles) are plotted in (B)–(F). The plots are as follows: A, C7(H5)–C7(H6) filled circles and C8(H5)–C5(H6) open circles; B, G2(H8)–T3(CH₃) filled circles and T3(H6)–T3(CH₃) open circles; C, T3(H6)–G2(H2') filled circles and T3(H6)–T3(H2') open circles; D, T3(H6)–G2(H2'') filled circles and T3(H6)–T3(H2'') open circles; E, T3(H6)–G2(H1') filled circles and T3(H6)–T3(H1') open circles; F, T3(H6)–G2(H3') filled circles and T3(H6)–T3(H3') open circles.

47

Table 1. Distance bounds (in Å) between non-exchangeable protons for the d(G–G–T–A–T–A–C–C) duplex at 25°C.

| Residue | Intranucleotide distance constraints | | | | | | |
	H1'–H8/H6	H2'–H8/H6	H2"–H8/H6	H3'–H8/H6	H1'–H2'	H1'–H2"	H1'–H3'
G1	3.6–4.0	2.3–2.7	2.9–3.3	—	2.8–3.2	2.3–2.7	—
G2	3.6–4.0	2.4–2.8	3.1–3.5	3.7–4.1	2.5–2.9	2.2–2.6	3.2–3.6
T3	3.0–3.4	2.1–2.5	2.5–2.9	3.4–3.8	2.9–3.2	2.2–2.6	3.4–3.8
A4	3.2–3.6	2.3–2.7	2.8–3.2	3.7–4.1	2.6–3.0	2.2–2.6	3.3–3.7
T5	3.1–3.5	2.1–2.5	2.8–3.2	3.4–3.8	2.7–3.1	2.2–2.6	3.7–4.1
A6	3.3–3.7	2.3–2.7	2.8–3.2	3.3–3.7	2.6–3.0	2.2–2.6	3.5–3.9
C7	3.2–3.6	2.3–2.7	2.7–3.1	—	2.8–3.2	2.1–2.5	—
C8	3.2–3.6	—	—	—	—	—	—

| Step | Internucleotide distance constraints | | | | |
	H1'–H8/H6	H2'–H8/H6	H2"–H8/H6	H8/H6–H5/CH$_3$	H3'–H8/H6
G1–G2	3.5–3.9	2.9–3.3	2.6–3.0	—	—
G2–T3	3.1–3.5	2.8–3.2	2.6–3.0	3.1–3.5	—
T3–A4	3.1–3.5	2.8–3.2	2.7–3.1	—	4.4–4.8
A4–T5	3.5–3.9	2.8–3.2	2.5–2.9	2.7–3.1	3.9–4.3
T5–A6	3.4–3.8	3.6–4.0	3.2–3.6	—	3.6–4.0
A6–C7	2.9–3.3	3.1–3.5	2.7–3.1	3.6–4.0	—
C7–C8	3.6–4.0	3.0–3.4	2.9–3.3	3.9–4.3	—

study was collected by recording NOESY spectra on the d(G–G–T–A–T–A–C–C) duplex in D$_2$O buffer at 25°C at mixing time values of 30, 60, 90, 120, 200 and 300 ms. The volume integral of each resolved NOE crosspeak was measured as a function of mixing time with typical examples of build-up curves plotted in Fig. 1. These include the build-up between the H6 and H5 protons (fixed 2.45 Å separation) of C7 and C8 (Fig. 1A), between the CH$_3$ proton of T3 and its own and 5'-flanking base protons (Fig. 1B) and between the H6 proton of T3 and its own and 5'-flanking sugar H2' (Fig. 1C), H2" (Fig. 1D), H1' (Fig. 1E) and H3' (Fig. 1F) protons. Intensities from longer mixing times which clearly, from visual inspection, diverged from the linearity of the initial build-up rates of these plots were deleted. A first-order polynomial line fit was then performed on the remaining peaks, and the slope of this line was used as η_{ij} in eqn. (1) to calculate the distance r_{ij}. As a standard r_{kl} was set to 2.45 Å, the cytidine H5–H6 distance, and η_{kl} was set to 1.06, the build-up rate measured for the H5–H6 crosspeak of C7 (Fig. 1A). The H5–H6 crosspeak of residue C8 gave an η_{kl} values of 1.14 (Fig. 1A), and using this value as a standard would have resulted in a difference of less than 2% in the measurement of an unknown distance of about 3 Å. We chose to use only the C7 standard because of concerns of motional fraying at the end base pair. The estimated distances were given lower and upper bounds of −0.2 Å and +0.2 Å, respectively, because of the approximations inherent in the isolated spin pair approximation. These are tight bounds and, in retrospect, the longer distances should have been given larger upper bounds. We obtained a set of 74 non-exchangeable proton distance constraints for one symmetrical half of the d(G–G–T–A–T–A–C–C) duplex and these are listed in Table 1.

Distance-refined structures. The major goal of this work was to characterize the structural differences which arise as a consequence of relaxation matrix refinement as a work-around for the spin diffusion artefacts prevalent in distance-restrained dynamics DNA structures. We have therefore examined several 'distance-refined' structures in order to discover the changes brought about by relaxation matrix refinement. We ran several refinements using the dynamics protocol of Scheme 1, starting, alternatively, from models of the octanucleotide constructed

Table 2.

RMSD (in Å) among structures refined using approximate distance bounds.

	DA2	DB1	DB2
DA1	0.653	0.855	1.676
DA2		0.923	1.681
DB1			1.313

RMSD among relaxation matrix-refined structures.

	RA2	RB1	RB2
RA1	0.629	0.772	1.068
RA2		0.641	0.758
RB1			1.186

RMS coordinate difference resulting from relaxation matrix refinement.

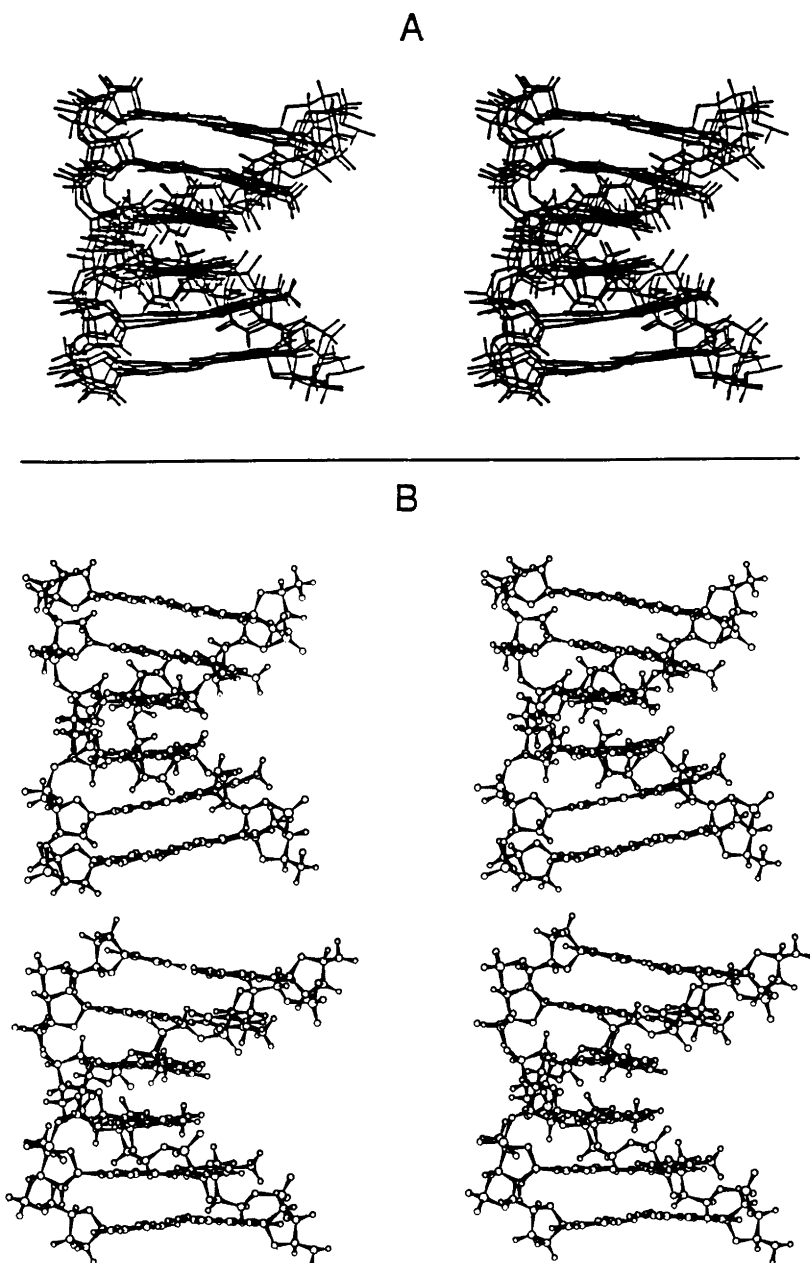| DA1 → RA1	1.656
DA2 → RA2	1.799
DB1 → RB1	1.277
DB2 → RB2	2.144

A

B

Fig. 2. (A) A stereo view of the superposition of four distance-refined structures of the central six base pairs of the d(G–G–T–A–T–A–C–C) duplex. Two structures designated DA1 and DA2 were refined starting from A-DNA while two structures designated DB1 and DB2 were refined starting from B-DNA. The RMS coordinate deviations among these distance-refined structures are listed in Table 2. (B) Stereoviews of the central six base pairs of the d(G–G–T–A–T–A–C–C) duplex. Structure after distance-refinement, DA1, shown above, structure after subsequent relaxation matrix-refinement, RA1, shown below. The RMS coordinate deviations among the structures are listed in Table 2.

from either A-form or B-form DNA. The degree of determination of the structure, using the Nilsson and Karplus[5] empirical potential in addition to the approximate distance bounds, is fairly high. Fig. 2A shows a stereo view of four superimposed 'final' distance-refined structures, i.e., the 12–20 ps dynamics-averaged structures of the molecular dynamics protocol of Scheme 1. Two of the structures (designated DB1 and DB2) were refined starting from canonical B-DNA coordinates, and two

structures (designated DA1 and DA2) were refined from canonical A-DNA coordinates. It is clear from Fig. 2A that the approximate NOE-based distance data, in concert with the other terms of the potential, is sufficient to drive a convergence to one family of similar conformations from structures as disparate as A-form and B-form DNA. The RMS deviations among these structures are listed in Table 2.

Helical parameters for the distance-refined structures

are plotted in Fig. 5 with the boxed region showing the range of values observed for a particular parameter and the filled circles indicate the averaged value of that parameter. The parameters include helical twist, glycosidic torsion angle and pseudorotation angle, as well as helical rise, roll and propeller twist at individual steps for the central hexanucleotide step in the d(G–G–T–A–T–A–C–C) duplex. Individual parameters vary dramatically with sequence both in terms of range and average value.

Observation of the *R*-factor during and subsequent to the relaxation matrix refinement shows that the *R*-factor falls sharply and reaches a stable minimum during the first picosecond of relaxation-restrained dynamics (Fig. 3A). Concomitant observation of the 'angle' energy term of the nucleic acid empirical energy potential[5] show a similar decline and stabilization (Fig. 3B). This observation is quite significant, for the 'angle' energy term generally reflects the agreement of a structure with the covalent terms of the empirical potential. Thus, from
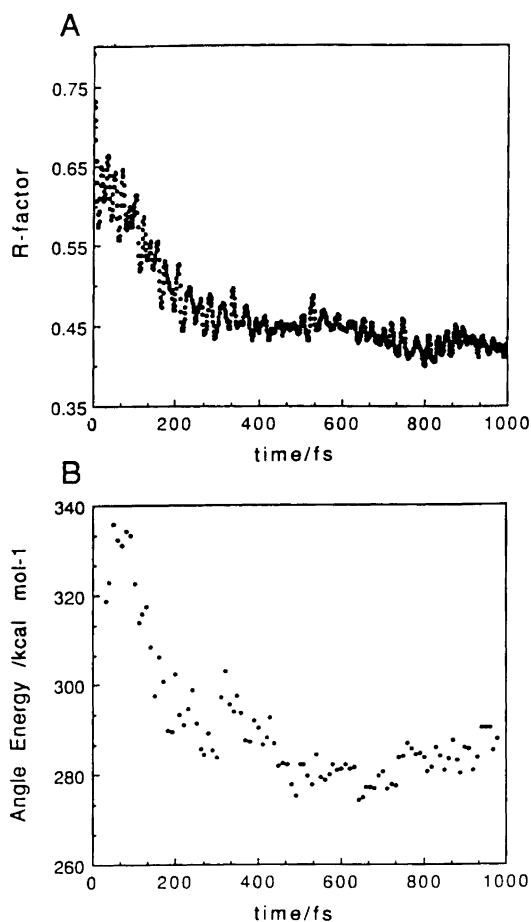
Fig. 3B, it becomes clear that experimental data obtained through a relaxation matrix approach is in better accord with the empirical potential than isolated spin pair approximation distance data.

Structural changes during relaxation matrix-restrained dynamics. To illustrate the structural changes achieved during relaxation matrix-restrained molecular dynamics which correspond to the spectral changes in the NOESY spectrum, we have undertaken a more in-depth study of the NMR relaxation environment of the protons of the T3 residue in the d(G–G–T–A–T–A–C–C) duplex. Fig. 4 shows plots of several interproton distances during the first picosecond of relaxation matrix-restrained molecular dynamics. Just as the overall *R*-factor reaches a constant minimum at around 1 ps of relaxation matrix-restrained molecular dynamics, the point labeled *c* in Scheme 1,

Fig. 3. (A) Decrease in the *R*-factor measured every 1 fs during the first ps of the relaxation matrix refinement starting from the DA1 structure using the molecular dynamics protocol presented in Scheme 1. (B) Concomitant decrease in the angle energy component which provides a measure of the fit to the empirical potential. Each point represents the running average over the previous 30 fs.
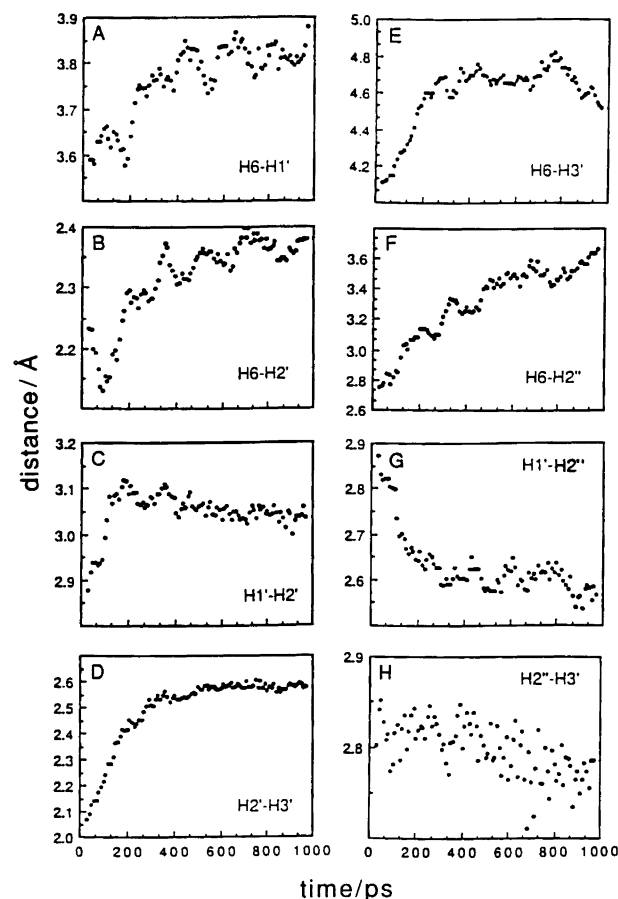

Fig. 4. The interproton separations within residue T3 are plotted every 10 fs during the first picosecond of relaxation matrix refinement starting from the DA1 structure using the molecular dynamics protocol presented in Scheme 1. Note that with a single exception, the interproton distances uniformly increase, indicating that the two-spin approximation systematically underestimates these distances in DNA. The plots are as follows: A, T3(H6)–T3(H1'); B, T3(H6)–T3(H2'); C, T3(H1')–T3(H2'); D, T3(H2')–T3(H3'); E, T3(H6)–T3(H3'); F, T3(H6)–T3(H2''); G, T3(H1')–T3(H2''); H, T3(H2'')–T3(H3').

these distances appear to converge to final values at this stage in the protocol. These plots, to some extent, represent the structural changes brought about by refinement against a relaxation matrix potential and, therefore, illustrate the divergences between the isolated spin pair approximation and complete relaxation matrix approaches.

Fig. 4H shows that the H2″ to H3′ distance for T3 remains relatively constant at a value around 2.8 Å during the period of dynamics observed here. In contrast, the H2′ to H3′ distance for T3 increases markedly from about 2.0 Å to 2.6 Å (Fig. 4D). Similar increases are observed for the H6 to H3′ distance for T3, which increases from roughly 4.1 Å to about 4.6 Å (Fig. 4E) and the H6 to H2″ distance for T3 which increases by nearly 1 Å from 2.7 Å to 3.6 Å (Fig. 4F). With only one exception, the inter-proton distances within T3 depicted in Fig. 4 uniformly increase as a result of the introduction of a relaxation matrix potential. This is apparently due to the systematic underestimation of distances derived from the isolated spin pair approximation.

Relaxation matrix-refined structures. Fig. 2B shows a comparison of the distance-refined structure DA1 (top) and the relaxation matrix-refined structure RA1 (bottom), i.e., the 22–24 ps dynamics-averaged structure of the molecular dynamics protocol in Scheme 1. These relaxation matrix-refined structures are designated RA1, RA2, RB1 and RB2 and were generated from starting distance-refined structures DA1, DA2, DB1 and DB2, respectively. The RMS deviations amongst these relaxation matrix-refined structures are listed in Table 2, and the corresponding interproton distances are tabulated in Table 3.

The helical parameters for the central six base pairs in the relaxation matrix-refined structures of the d(G–G–T–A–T–A–C–C) duplex are plotted in Fig. 6. The boxed region shows the range of values observed for a particular parameter and the filled circles indicate the averaged value of that parameter. The individual parameters for the relaxation matrix-refined structures (Fig. 6) do not show the dramatic sequence-dependent variations both in terms of range and average values observed for the distance refined structures (Fig. 5).

The central A4–T5 step has an average twist value of 31° which is lower than the average twist values of 38° observed at the other non-terminal steps (Fig. 6A). The average value of the glycosidic torsion angles vary over a narrow range of $-100°$ to $-115°$ (Fig. 6B) while the average value of the pseudorotation angle covers the range 140° to 175° (Fig. 6C). The average value of the helical rise is greater at the G2–T3 and A6–C7 steps (3.4 Å) compared with the internal T3–A4, A4–T5 and T5–A6 steps (3.0 to 3.1 Å) (Fig. 6D). The average value of the roll angles cover the range $-2°$ ($\pm 3°$) (Fig. 6E) while the average propeller twist falls within the 7° to 13° range (Fig. 6F).

Relaxation matrix refinement and R-*factor.* The graph of R-factor versus time during molecular dynamics presented in Fig. 3A provides a useful measure of the refinement process. However, we have found it necessary to compare each calculated and observed NOE intensity to avoid the possibility of single large errors which would escape detection as part of an overall R-factor. We have used the program GNOE (Hare Research, Inc., Woodinville, WA) to generate calculated two-dimen-
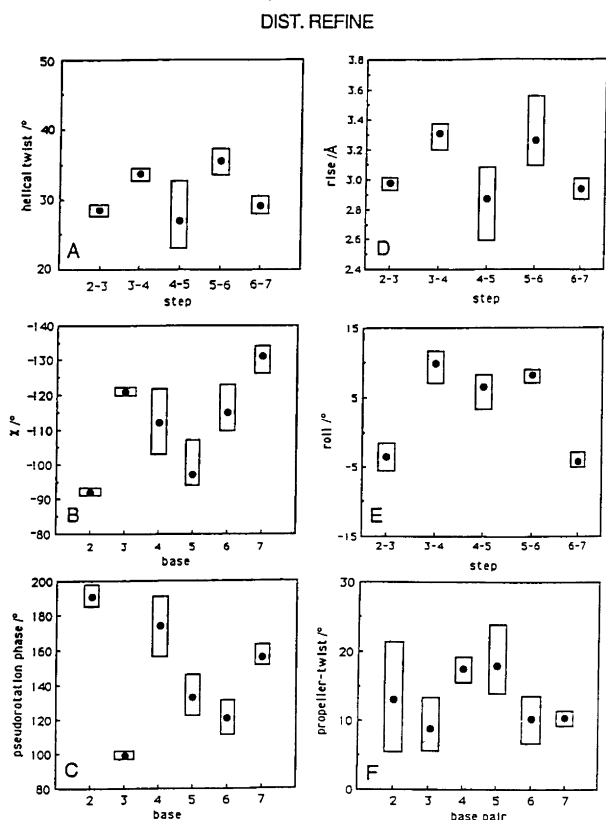
Table 3. Interproton distance (in Å) between non-exchangeable protons for the relaxation matrix-refined structure of d(G–G–T–A–T–A–C–C) duplex at 25°C.

| Residue | Intranucleotide distances | | | | | | |
	H1′–H8/H6	H2′–H8/H6	H2″–H8/H6	H3′–H8/H6	H1′–H2′	H1′–H2″	H1′–H3′
G1	3.86	4.15	4.79	4.95	3.06	2.62	3.88
G2	3.97	2.58	3.89	4.77	3.02	2.29	3.77
T3	3.83	2.26	3.43	4.62	3.06	2.57	3.76
A4	3.92	2.38	3.74	4.57	3.04	2.32	3.82
T5	3.80	2.14	3.33	4.57	3.05	2.38	3.77
A6	3.99	2.54	3.71	4.82	3.09	2.61	3.84
C7	3.79	2.30	3.54	4.31	3.05	2.35	3.72
C8	3.77	3.05	4.28	4.38	3.08	2.50	3.78

| Step | Internucleotide distances | | | | |
	H1′–H8/H6	H2′–H8/H6	H2″–H8/H6	H8/H6–H5/CH$_3$	H3′–H8/H6
G1–G2	3.86	3.75	2.23	—	4.84
G2–T3	3.09	4.12	2.50	4.06	5.15
T3–A4	3.43	4.31	2.93	—	5.54
A4–T5	3.75	3.79	2.42	3.65	5.12
T5–A6	3.53	4.51	2.94	—	5.72
A6–C7	3.80	3.88	2.74	3.76	5.46
C7–C8	3.81	4.51	2.77	4.84	5.26

DIST. REFINE

Fig. 5. Helical parameters for the distance-refined structures, DA1, DA2, DB1 and DB2 of the d(G–G–T–A–T–A–C–C) duplex. The boxed regions show the range of values observed for a particular parameter and the filled circles indicate the averaged structure. The plots are as follows: A, helical twist between adjacent base pairs; B, glycosidic torsion angles of individual residues; C, pseudorotation angles for individual sugar rings; D, rise between adjacent base pairs; E, roll angle between adjacent base pairs; F, propeller twist.

RELAX. REFINE

Fig. 6. Helical parameters in the relaxation matrix-refined structures, RA1, RA2, RB1 and RB2 of the d(G–G–T–A–T–A–C–C) duplex. The boxed regions show the range of values observed for a particular parameter and the filled circles indicate the averaged structure. The plots are as follows: A, helical twist between adjacent base pairs; B, glycosidic torsion angles at individual residues; C, pseudorotation angles for individual sugar rings; D, rise between adjacent base pairs; E, roll angle between adjacent base pairs; F, propeller twist.

sional NOE spectra from lists of calculated NOE intensities. The program generates Gaussian line-shapes using line-width and chemical shift parameters measured from the experimental spectra.

A more direct comparison between the experimental NOESY spectrum and the corresponding calculated spectrum is best visualized by recording stacked plots of specific expanded regions. We show four such expanded regions where the experimental stacked plot is compared with calculated stacked plots for A-DNA, B-DNA, distance-refined structure DA1 and relaxation matrix-refined structure RA1. The expanded regions of interest are base to H1' (Fig. 7, left panel), base to H3' (Fig. 7, right panel), base to H2', 2" (Fig. 8, left panel) and H1' to H2', 2" (Fig. 8, right panel).

It is readily apparent that the calculated NOESY stacked plots for A-DNA (Figs. 7E, 7J, 8E and 8J) differ significantly from their experimental counterparts (Figs. 7A, 7F, 8A and 8F) for data presented at a mixing time of 300 ms. The sugar pucker is in a C3'-*endo* conformation in A-DNA resulting in short distances between
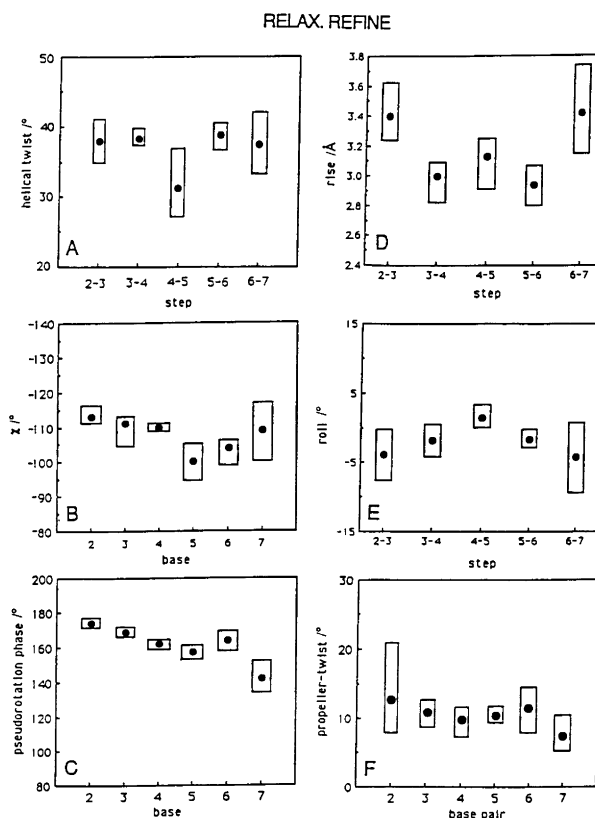
base (purine H8 and pyrimidine H6) and their own sugar H3' protons. This accounts for the strong NOE crosspeaks between base and H3' protons in A-DNA (Fig. 7J) which were not observed for their experimental counterparts (Fig. 7F). The short distance between the base proton and the H2' proton of its 5'-flanking residue

Table 4. Weighted and unweighted *R*-factors for initial, distance-refined and relaxation-matrix refined structures.

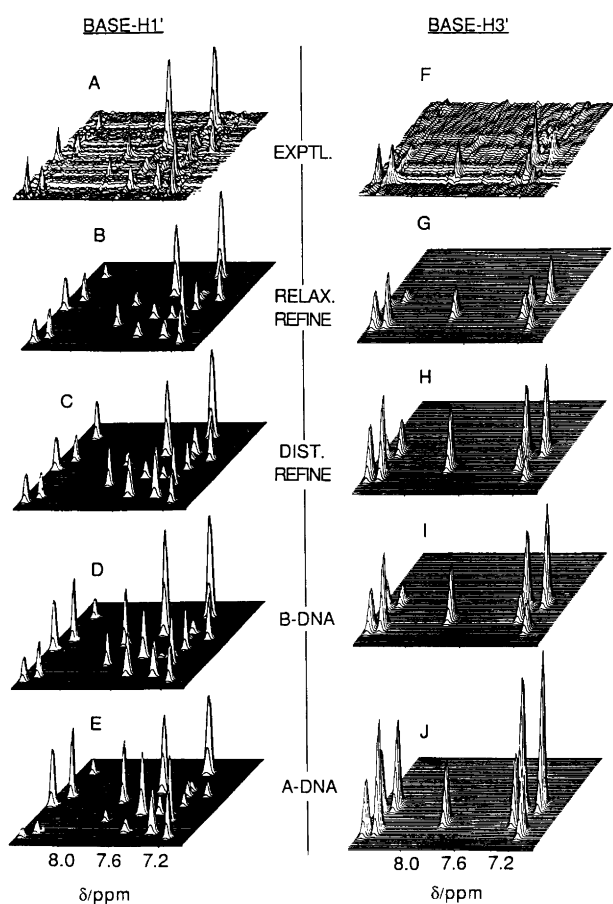|	Unweighted *R*-factor	Weighted *R*-factor
A-DNA	0.76	1.57
B-DNA	0.60	1.36
DA1	0.48	0.78
DA2	0.51	1.05
DB1	0.48	0.98
DB2	0.46	0.98
RA1	0.33	0.41
RA2	0.33	0.43
RB1	0.32	0.39
RB2	0.32	0.38

Fig. 7. (Left-hand panel). A comparison of stacked plots of the base to H1' region for the d(G–G–T–A–T–A–C–C) duplex, 25°C, recorded at a mixing time of 300 ms for (A) the experimental spectrum with (B) the corresponding relaxation matrix-refined RA1 structure, (C) distance-refined DA1 structure, (D) B-DNA structure and (E) A-DNA structure. (Right-hand panel). A comparison of stacked plots of the base to H3' region for the d(G–G–T–A–T–A–C–C) duplex recorded at a mixing time of 300 ms for (F) the experimental spectrum with (G) the corresponding relaxation matrix-refined RA1 structure, (H) distance-refined DA1 structure, (I) B-DNA structure and (J) A-DNA structure.

Fig. 8. (Left-hand panel). A comparison of stacked plots of the base to H2', 2" region for the d(G–G–T–A–T–A–C–C) duplex, 25°C, recorded at a mixing time of 300 ms for (A) the experimental spectrum with (B) the corresponding relaxation matrix-refined RA1 structure, (C) distance-refined DA1 structure, (D) B-DNA structure and (E) A-DNA structure. (Right-hand panel). A comparison of stacked plots of the H1' to H2', 2" region for the d(G–G–T–A–T–A–C–C) duplex recorded at a mixing time of 300 ms for (F) the experimental spectrum with (G) the corresponding relaxation matrix-refined RA1 structure, (H) distance-refined DA1 structure, (I) B-DNA structure and (J) A-DNA structure.

in A-DNA results in a calculated NOE plot for this region (Fig. 8E) which also differs significantly from the experimental counterpart (Fig. 8A). Similarly, pronounced differences are detected between the calculated spectrum for A-DNA (Fig. 7E) and the experimental spectrum (Fig. 7A) in the base to sugar H1' stacked plot regions. These comparisons establish that the A-DNA conformation cannot account for the NMR spectral parameters of the d(G–G–T–A–T–A–C–C) duplex in solution.

The difference between experimental and calculated NOESY plots is reduced on proceeding from A-DNA (Figs. 7E, 7J, 8E and 8J) to B-DNA (Figs. 7D, 7I, 8D and 8I) and reduced further on proceeding to the distance refined structure DA1 (Figs. 7C, 7H, 8C and 8H). We do, however, detect strong NOEs between the H5 protons of
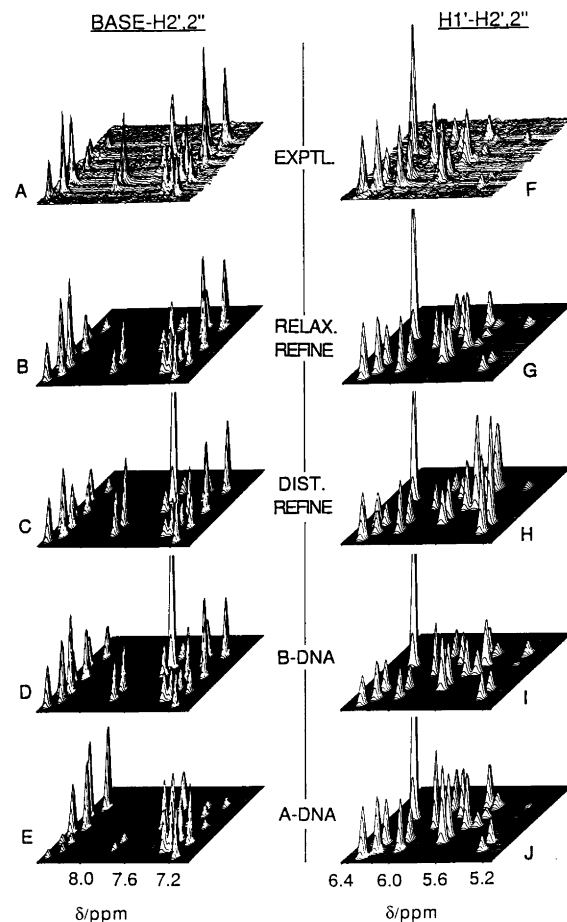
C7 (and C8) and the H2', 2" protons of their 5'-flanking residues for the DA1 structure (Fig. 8H) which were not detected in the experimental spectrum (Fig. 8F). Distance constraints involving these protons were not included in the distance restrained dynamics simulation and point out the importance of including as complete a distance bounds set as possible. Thus, if either lower or upper distance bounds are available, they should be included to the distance constraint list.

We observe an excellent match on comparing the experimental stacked plots (Figs. 7A, 7F, 8A and 8F) with those calculated for the relaxation restrained-refined structure RA1 (Figs. 7B, 7G, 8B and 8G). We point out that the conformation of the terminal residues of the RA1 conformation of d(G–G–T–A–T–A–C–C) are significantly distorted (data not shown) in order to

achieve a match between experimental and calculated NOEs for the terminal residues. This reflects fraying at the ends of the duplex[45] so that no single conformation of the terminal base pairs can account for the experimental data.

It is important to point out that the experimental and calculated NOESY plots are fit as a function of mixing time in the current study. This can best be visualized by the comparison of the NOE stacked plots of the base to H2', 2" region for the experimental data (left-hand panel, Fig. 9) and the relaxation matrix restrained-refined structure RA1 (right-hand panel, Fig, 9) at mixing time values of 30, 60, 90, 120 and 200 ms.

Backbone torsion angle constraints. We have used dihedral angle restraints in addition to empirical potentials for the backbone atoms in both the distance-restrained refinements and in the relaxation matrix-restrained refinements. We have restrained the α, β, ϵ and ζ dihedral angles, for which there are no relevant

experimental NOE data, to $\pm 40°$ of their B_1 values of $-70°$, $180°$, $180°$ and $-85°$, respectively.

It should be noted that these values encompass the backbone angles observed both in canonical A- and B-form DNA. Moreover, we note that the energy of restraining these angles is essentially zero. The dihedral constraint pseudoenergy begins and ends the refinement at 0.0 kcal mol^{-1}, and only occasionally rises as high as 0.1 kcal mol^{-1}. The primary purpose of these restraints is to overcome a difficulty in the molecular dynamics refinement method in which motions of completely unrestrained heavy atoms can propagate distortions. The final distribution of the backbone angles is not seen predominantly at the edges of the dihedral angle restraints, further indicating that the inclusion of these restraints is not a clear source of artificial structural determination.

Spin diffusion pathways. Three-dimensional NOE–NOE spectra facilitate the direct visualization of spin diffusion
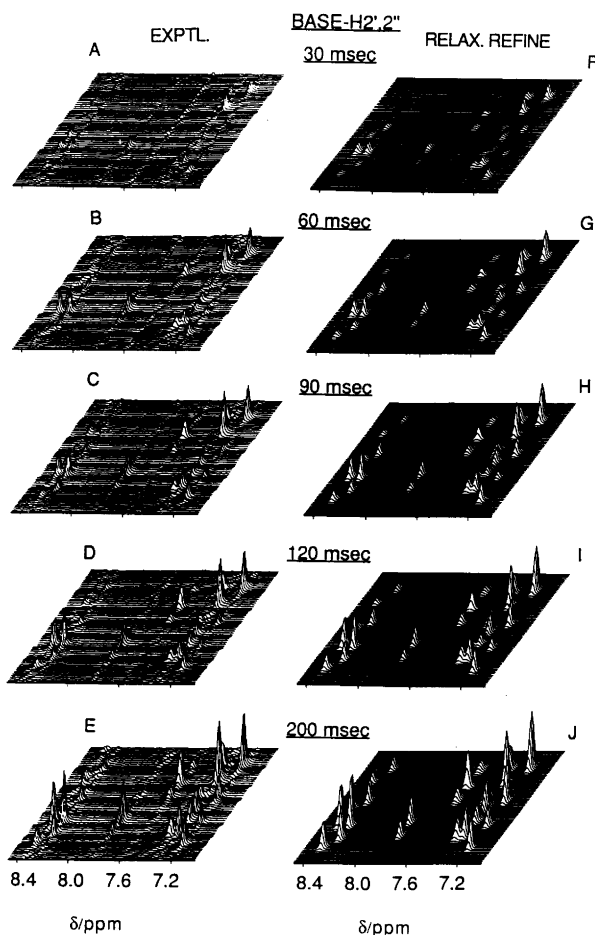

Fig. 9. Stacked plots showing the NOE build-up in the base to H2', 2" region of the NOESY spectrum of d(G–G–T–A–T–A–C–C) duplex, 25°C, at mixing times of 30, 60, 90, 120 and 200 ms. The experimental data are shown in the left-hand panel (plots A–E) while the corresponding calculated spectra for the relaxation matrix-refined RA1 structure are shown in the right-hand panel (plots F to J).
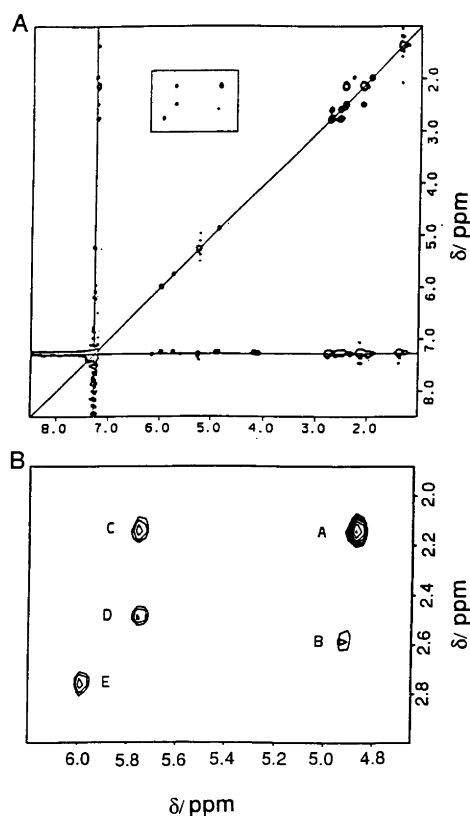

Fig. 10. (A) A contour plot of the D1 = T3(H6) plane of the three dimensional NOESY (120 ms mixing time)–NOESY (120 ms mixing time) spectrum of the d(G–G–T–A–T–A–C–C) duplex in 0.1 M NaCl, D$_2$O buffer at 25°C. (B) An expanded region of the boxed region in panel A. Cross peak A reflects a very strong spin diffusion pathway from T3(H6) → T3(H2') → T3(H3'). The other spin diffusion pathways are T3(H6) → G2(H2') → G2(H3') as represented by cross peak B, T3(H6) → T3(H2') → T3(H1') as represented by cross peak C, T3(H6) → T3(H2") → T3(H1') as represented by cross peak D, and T3(H6) → G2(H2") → G2(H1') as represented by cross peak E.

pathways as distinct off-diagonal crosspeaks. Thus, a three-spin diffusion pathway representing magnetization transfer from proton i to proton j to proton k is evidenced by the observation of a three-dimensional crosspeak at chemical shift position (w_i, w_j, w_k). Three-dimensional NOE–NOE spectroscopy can give insight into the processes which give rise to crosspeaks in two-dimensional NOESY spectra.

The base-H3' two-dimensional NOE crosspeak is, essentially, ubiquitously observed in NOESY spectra of DNA. The base-H3' distances observed in our relaxation matrix-refined structures are quite long, generally greater than 4.5 Å (Table 3) and thus direct NOE effects cannot account for the intensity of this crosspeak. Two lines of explanation can be drawn to account for the ubiquitous nature of this NOE interaction. The first envisions the base-H3' crosspeak as arising from spin diffusion effects. This model is effectively assumed in our relaxation matrix refinements. The second alternative postulates that the deoxy sugars exist in a fast equilibrium between C2'-endo and C3'-endo conformations. In the C3'-endo state, the H3' proton would be near to the base protons, as in A-DNA, and thus exhibit NOE crosspeaks through a direct rather than a spin diffusion interaction. Although we cannot rule out this second model, we can be certain that at least some of the base-H3' NOE intensity does arise from spin diffusion. In Fig. 10B the three-dimensional NOE–NOE spectrum shows a very strong H6 → H2' → H3' spin diffusion pathway. Thus the intensity of the H6–H3' two-dimensional NOE crosspeak, at least in some measure, arises from this spin diffusion interaction. This spin diffusion effect must arise from a C2'-endo-like sugar conformation rather than a C3'-endo conformation since the latter would show direct NOEs to the H3' proton which would not involve magnetization transfer through the H2' proton. The magnitude of the contribution of this spin diffusion pathway to the base-H3' crosspeak is probably responsible for the severe underestimation of this distance, which is one of the most prominent systematic errors in ISPA analyses of DNA.

Discussion

NMR relaxation matrix refinement should obviate the systematic errors in distance estimation which are inherent in the ISPA approach to distance measurement. This point is graphically illustrated in Scheme 2. The
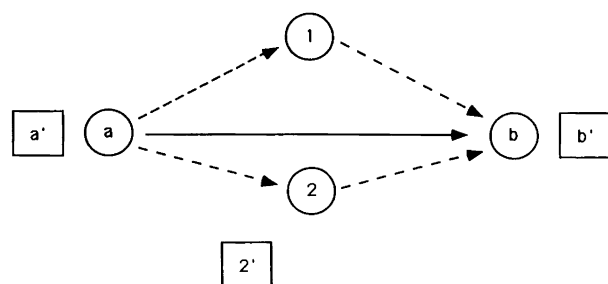
Scheme 2.

circles labeled 1, 2, a, and b represent hydrogen nuclei. Imagine that an NOE has been observed between spins a and b. This NOE will arise through a combination of direct and indirect interactions, which are depicted by the arrows. Hence, the solid arrow joining protons a and b represent direct spin–spin relaxation, whereas the dashed arrows represent spin diffusion interactions mediated by spins 1 and 2. The ISPA approach takes into account only the direct interaction, and ignores the contribution of the spin diffusion interactions to the NOE intensity. We have shown here that, for some NOEs in DNA, this contribution can be quite extensive.

The contribution of indirectly involved spins to the intensity of NOE crosspeaks adds a necessary complication to structure refinement. Say that the NOE calculated from a model structure between protons a and b in Scheme 2 was more intense than that measured from experimental NOESY spectra. The refinement algorithm could correct for this discrepancy in two distinct ways (or a combination thereof). The first possibility involves increasing the distance between spins a and b by, for example, moving these atoms to the boxed positions a' and b'. This approach will reduce the contribution of the direct interaction to the NOE intensity. Alternatively, the coordinates of one of the intermediary spins can be altered to reduce the contribution of spin diffusion interactions. This case is illustrated with a possible new position for spin 2 at the boxed position 2'.

When using the ISPA approach, one typically estimates a set of distances, and incorporates this approximate set of distances into the molecular dynamics potential for refinement. Thus, using the ISPA approach, the only possible refinement strategy is the first given above: atoms a and b must be moved to positions a' and b'. This is true because, in the distance potential, the positions of any mediating spins is irrelevant; they cannot be taken into account by this potential.

Relaxation matrix refinement is driven by the gradient of the relaxation matrix with respect to all proton coordinates. Thus, solutions which require altering of the positions of intermediary spins, even those to which NOEs may not be observed, should be well within the grasp of relaxation matrix refinement. Distance-based refinement schemes may artificially preclude some of these routes to refinement and may introduce yet another artificial bias into the final structures.

We have seen that in the case of the d(G–G–T–A–T–A–C–C) octamer, relaxation matrix refinement has had a large impact on evolving the final structure from a distance-refined intermediate structure. Interproton distances were observed almost uniformly to increase, as was predicted by theory, some by as much as 0.8 Å. Furthermore, many of the structural parameters of the octamer moved closer to canonical B-DNA values, and became more uniform. Finally, the empirical energy of the structure fell, as well. These observations support the conclusion that NMR relaxation matrix refinement has brought us to a solution structure which is nearer to

that which actually exists than we would have arrived at with distance-based refinement alone. This must be highlighted as the major conclusion of this work. Additionally, we also arrive at the disturbing possibility that many of the 'structural features' which have been observed in previous NMR studies of DNA oligomers may in fact have arisen as spin diffusion artefacts.

Relaxation matrix refinement has facilitated yet another advantage over distance-based refinement schemes, namely that the agreement of the structure with experimental data can now be quantified. In distance-based refinement schemes, this comparison could not be made; only the agreement with approximate distance bounds could be calculated, and this could not easily be translated into agreement with data. Relaxation matrix refinement has naturally given rise to an 'NMR R-factor' which quantifies the residual disagreement between observed and calculated NOESY spectra.

Zhou *et al.*[38] postulated, on the basis of a non-iterative relaxation matrix study of a limited number of DNA conformers, that the d(G–G–T–A–T–A–C–C) octamer exists in solution as a B–D–B duplex, i.e., that the G · C base pairs exist in a B-DNA conformation, whereas the A · T pairs exist in a D-DNA conformation. We have not reproduced these observations, but we cannot authoritatively dispute them. Although the Zhou structure gave a slightly higher R-factor in our relaxation matrix calculations, we could not achieve convergence to a totally B-like structure starting from this model.[46]

References

1. Noggle, J. H. and Schirmer, R. E. *The Nuclear Overhauser Effect*, Academic Press, New York 1971.
2. Hare, D. R. and Reid, B. R. *Biochemistry 25* (1986) 5341.
3. Hare, D. R., Shapiro, L. and Patel, D. J. *Biochemistry 25* (1986) 7445.
4. Hare, D. R., Shapiro, L. and Patel, D. J. *Biochemistry 25* (1986) 7456.
5. Nilsson, L. and Karplus, M. *J. Comput. Chem. 7* (1986) 691.
6. Nilges, M., Clore, G. M., Gronenborn, A. M., Piel, N. and McLaughlin, L. W. *Biochemistry 26* (1987) 3734.
7. Nilges, M., Clore, G. M., Gronenborn, A. M., Brünger, A. T., Karplus, M. and Nilsson, L. *Biochemistry 26* (1987) 3718.
8. Happ, C. S., Happ, E., Nilges, M., Gronenborn, A. M. and Clore, G. M. *Biochemistry 27* (1988) 1735.
9. Patel, D. J., Shapiro, L. and Hare, D. *Ann. Rev. Biophys. Chem. 16* (1987) 423.
10. Patel, D. J., Shapiro, L. and Hare, D. *Q. Rev. Biophys. 20* (1987) 35.
11. Kalk, A. and Berendsen, J. C. *J. Magn. Reson. 24* (1976) 343.
12. van de Ven, F. J. M. and Hilbers, C. W. *Eur. J. Biochem. 178* (1988) 1.
13. Macura, S. and Ernst, R. R. *Mol. Phys. 41* (1980) 95.
14. Keepers, J. W. and James, T. L. *J. Magn. Reson. 57* (1984) 404.
15. Lefevre, J. F., Lane, A. N. and Jardetzky, O. *Biochemistry 26* (1987) 5076.
16. Suzuki, E., Nagarajan, P., Zon, G. and James, T. L. *Biochemistry 25* (1986) 6854.
17. Banks, K. M., Hare, D. R. and Reid, B. R. *Biochemistry 28* (1989) 4828.
18. Nerdal, W., Hare, D. R. and Reid, B. R. *J. Mol. Biol. 201* (1988) 717.
19. Metzler, W. J., Wang, C., Kitchen, D. B., Levy, R. M. and Pardi, A. *J. Mol. Biol. 214* (1990) 711.
20. Borgias, B. A. and James, T. L. *J. Magn. Reson. 79* (1988) 493.
21. Boelens, R., Koning, T. M. G., van der Marel, G. A., van Boom, J. H. and Kaptein, R. *J. Magn. Reson. 82* (1989) 290.
22. Borgias, B. A. and James, T. L. *Methods Enzymol. 176* (1989) 169.
23. Borgias, B. A., Gochin, M., Kerwood, D. J. and James, T. L. *Prog. Nucl. Magn. Reson. Spectrosc. 22* (1990) 83.
24. Borgias, B. A. and James, T. L. *J. Magn. Reson. 87* (1990) 475.
25. Pieters, J. M., de Vroom, E., van der Marel, G. A., van Boom, J. H., Koning, T. M., Kaptein, R. and Altona, C. *Biochemistry 29* (1990) 788.
26. Nikonowicz, E. P. and Gorenstein, D. G. *Biochemistry 29* (1990) 8845.
27. Nikonowicz, E. P., Meadows, R. P. and Gorenstein, D. G. *Biochemistry 29* (1990) 4193.
28. Powers, R., Jones, C. R. and Gorenstein, D. G. *J. Biomol. Struct. Dyn. 8* (1990) 253.
29. Baleja, J. D., Moult, J. and Sykes, B. D. *J. Magn. Reson. 87* (1990) 375.
30. Baleja, J. D., Pon, R. T. and Sykes, B. D. *Biochemistry 29* (1990) 4828.
31. Baleja, J. D., German, M. W., van de Sande, J. H. and Sykes, B. D. *J. Mol. Biol. 215* (1990) 411.
32. Yip, P. and Case, D. A. *J. Magn. Reson. 83* (1989) 643.
33. Nilges, M., Habazettl, J., Brünger, A. T. and Holak, T. A. *J. Mol. Biol. 219* (1991) 499.
34. Shakked, Z. and Kennard, O. In: Jurnak, F. A. and McPherson, A., Eds., *Biological Macromolecules and Assemblies*, Vol. 2, Wiley, New York 1985.
35. Reid, D. G., Salisbury, S. A., Bellard, S., Shakked, Z. and Williams, D. H. *Biochemistry 22* (1983) 2019.
36. Jaimin, N., James, T. L. and Zon, G. *Eur. J. Biochem. 152* (1985) 157.
37. Patel, D. J., Shapiro, L. and Hare, D. *Biopolymers 25* (1986) 683.
38. Zhou, N., Bianucci, A. M., Pattabiraman, N. and James, T. L. *Biochemistry 26* (1987) 7905.
39. Brünger, A. T. X-PLOR, Copyright by the Trustees of Harvard University and the Trustees of Yale University.
40. Koning, T. M. G., Boelens, R. and Kaptein, R. *J. Magn. Reson. 90* (1990) 111.
41. States, D. J., Haberkorn, R. A. and Ruben, D. J. *J. Magn. Reson. 48* (1982) 286.
42. Boelens, R., Vuister, G. W., Koning, T. M. G. and Kaptein, R. *J. Am. Chem. Soc. 111* (1989) 8525.
43. Nilsson, L., Clore, G. M., Gronenborn, A. M., Brünger, A. T. and Karplus, M. *J. Mol. Biol. 188* (1986) 455.
44. Ryckaert, J. P., Cicotto, G. and Berendsen, H. J. C. *J. Comput. Phys. 23* (1977) 327.
45. Patel, D. J. and Hilbers, C. W. *Biochemistry 14* (1982) 2651.
46. Shapiro, L. and Eriksson, M. *Unpublished results.*