Journal of Biomolecular NMR, 6 (1995) 390–402
ESCOM

J-Bio NMR 314

Interproton distance bounds from 2D NOE intensities: Effect of experimental noise and peak integration errors

He Liu[a], H. Peter Spielmann[b], Nikolai B. Ulyanov[a], David E. Wemmer[b] and Thomas L. James[a,*]

[a]*Department of Pharmaceutical Chemistry, University of California, 926 Medical Science, San Francisco, CA 94143-0446, U.S.A.*
[b]*Department of Chemistry, University of California, Berkeley, CA 94720, U.S.A.*

Received 20 June 1995
Accepted 7 September 1995

Keywords: NOE distance determination; Random noise; High-resolution structure; Psoralen–DNA interstrand cross-link

Summary

The effect of experimental and integration errors on the calculation of interproton distances from NOE intensities is examined. It is shown that NOE intensity errors can have a large impact on the distances determined. When multiple spin ('spin diffusion') effects are significant, the calculated distances are often underestimated, even when using a complete relaxation matrix analysis. In this case, the bias of distances to smaller values is due to the random errors in the NOE intensities. We show here that accurate upper and lower bounds of the distances can be obtained if the intensity errors are properly accounted for in the complete relaxation matrix calculations, specifically the MARDIGRAS algorithm. The basic MARDIGRAS algorithm has been previously described [Borgias, B.A. and James, T.L. (1990) *J. Magn. Reson.*, **87**, 475–487]. It has been shown to provide reasonably good interproton distance bounds, but experimental errors can compromise the quality of the resulting restraints, especially for weak cross peaks. In a new approach introduced here, termed RANDMARDI (random error MARDIGRAS), errors due to random noise and integration errors are mimicked by the addition of random numbers from within a specified range to each input intensity. Interproton distances are then calculated for the modified intensity set using MARDIGRAS. The distribution of distances that define the upper and lower distance bounds is obtained by using N randomly modified intensity sets. RANDMARDI has been used in the solution structure determination of the interstrand cross-link (XL) formed between 4'-hydroxymethyl-4,5',8-trimethylpsoralen (HMT) and the DNA oligomer d(5'-GCGTACGC-3')$_2$ [Spielmann, H.P. et al. (1995) *Biochemistry*, **34**, 12937–12953]. RANDMARDI generates accurate distance bounds from the experimental NOESY cross-peak intensities for the fixed (known) interproton distances in XL. This provides an independent internal check for the ability of RANDMARDI to accurately fit the experimental data. The XL structure determined using RANDMARDI-generated restraints is in good agreement with other biophysical data that indicate that there is no bend introduced into the DNA by the cross-link. In contrast, isolated spin-pair approximation calculations give distance restraints that, when applied in a restrained molecular dynamics protocol, produce a bent structure.

Introduction

The determination of distance restraints from 2D NOE intensities is complicated by indirect magnetization transfer (i.e., spin diffusion), experimental and peak integration errors, and the local dynamics of the molecule. The determination of accurate distances from weak NOE intensities is especially challenging, since these intensities are often dominated by spin diffusion, and they are more susceptible to experimental and integration errors. Such distances are significantly underestimated when NOE data are interpreted using the isolated spin-pair approximation (ISPA), which ignores the effects of spin diffusion. To overcome this problem, a complete relaxation matrix approach, which explicitly takes into account all magnetization transfer pathways, is required. The MARDIGRAS algorithm (Borgias and James, 1990) has been developed to calculate accurate interproton distances from NOE

*To whom correspondence should be addressed.
Abbreviations: NOE, nuclear Overhauser effect; SD, standard deviation; HMT, 4'-hydroxymethyl-4,5',8-trimethylpsoralen; XL, psoralen–DNA interstrand cross-link.

intensities through an iterative complete relaxation matrix procedure. Similar to other complete relaxation matrix algorithms (Boelens et al., 1988; Borgias and James, 1988; Post et al., 1990; Madrid et al., 1991), a large number of NOE intensities, which may not be measurable because they are under the noise level or overlapped with other peaks, are supplied by simulated intensities based on an assumed dynamic model, an initial structure and an estimated correlation time. MARDIGRAS yields more accurate distances by making iterations to improve the internal consistency of the hybrid relaxation matrix and the agreement between the observed and simulated intensities (Borgias and James, 1990). More importantly, MARDIGRAS outputs upper and lower distance bounds for subsequent distance geometry (DG) and restrained molecular dynamics (RMD) calculations according to a user-defined noise level of the NOESY spectrum. In this work, we introduce a more rigorous approach, i.e., random error MARDIGRAS calculations (RANDMARDI), that take into account the effect of random noise and integration errors and yield more accurate distances and distance bounds.

With the complete relaxation matrix approach, intensities and their errors do not have a one-to-one correspondence to the calculated distances, i.e., an underestimated intensity does not necessarily yield an overestimated distance and vice versa. This poses several challenges to the accuracy of NOE-derived distances by any complete relaxation matrix approach. (i) Errors in the intensity of a given cross peak not only affect the distance calculated for that proton pair, but also propagate to other pairs due to spin diffusion. The interproton distances calculated for weak NOE intensities thus suffer not only from errors in their measured intensities, but also from errors in the measured intensities of strong peaks that are coupled by dipolar relaxation through spin diffusion. The *combination* of errors in different intensities is as important as the errors themselves. (ii) Weak peaks are the victims of both the absolute errors (such as the random noise of the spectrum) and the relative random errors of peak integration. The absolute noise per se has a larger impact on weaker peaks. Although the intensity errors have the same effect on all peaks, the accuracy of distances determined from weak peaks suffers seriously from the relative errors of strong peaks due to spin diffusion. (iii) Negative eigenvalues may occur with the hybridized NOE matrix due to experimental errors or inappropriate theoretical intensities. This eigenvalue problem consequently gives cross-relaxation rates with a wrong sign (conventionally, the rate is negative)*. A positive rate implies that, due to errors, the corresponding intensity (which emanates from spin diffusion as well as direct cross-relaxation) is weaker than the indirect magnetization transfers of spin diffusion. In this case, the distance cannot be calculated, and the intensity is thus termed 'nonphysical' or 'bad'. The problem of positive rates (originating from the eigenvalue problem) is almost always connected with weak intensities, where spin diffusion dominates the cross-peak intensity, and it prevents the determination of distance upper bounds. This observation is closely related to the geometry problems previously reported (Landy and Rao, 1993).

RANDMARDI effectively addresses all of these problems by incorporating the effects of noise and integration errors into the complete relaxation matrix calculations. Simulated random noise and integration errors are added to the original experimental intensities, and distances are calculated using MARDIGRAS. A distribution of distance estimates is obtained when N (default 30) modified data sets with the addition of random errors are used. This distribution for a given noise level and integration errors defines accurate distance bounds, suitable for use in DG or RMD calculations. More scattered distributions, with consequently wider bounds, will result for the distances that are relatively uncertain due to small intensity or strong spin diffusion. On the other hand, tight bounds will result if the intensity is strong and spin diffusion is weak. By randomly varying the intensities, RANDMARDI can calculate distances from intensities that are 'bad' (i.e., conflicting with other intensities, so that the corresponding distance cannot be calculated) in the original data set.

The effect of random intensity errors on the calculated distances was evaluated for three model systems using RANDMARDI. The ability of RANDMARDI to yield accurate distance bounds was verified using the fixed distances in the interstrand cross-link (XL) formed between 4'-hydroxymethyl-4,5',8-trimethylpsoralen (HMT) and the DNA oligomer d(5'-GCGTACGC-3')$_2$. By accurate distance bounds, we mean that the actual distance lies between the bounds determined. In the bonding unit T4-HMT-T12, 20 distances are covalently fixed (vide infra). The fixed distances range from 2.35 to 6.94 Å, and the types of distances include proton–proton, proton–methyl and methyl–methyl. Distance bounds for XL were calculated from experimental NOESY intensities using RANDMARDI. The average violation of the calculated bounds for the 20 fixed distances was 0.05 Å and the maximum violation was 0.26 Å.

Methods

The errors in NOE intensities can be divided into two types: experimental random noise and peak integration errors. The experimental noise is simulated by a series of

*The MARDIGRAS algorithm attempts to circumvent the potential problem of obtaining the logarithm of a negative eigenvalue by replacing that eigenvalue with its absolute value, or it can interpolate from correct eigenvalues.

random numbers with a randomly chosen sign, and is scaled by a factor so that the noise falls within a chosen level. The noise level is a constant, determined by the sensitivity of the experimental spectrum, and can be conservatively designated as the size of the weakest observed peaks. Random noise is an absolute error that is independent of the peak intensities. The integration error is defined as a percentage of the peak intensity; it is ascribed a random sign and is multiplied by a random number between 0 and 1 each time the error is added to the intensity. In the RANDMARDI procedure, the simulated random noise and random integration errors are added to the original experimental NOE intensities. N different sets of such randomly modified intensity data are obtained to imitate different possible errors and their combinations. The modified intensity A_{ij}^k in the *k*th set can be written as:

$$A_{ij}^k = A_{ij} + \sigma_1 \xi_1 E + \sigma_2 \xi_2 E_{ij} A_{ij} \quad (1)$$

where A_{ij} is the original experimental NOE intensity for protons i and j, E is the noise level (practically speaking, a constant determined by the weakest observed peak), E_{ij} is an estimated percentage integration error for peak A_{ij} (for example, the difference between intensities above and below the diagonal divided by their average), σ_1 and σ_2 are randomly chosen +1 or −1, and ξ_1 and ξ_2 are random numbers uniformly distributed between 0 and 1. By default, the program uses the original intensities as one of the N randomly varied intensity sets (the first set). Distances calculated from this set are useful as references, but they are not any better or worse than distances resulting from other sets if the intensity modification is based on well-estimated experimental and integration errors.

The distances are calculated from each of the N modified intensity sets using the MARDIGRAS algorithm. The number of intensity sets chosen (N) should be sufficiently large, so that the variation of each observed intensity represents a uniform sampling of errors within a given error level. Statistical tests have been done using two different random number generators. The mean value and variance for N random numbers were calculated. When N = 30, the deviations of the mean value and variance are within 6–7% of the values expected for uniform sampling. N = 30 is thus used as the RANDMARDI default. In applications, the number of times a distance can be calculated from a 'bad' intensity with N random variations can be used as a criterion of whether N is large enough. The distance for a 'bad' intensity should be calculated a reasonable number of times to allow an estimation of the distance bounds. In extreme cases, only one or two distances may result from a reasonably large number of intensity variations for a given proton pair using a well-estimated error level. This intensity should then be discarded.

For two isolated spins, the addition of random errors to the NOE intensity results in a distribution of distances (provided N is large enough) that is likely centered around neither the 'true' distance d_0 (which is normally unknown) nor the distance d_1, corresponding to the original intensity (i.e., without introduction of random error), but around a larger distance. The bias of the average distance is due to the approximate sixth power relation between distances and NOE intensities. Indeed, when the intensity is randomly decreased by some amount, the resulting distance is overestimated by an amount greater than it is underestimated due to the same amount of increase in intensity. For more than two spins, the deviation of average distances from d_1 due to the sixth power relation may be reversed by spin diffusion. When spin diffusion of the surrounding spins is dominating and is underestimated due to an error, the distance determined is smaller, and vice versa.

There are a number of ways to define the upper and lower distance bounds based on distances calculated from N sets of modified intensities. A definition used in RANDMARDI is:

$$d_{upper} = d_{avg} + SD \quad (2)$$

$$d_{lower} = d_{avg} - SD \quad (3)$$

where d_{avg} is the average distance calculated from the N intensity sets, and SD is the standard deviation. Another definition is:

$$d_{upper} = d_{max}(P) \quad (4)$$

$$d_{lower} = d_{min}(P) \quad (5)$$

where $d_{max}(P)$ and $d_{min}(P)$ are respectively the maximum and minimum of P percent of the N calculated distances left after (100 − P)/2 percent of the largest and smallest values were discarded. This definition helps to avoid a possible bias in d_{avg} and SD due to a few unrealistically large or small calculated distances. If the distance distribution is Gaussian, bounds defined by $d_{max}(67\%)$ and $d_{min}(67\%)$ are equivalent to $d_{avg}+SD$ and $d_{avg}-SD$. In general, a relatively conservative noise level and integration error should be used if $d_{max}(100\%)$ and $d_{min}(100\%)$, i.e., the maximum and minimum of all calculated values, are chosen as the distance bounds.

For weak intensities, the distance may not be calculated for all N variations. As mentioned above, a distance may not be calculated because its integrated NOE intensity is too weak (weaker than the spin-diffusion contributions from the neighbor). These 'bad' intensities correspond to the distance upper bounds (if calculated). When the number of times a distance can be calculated, n, is much smaller than the number of trials N, the upper bound d_{upper} may be underestimated. A correction to the

upper bounds may be applied based on the value of n. RANDMARDI provides the option to add $(N/n-1)*q$ to d_{upper}. Here q is a user-defined real number (the default is 0.04 Å). On the other hand, the distance can always be calculated if the intensity is strong due to an error, i.e., no matter how bad the error is, the lower bounds can always be calculated.

The distance bounds should bracket the true distances without using unrealistically large intensity errors. An important criterion for accurate distance bounds is that the fixed distances in the molecule fall within the calculated distance bounds. Ideally, these fixed distances should span the range of observable NOE values, and the protons defining the fixed distances should not be subject to serious local dynamics. The input error levels and the number of random variations may be adjusted to achieve better accuracy of distance bounds.

RANDMARDI calculates average distances and their bounds using NOESY intensities acquired at one mixing time, and based on one starting structure and one given correlation time. The final distance bounds for different mixing times, starting structures and correlation time estimations may be obtained using the in-house written program AVGBNDS. This program generates three different types of bounds: (i) average upper and lower bounds; (ii) average upper bounds plus their standard deviations and average lower bounds minus their standard deviations; and (iii) maximum upper bounds and minimum lower bounds. The results depend on the amount of error added to the intensities and on the data set itself. The bounds calculated for fixed distances in the system serve as an important criterion for selecting the appropriate set of bounds.

Model calculations

Our goal in the following calculations is to investigate the effect of different random noise and integration errors in NOE intensities on the derived distances and their distributions. Using three model systems, we will show that it is impossible to obtain accurate distances from erroneous NOE intensities, but it is feasible to determine accurate distance bounds that bracket the true distance. NOE intensities were simulated for the model systems using CORMA (Borgias and James, 1988,1990) for 500 MHz spectra and mixing times of $\tau_m = 50$, 150 and 250 ms, assuming isotropic motion with correlation times of $\tau_c = 3$ and 4 ns. In the simulations, NOE intensities and the noise levels are expressed in terms of 'normalized NOE intensity units' (n.u.). In this definition, the diagonal peak of a spin at $\tau_m = 0$ is 1.0 n.u., and the weakest observable NOE in a NOESY spectrum of 'good' sensitivity is typically 0.001 n.u. Roughly speaking, with ISPA, the distance ranges of 1.8–2.6, 2.4–3.6 and 3.4–5.0 Å correspond, respectively, to NOE intensities of 0.36–0.07 n.u. ('strong'), 0.10–0.01 n.u. ('medium'), and 0.014–0.001 n.u. ('weak') when $\tau_c = 4$ ns and $\tau_m = 100$ ms.

RANDMARDI calculations with 10 different absolute errors (0.001 to 0.01 n.u., with a step size of 0.001) and 11 relative errors (0 to 50%, with a step size of 10%) were carried out for the three model systems, using the simulated intensities as the original intensity input. Each RANDMARDI run used N = 30 random variations in the intensities. The purpose of these calculations was to assess the effect of different random errors on the estimated distances. The deviations of the average distances from the actual values and the distance distributions for a mixing time of 150 ms are plotted in Figs. 2, 3 and 4.

In another group of calculations, we investigated whether accurate distance bounds can be obtained from 'noisy' input intensities using RANDMARDI. The results were compared with ISPA and regular MARDIGRAS calculations. An absolute noise E = 0.002 n.u. and an integration error $E_{ij} = 10\%$ were chosen as being representative of typical experimental errors. For each model system, 30 randomly modified intensity sets were computed as described above (Eq. 1). A single 'worst' set that gave the largest errors* in the calculated distances was then selected, and further used as 'experimental' input intensities for the RANDMARDI calculations. RANDMARDI was run with the same noise levels (E = 0.002 n.u. and $E_{ij} = 10\%$) that were used to produce the set, and with the number of random variations N = 30. The resulting average distances and their bounds are listed in Tables 1–3.

HHH model: Three protons in a line

Three-proton systems have been studied previously to understand the error in distances due to spin diffusion of different geometries (Landy and Rao 1989,1993). It has been pointed out that with a moderate error of 0.0025 n.u. in the intensities, the distance determined for two protons i and j can be underestimated by over 30% if a third proton k is placed in-between the two protons i and j (the worst geometry for spin diffusion). The distance error arises from the large eigenvalue ratio of the short-distance interactions (often the spin-diffusion component) versus the long-distance interaction (in this case, the direct relaxation). The component with a large eigenvalue decays rapidly, and may fall below the noise level within typical mixing times (Landy and Rao, 1993). The distance has to be determined with uncertain knowledge about the spin diffusion. We show below that an estimation of the distance bounds for this system can be obtained by random sampling of the intensity errors. A correction to the

*The largest errors occur in the long distances when strong spin diffusion is involved; the errors are often negative, i.e., such distances are underestimated (Landy and Rao, 1993).

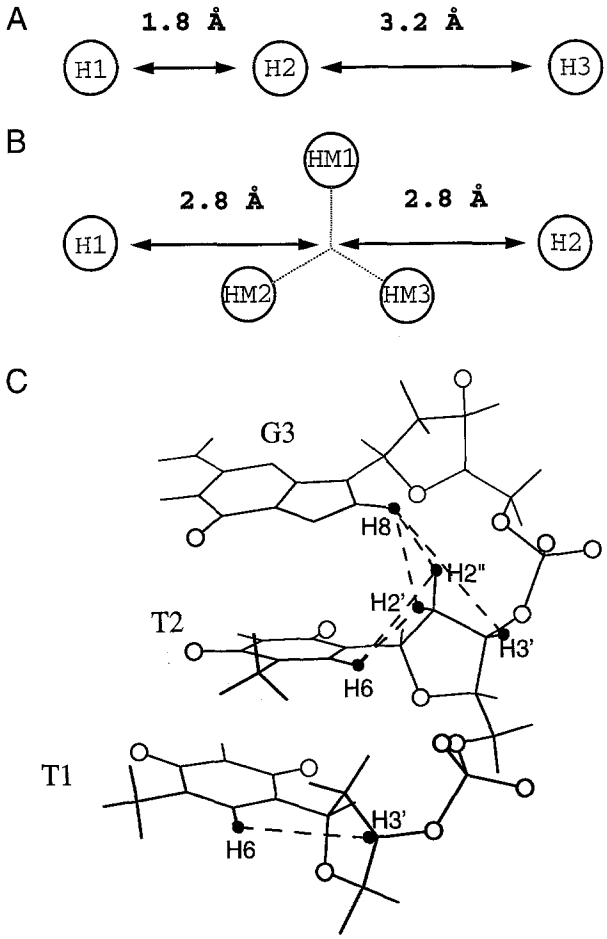

Fig. 1. (A) Model system HHH. The three protons are in a line and the distances H1-H2, H2-H3 and H1-H3 are 1.8, 3.2 and 5.0 Å, respectively. (B) The geometry of system HMH, where a methyl group is in line with two protons and the distances H1-M, H2-M and H1-H2 are 2.8, 2.8 and 5.6 Å, respectively. The methyl group is oriented in such a way that HM2 and HM3 are staggered by 30° from H1 and H2, respectively. (C) The structure of the trinucleotide dTTG. The open circles are oxygen atoms and the dashed lines indicate the interproton distances of interest. These distances can be found in Table 3.

upper bounds is needed if the eigenvalue problem is severe (vide supra).

Distances in the HHH system (Fig. 1A) were chosen so that the corresponding intensities are 'strong', 'medium', and 'weak', i.e., H1-H2 = 1.8, H2-H3 = 3.2 and H1-H3 = 5.0 Å. The intensities for proton pair H1-H2 (1.8 Å) are strong and not sensitive to either the absolute or the relative errors. The distance deviations of H2-H3 and H1-H3 for different relative errors are plotted respectively in Figs. 2A and B as a function of absolute error. The average distances of H2-H3 are slightly biased towards larger values. The maximum error in the average distances is ≤ 0.02 Å for absolute intensity errors below 0.004 n.u. The distance determined for H1-H3 is largely underestimated, because a large portion of the intensity (which is weak) comes from spin diffusion through the strong interaction between H1 and H2. The deviation of the average distance is as large as −1.0 Å for a noise level of 0.004 n.u. The distance deviation decreases slightly with longer mixing time due to stronger intensity (data not shown). The histograms of the distance distributions of the H1-H3 distance for different absolute errors and relative errors are plotted in Figs. 2C and D, respectively. It is seen that 30 random calculations yield a distribution of distances with the maximum occurrence shifted to a smaller value relative to the true distance (the vertical solid line). However, for the given errors, the true distances are within the bounds defined by the standard deviations of the distances, i.e, approximately d_{min} (67%) and d_{max} (67%).

Table 1 shows the distances calculated from the 'worst experimental' intensity set. ISPA gives the worst underestimated distances for H1-H3 and also overestimates the H1-H2 and H2-H3 distances. A single MARDIGRAS calculation, taking into account spin diffusion, yields better distances for all three proton pairs, but the large H1-H3 distance is still underestimated. RANDMARDI calculations with parameters E = 0.002, E_{ij} = 10% and N = 30 result in only a slightly better average distance compared to a single MARDIGRAS run, but the error bounds are accurate, i.e., the 'true' distances fall within the bounds defined by d_{min} and d_{max}. It should be noted that, for the 'worst' intensity set, the distance bounds defined by SD bounds are too narrow. It is also important to note that in some cases the distance lies almost at the exact boundary of the interval (d_{min}, d_{max}), e.g., H1-H3 at τ_m = 50 ms and H1-H3 at τ_m = 150 ms. This is because the noise level used with the RANDMARDI calculations corresponded exactly to the amount of noise in the input intensities (vide supra). This illustrates the importance of a correct estimation of experimental errors: if the noise parameters E and E_{ij} used in RANDMARDI were less than the errors in input intensities, the true distances might have been well outside of the calculated (d_{min}, d_{max}) bounds.

HMH model: In-line methyl system

The in-line geometry of HMH sketched in Fig. 1B is an exaggerated real situation encountered in the psoralen derivative (vide infra). H1 and H2 are in the plane formed by the methyl group protons, and they both are separated by 2.8 Å from the center of the methyl protons. The methyl group is responsible for the observation of a cross peak between H1 and H2. The rapid rotation of the methyl group can be described by discrete jumps (Tropp, 1980). The NOE intensity between a proton and a methyl group depends on their separation and the orientation of the methyl group. MARDIGRAS calculates the distance between a proton and the geometric center of a methyl group by iteratively varying the distances to fit all observed NOE intensities (Liu et al., 1991).

The distance deviations for H1-M (d = 2.8 Å) are plotted in Fig. 3A. The intensity between proton H1 and the

methyl is strong, and the random noise errors hardly make any difference for the derived distances; the average distance is almost unnoticeably biased towards longer distances. The maximum error in the average distance is within 0.02 Å. Figure 3B shows the calculated distance deviation for H1-H2 (d = 5.6 Å). The corresponding intensity is very weak and arises mainly from spin diffusion through the methyl group. The distance is underestimated by 1.5 Å if the noise level is 0.004 n.u. The histograms of the distance distributions of H1-H2 are plotted in Figs. 3C and D. The number of times the H1-H2 distance was calculated in the 30 trials decreased from 21 to 20 for noise levels of 0.001 to 0.004 n.u., and from 21 to 15 for relative errors of 5 to 20% (compared with the decrease from 30 to 21 for H1-H3 of the HHH system). About one third of the distances, corresponding to larger distance values, cannot be calculated. Subsequently, the distance distributions bias obviously to the smaller side. An additional correction to the upper bounds, based on the number of trials (N) and the number of times the distance can be calculated (n), should be applied (see the Methods section).

Table 2 shows the distances calculated using the 'worst' intensity set. Both MARDIGRAS and RANDMARDI yielded accurate values for the H-M distances, where the NOE intensities are strong and less vulnerable to errors. The H1-H2 distance is underestimated, but similar to the case of the HHH system, the RANDMARDI bounds are able to bracket the true distances.

TTG model: Trinucleotide

The in-line geometries considered above represent hypothetical worst cases with respect to interference from spin diffusion. Similar situations, however, are encoun-

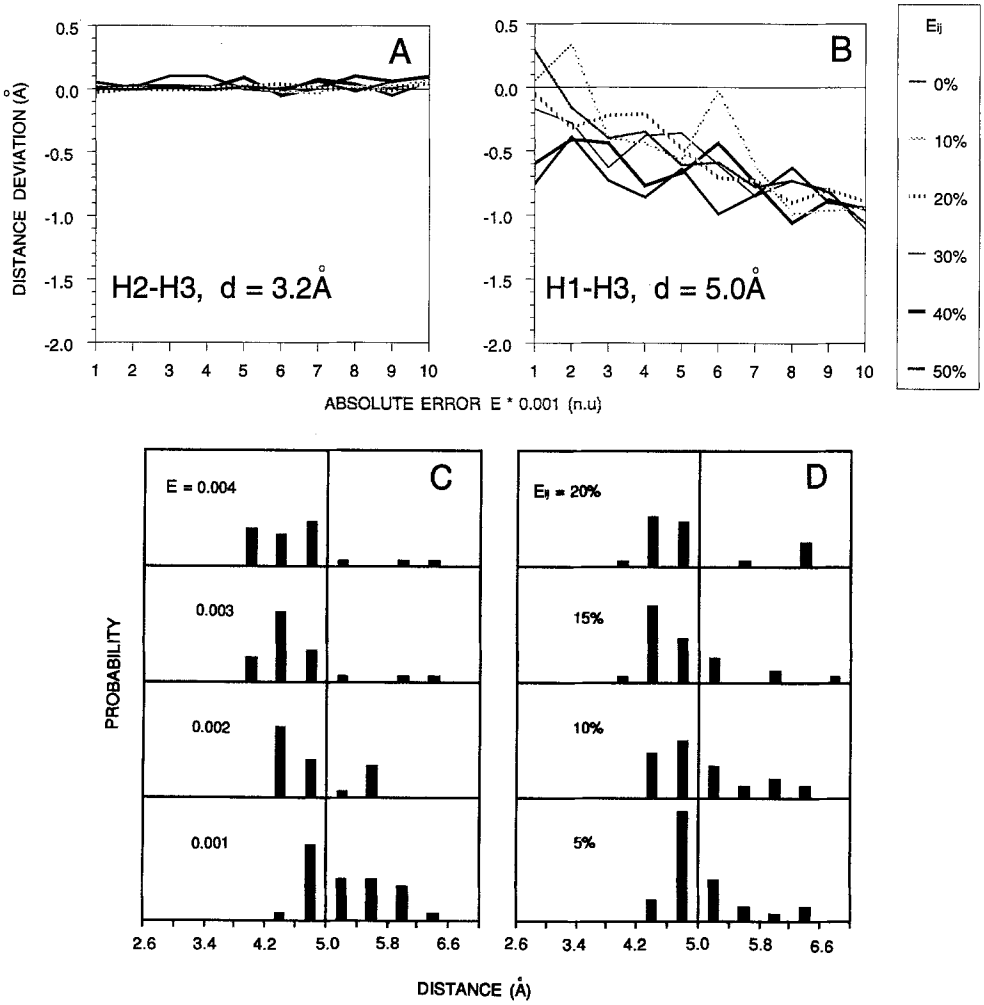

Fig. 2. The deviations of average distance from the actual value due to different noise and integration errors for model HHH at a mixing time of 150 ms for the H2-H3 distance (A) and the H1-H3 distance (B). H2-H3 yields a strong NOE intensity, and the resultant distances are slightly biased to larger distances. The intensity of H1-H3 is largely established by spin diffusion through the H2 proton. The distance deviations of H1-H3 due to random intensity errors are negative, except for a few cases where the noise is very low. The horizontal axis is the absolute error (noise level) in normalized NOE intensity units. Different lines represent relative errors from 0% to 50%. The distance distributions calculated for H1-H3 due to absolute errors of 0.001, 0.002, 0.003 and 0.004 are plotted in (C), and those due to relative errors of 5, 10, 15 and 20% in (D). The solid vertical lines designate the actual value of the distance.

TABLE 1
DISTANCES CALCULATED BY ISPA, MARDIGRAS AND RANDMARDI FOR AN HHH IN-LINE SYSTEM[a]

Atom_i	Atom_j	Ideal NOE	Error	d_true[b]	ISPA[c]	MARDI[c]	RANDMARDI[c]			
							d_avg[b]	d_min[b]	d_max[b]	SD[c]
τ_m = 50 ms										
H1	H2	0.23805	0.00455	1.80	1.90	1.80	1.80	1.80	1.80	0.00
H2	H3	0.00904	−0.00126	3.20	3.36	3.32	3.35	3.17	3.56	0.10
H1	H3	0.00201	0.00207	5.00	3.75	3.87	3.97	3.44	5.09	0.35
τ_m = 150 ms										
H1	H2	0.40850	0.03380	1.80	2.06	1.79	1.79	1.78	1.80	0.01
H2	H3	0.02208	−0.00317	3.20	3.48	3.36	3.38	3.24	3.51	0.07
H1	H3	0.01001	0.00103	5.00	3.81	4.27	4.28	3.90	5.21	0.32
τ_m = 250 ms										
H1	H2	0.43965	−0.03785	1.80	2.28	1.82	1.82	1.78	1.86	0.02
H2	H3	0.03244	−0.00174	3.20	3.50	3.24	3.23	3.11	3.38	0.07
H1	H3	0.01942	0.00172	5.00	3.72	4.09	4.27	3.63	5.63	0.47

[a] Input intensities were simulated with an absolute noise level of 0.002 n.u. and a relative error of 10%, as described in the text.
[b] d_true = the actual known distances; d_avg = RANDMARDI output distances; d_min and d_max = the maximum and minimum distances, respectively, of 30 RANDMARDI data sets.
[c] ISPA = isolated spin-pair approximation; MARDI = regular MARDIGRAS; RANDMARDI = calculations using 30 randomly varied intensity sets; SD = standard deviation of the distance distribution.

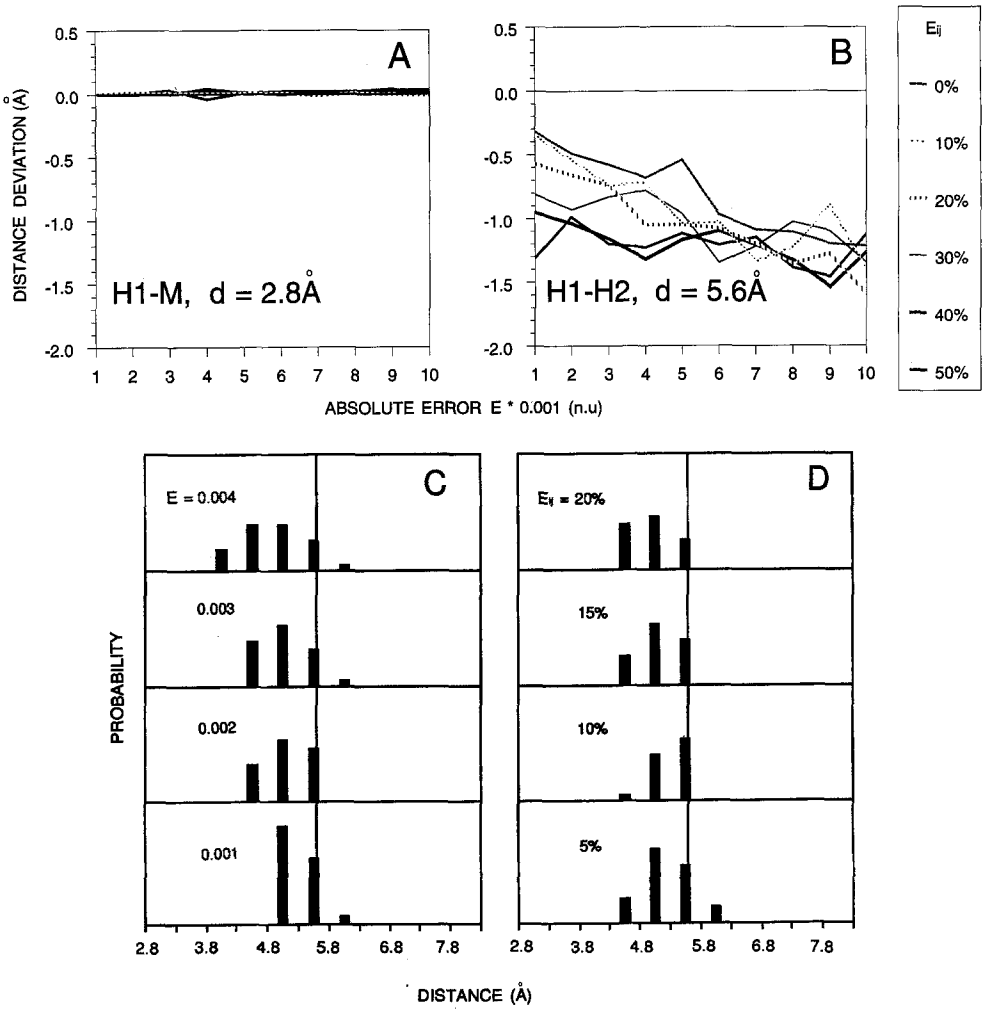

Fig. 3. (A and B) Distance deviations as a function of the absolute error for the HMH model at a mixing time of 150 ms. Different lines represent relative errors from 0 to 50%. (A) Distance errors of H1-M (d = 2.8 Å) are almost zero. (B) Distance H1-H2 (5.6 Å) is seriously underestimated. In (C) and (D), the distance distributions of H1-H2 (d = 5.6 Å) due to absolute errors of 0.001, 0.002, 0.003 and 0.004 and relative errors of 5, 10, 15 and 20% are plotted, respectively. The actual distance is represented by the solid vertical lines.

TABLE 2
DISTANCES CALCULATED BY ISPA, MARDIGRAS AND RANDMARDI FOR AN HMH IN-LINE SYSTEM[a]

Atom_i	Atom_j	Ideal NOE	Error	d_true	ISPA	MARDI	RANDMARDI			
							d_avg	d_min	d_max	SD
τ_m = 50 ms										
M	H1	0.08446	−0.00184	2.80	2.27	2.82	2.82	2.78	2.86	0.02
M	H2	0.08446	0.00248	2.80	2.25	2.79	2.80	2.76	2.83	0.02
H1	H2	0.00168	0.00184	5.60	3.84	4.10	4.22	3.67	5.83	0.49
τ_m = 150 ms										
M	H1	0.20029	−0.01649	2.80	2.39	2.82	2.81	2.77	2.85	0.02
M	H2	0.20029	−0.00639	2.80	2.36	2.79	2.79	2.76	2.83	0.02
H1	H2	0.01053	0.00143	5.60	3.76	4.50	4.60	4.00	5.90	0.43
τ_m = 250 ms										
M	H1	0.26619	−0.00319	2.80	2.45	2.80	2.80	2.76	2.84	0.02
M	H2	0.26619	−0.00869	2.80	2.46	2.81	2.81	2.77	2.85	0.02
H1	H2	0.02279	0.00260	5.60	3.61	4.44	4.51	3.86	5.97	0.56

[a] For an explanation of the parameters, see footnotes to Table 1.

tered during refinement of solution structures, for example with intranucleotide distances between base and sugar protons in DNA oligonucleotides. Accurate measurement of such distances is very important for the determination of sugar conformations (Schmitz et al., 1993; Ulyanov et al., 1995). However, the accuracy here can be compromised by strong spin diffusion via the geminal protons H2' and H2". To investigate this effect, we selected a single-stranded trinucleotide dTTG (Fig. 1C), which is part of a recently solved structure of a DNA duplex in solution (Weisz et al., 1994). An interesting feature of this structure is a sharp bend in the direction of the major groove, localized in the complementary dinucleotide TG : CA (Weisz et al., 1994). Such a bend is a common feature of TG sequences in solution (Ulyanov and James, 1994), in crystals (Gorin et al., 1995), and on polyacrylamide gels (Beutel and Gold, 1992).

For the purpose of this study, the atomic coordinates of the trinucleotide dTTG were calculated using the idealized bond length and bond angle values and the helical

TABLE 3
BASE-TO-SUGAR DISTANCES IN dTTG, CALCULATED BY ISPA, MARDIGRAS AND RANDMARDI[a]

Atom_i	Atom_j	Ideal NOE	Error	d_true	ISPA	MARDI	RANDMARDI			
							d_avg	d_min	d_max	SD
τ_m = 50 ms										
H2"(T2)	H8(dG3)	0.08708	−0.00391	2.02	2.13	2.02	2.02	1.98	2.07	0.02
H2'(T2)	H6(T2)	0.06042	−0.00274	2.15	2.26	2.14	2.14	2.10	2.19	0.02
H2'(T2)	H8(dG3)	0.02749	0.00044	2.68	2.60	2.72	2.72	2.60	2.87	0.07
H2"(T2)	H6(T2)	0.01053	0.00026	3.66	3.06	3.84	3.85	3.32	5.87	0.55
H3'(T1)	H6(T1)	0.00442	−0.00165	3.75	3.34	3.47	3.51	3.25	3.92	0.19
H3'(T2)	H8(dG3)	0.00279	−0.00164	4.30	3.52	3.71	3.74	3.33	4.44	0.29
τ_m = 150 ms										
H2"(T2)	H8(dG3)	0.14641	0.00221	2.02	2.36	2.03	2.04	1.97	2.09	0.03
H2'(T2)	H6(T2)	0.10205	0.00710	2.15	2.54	2.20	2.20	2.13	2.24	0.03
H2'(T2)	H8(dG3)	0.07557	−0.00184	2.68	2.62	2.66	2.67	2.49	3.00	0.13
H2"(T2)	H6(T2)	0.03936	−0.00389	3.66	2.89	3.04	3.10	2.81	3.97	0.28
H3'(T1)	H6(T1)	0.01714	−0.00012	3.75	3.37	3.83	3.84	3.58	4.37	0.19
H3'(T2)	H8(dG3)	0.01363	−0.00145	4.30	3.44	3.98	4.03	3.57	5.58	0.46
τ_m = 250 ms										
H2"(T2)	H8(dG3)	0.15025	−0.00875	2.02	2.53	1.96	1.90	1.71	2.05	0.09
H2'(T2)	H6(T2)	0.10492	0.00312	2.15	2.73	2.22	2.21	2.12	2.30	0.05
H2'(T2)	H8(dG3)	0.09744	0.00272	2.68	2.76	3.19	3.34	2.50	5.08	0.60
H2"(T2)	H6(T2)	0.05606	−0.00432	3.66	2.98	2.90	2.97	2.56	3.74	0.31
H3'(T1)	H6(T1)	0.03000	−0.00051	3.75	3.34	3.76	3.81	3.45	4.82	0.25
H3'(T2)	H8(dG3)	0.02562	−0.00083	4.30	3.42	4.06	4.08	3.49	6.77	0.64

[a] For an explanation of the parameters, see footnotes to Table 1.

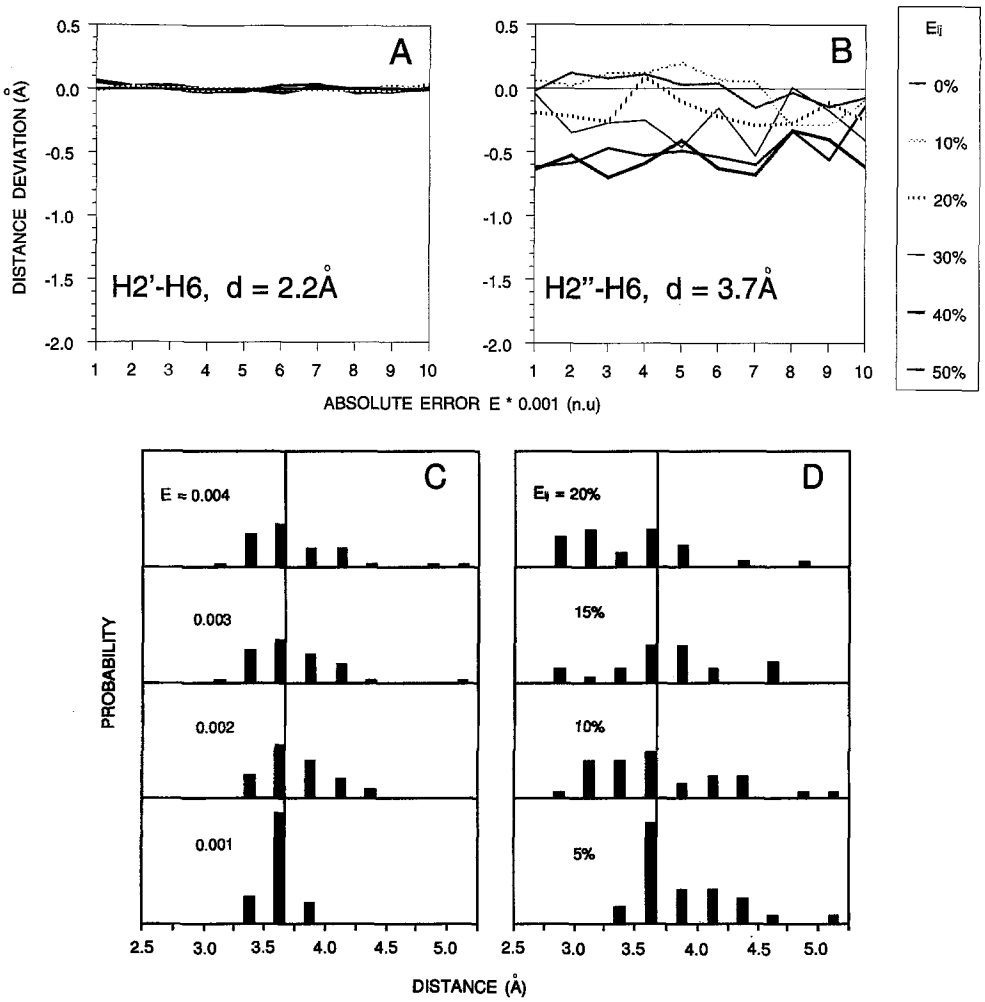

Fig. 4. For the dTTG trinucleotide model, (A) and (B) show the distance deviations on the second thymine of H2'-H6 (2.16 Å) and H2"-H6 (3.66 Å), respectively. The distance errors are plotted against the absolute errors from 0.001 to 0.01. Different lines represent relative errors from 0 to 50%. The deviation of distance H2'-H6 is small, with no obvious bias; the distance H2"-H6 is underestimated for most relative errors by up to 0.7 Å. In (C) and (D), the distance distributions of H1-H2 (5.6 Å) due to absolute errors of 0.001, 0.002, 0.003 and 0.004 and relative errors of 5, 10, 15 and 20% are plotted, respectively. Due to complex interactions in the dTTG system, a large standard deviation is expected when the relative error is significant. The actual value of the distance is represented by the solid vertical lines.

parameters of the solution DNA decamer (Weisz et al., 1994). Protons H2', H2" and H6 of the central thymine are positioned almost in line (Fig. 1C), the distances H2'-H2", H2'-H6 and H2"-H6 being 1.8, 2.2 and 3.7 Å, respectively. In addition, there are a number of other protons in the vicinity, creating a complex pattern of magnetization transfer pathways.

Figure 4A shows the distance deviation for H2'-H6 (2.16 Å) of the second thymine. There is no obvious distance bias, and the errors are within 0.1 Å. Figure 4B shows the distance deviation for H2"-H6 (3.66 Å). Subject to spin diffusion, H2" → H2' → H6, this distance is underestimated for almost all noise levels. Figures 4C and D display the histograms of the distance distributions of H2"-H6. Due to more complex interactions and strong spin diffusion of the geminal protons (H2' and H2") in the dTTG system, a relatively broader distribution is obtained for the H2"-H6 distance, especially when the relative error is large. The actual distance is well within the distance bounds defined by the standard deviation, but wider bounds are expected for these scattered distributions.

In Table 3, selected distances in dTTG, simulated using the 'worst' intensity set, are listed. MARDIGRAS gives accurate distances when the intensity is strong and therefore the effect of error is not important, e.g., for H2'(T2)-H6(T2) and H2"(T2)-H8(dG3). When the intensity is of the same order of magnitude ($\sim 10^{-3}$) as the absolute errors, the distance errors become significant, although much better than in the case of ISPA. The average distances calculated using RANDMARDI with parameters $E = 0.002$, $E_{ij} = 10\%$ (i.e., with the same amount of noise as present in the input intensities), and $N = 30$ are almost identical to those obtained with a single MARDIGRAS calculation. However, the calculated distance bounds are able to bracket the true distances. RANDMARDI in

general gives tight bounds for distances corresponding to strong NOE intensities that are not affected significantly by spin diffusion, and wide bounds for distances whose NOE intensities are weak and/or suffer from severe spin diffusion. Note that, in the most unfavorable case, the (d_{min}, d_{max}) bounds are so wide that they may not be of great use during structure refinement (e.g., $d_{min} = 3.3$ and $d_{max} = 5.9$ Å, when calculated from the 'noisy' intensities at $\tau_m = 50$ ms, while the true distance H2"(T2)-H6(T2) is 3.8 Å; see Table 3). However, this is a more acceptable situation than unrealistically tight but inaccurate bounds, which could lead to an incorrectly refined structure.

Distance bounds determined for the HMT–DNA cross-linked complex (XL)

RANDMARDI has been applied to obtain distance restraints for the determination of the structural changes induced in the DNA oligomer d(5'-GCGTACGC-3')$_2$ upon conversion of the 4'-hydroxymethyl-4,5',8-trimethylpsoralen (HMT) furanoside monoadduct (MAf) to the interstrand cross-link (XL) (Spielmann et al., 1995a,b). The distances that are fixed by the rigidness of the covalent bonds in XL range from 2.4 to 6.9 Å and span the range of observable NOE contacts. Therefore, they provide an excellent system for examining the accuracy of distance bounds obtained with different methods. NOESY spectra with mixing times of 50, 100, 175 and 250 ms were recorded for XL. The spectra were integrated using FELIX (Biosym Technologies, San Diego, CA). The cross peaks above and below the diagonal were treated independently. This gave eight sets of NOE intensities. The numbers of integrated peaks for the four mixing times are 205, 236, 259 and 263, respectively. There are 26 NOEs observed in the T4-HMT-T12 unit (Fig. 5), and 20 of them correspond to distances that are covalently fixed. Distances were determined using ISPA, regular MARDIGRAS and RANDMARDI. The reliability of these methods was evaluated by comparing the results to the actual fixed distances in HMT–thymine adducts (see Table 4).

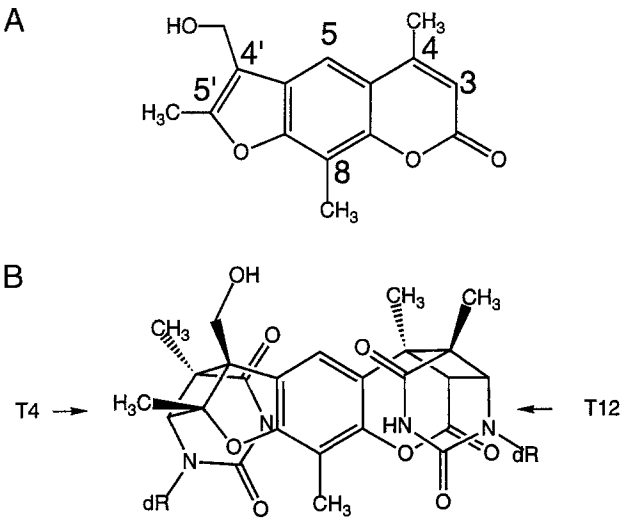

Fig. 5. (A) HMT (4'-hydroxymethyl-4,5',8-trimethylpsoralen); and (B) thymine-HMT-thymine interstrand cross-link. The T4-HMT-T12 unit in XL has 20 covalently fixed distances that correspond to observable NOE peaks.

ISPA

The buildup curves for mixing times of 50, 100, 175 and 250 ms were fitted to a second order polynomial function to determine the initial NOE buildup rates, and the distances were obtained by the sixth power ratio against the known distance of H5-H6 (2.52 Å) in the cytosines. The lower and upper bounds of a particular distance are the minimum and maximum of eight values determined using two sets of initial rates obtained from NOE peaks above and below the diagonal and taking measured intensities of four H5-H6 cytosine cross peaks (for each diagonal) as a reference. In Table 4, the ISPA distance bounds determined for the 20 fixed distances in the HMT–thymine adducts are shown. Among 15 distances whose bounds were calculated, only two distances are not violated by ISPA bounds. The average violation per distance is 0.5 Å; the worst violation is as great as −1.95 Å. There are five large distances whose bounds cannot be calculated because their NOE values were observed in only one of the four experiments, i.e., that with the longest mixing time (250 ms). In this case, the buildup rates cannot be calculated.

MARDIGRAS

The complete relaxation matrix method MARDIGRAS has been applied to the data. A correlation time of 3.75 ns was used in these simulations. The starting structure was obtained using Discover (Biosym Technologies) by energy minimization of the B-form DNA with HMT covalently bound to thymidines T4 and T12. The distances were averaged over the eight data sets, and the distance bounds were determined from the standard deviations. The distance bounds and their violation from the actual values are listed in Table 4. Out of a total of 20, 18 fixed distances were calculated by MARDIGRAS. The average violation of the actual distances from the bounds is 0.19 Å. The distance HMT_M4-T_4:M7 = 6.83 Å, which has the largest error (1.22 Å), cannot be calculated by ISPA. Distances cannot be calculated for HMT_M4-HMT_M8 (6.94 Å) and T_4:M7-HMT_M8 (5.84 Å) because the experimentally measured intensities correspond to bad intensities in the complete relaxation matrix analysis.

RANDMARDI

The experimental noise in the NOESY spectra has been taken into account by the RANDMARDI procedure. The noise level was set to one to five times the

TABLE 4
DISTANCE BOUNDS DETERMINED FOR THE FIXED DISTANCES[a] IN THE T4-HMT-T12 UNIT USING ISPA, MARDIGRAS AND RANDMARDI

Atom_i	Atom_j	d_{fix}	ISPA[b]				MARDIGRAS[c]				RANDMARDI[d]			
			lower	upper	width	viol.	lower	upper	width	viol.	lower	upper	width	viol.
T_12:H6	HMT_H3	2.35	2.24	2.41	0.17	0.00	1.96	2.15	0.19	0.20	1.89	2.19	0.30	0.16
HMT_H5	HMT_M4	2.67	2.24	2.40	0.16	0.27	2.58	2.87	0.29	0.00	2.57	2.91	0.34	0.00
T_12:H6	T_12:M7	2.70	2.49	2.67	0.18	0.03	2.70	2.85	0.15	0.00	2.61	2.96	0.35	0.00
T_4:H6	T_4:M7	2.74	2.40	2.57	0.18	0.17	2.71	2.84	0.13	0.00	2.66	2.90	0.24	0.00
T_4:H6	HMT_M5'	2.88	2.23	2.40	0.16	0.48	2.46	2.60	0.14	0.28	2.43	2.64	0.21	0.24
HMT_H3	HMT_M4	2.89	2.56	2.75	0.19	0.14	2.75	3.20	0.45	0.00	2.71	3.33	0.62	0.00
HMT_M4	T_12:M7	3.14	2.07	2.22	0.15	0.92	2.91	3.20	0.29	0.00	2.91	3.23	0.32	0.00
T_12:H6	HMT_M4	3.70	2.72	2.92	0.20	0.78	3.32	3.67	0.35	0.03	3.21	4.48	1.27	0.00
HMT_H3	T_12:M7	4.07	4.27	4.58	0.31	0.20	3.18	6.61	3.42	0.00	3.15	6.16	3.01	0.00
HMT_H5	T_12:M7	4.18	2.08	2.23	0.15	1.95	3.59	4.02	0.43	0.16	3.44	4.50	1.06	0.00
HMT_M5'	T_4:M7	4.46	–	–	–	–	4.29	4.92	0.64	0.00	3.84	6.41	2.57	0.00
HMT_H5	T_4:M7	4.56	3.88	4.16	0.29	0.40	3.96	6.43	2.47	0.00	3.64	6.28	2.64	0.00
HMT_H5	HMT_H3	4.71	4.11	4.42	0.30	0.29	3.80	4.14	0.34	0.57	3.30	4.96	1.66	0.00
HMT_H5	T_12:H6	5.36	4.44	4.77	0.33	0.59	4.05	4.74	0.69	0.62	3.03	5.12	2.09	0.24
HMT_H5	HMT_M5'	5.51	3.91	4.20	0.29	1.31	4.42	5.57	1.15	0.00	3.55	6.60	3.05	0.00
HMT_H5	T_4:H6	5.53	5.25	5.64	0.39	0.00	4.25	5.11	0.87	0.42	3.46	5.27	1.81	0.26
T_4:M7	HMT_M8	5.84	–	–	–	–	–	–	–	–	5.61	6.96	1.35	0.00
T_4:H6	HMT_M8	6.10	–	–	–	–	4.59	6.55	1.95	0.00	3.78	5.92	2.14	0.18
HMT_M4	T_4:M7	6.83	–	–	–	–	4.53	5.61	1.07	1.22	4.15	7.12	2.97	0.00
HMT_M4	HMT_M8	6.94	–	–	–	–	–	–	–	–	6.04	7.22	1.18	0.00
Average violation						0.50				0.19				0.05
Worst violation						1.95				1.22				0.26

[a] Fixed distances are measured from a model structure built by covalent bonding of the HMT moiety with thymidines T4 and T12, followed by energy minimization with the AMBER force field. The cyclobutane rings (formed by the 4', 5' carbons of HMT and the 5, 6 carbons of T4; and by the 3, 4 carbons of HMT and the 5, 6 carbons of T12, see Fig. 5) may undergo small conformational changes, but the T4-HMH-T12 unit is fairly rigid and the distances listed are considered fixed.

[b] ISPA bounds of a particular distance are determined as the minimum and maximum of eight values determined using NOE intensities measured above and below the diagonal and measured intensities of four H5-H6 cytosine cross peaks as a reference (see text). Five distances were not calculated because NOEs were observed only at $\tau_m = 250$ ms.

[c] MARDIGRAS distance bounds were determined from the average distances of all mixing times plus and minus their SD. Two distances were not calculated because of the eigenvalue problem (see text).

[d] RANDMARDI bounds were the average upper and lower bounds of all mixing times plus or minus their SD. Five distance bounds violate the actual distance value by up to 0.26 Å. Three of them are large distances (5.36, 5.53 and 6.10 Å). Two are short distances (2.35 and 2.88 Å) in the cyclobutane rings, whose conformations may not be completely rigid.

integrated intensity of the smallest cross peak for the eight data sets, and a 5% integration error was assumed for all intensities. The dynamic range of measured peak intensities was 1000. RANDMARDI calculations for 30 random intensity sets were carried out for each mixing time. The resulting distances were averaged to give the distances and bounds for each data set. The upper and lower bounds were the average distance bounds over all mixing times plus or minus their SD. The distance bounds calculated by RANDMARDI for the fixed distances can be found in Table 4. All 20 fixed distances were calculated (in fact, distances could be calculated from all observed NOEs), with an average bounds violation of 0.05 Å and the worst violation being 0.26 Å.

The psoralen–DNA interstrand cross-link structures were determined using RANDMARDI and ISPA restraints. The starting models for the refinement were generated by covalently binding the HMT with thymidines T4 and T12 (see Fig. 5), and energy minimizing with the AMBER force field function of the Discover module (Biosym Technologies) of InsightII with 100 steps of steepest-descent minimization, followed by conjugate gradient refinement. The NOE-derived distance restraints were applied to the models with 20 ps of RMD: at 350 K for 4 ps, followed by heating to 400 K for 4 ps and then cooling to 200 K in 50 K steps of 3 ps each. The final structure was then energy minimized to a maximum derivative of 0.01. All of the NOE-derived distance restraints were assigned an upper and a lower bound force constant of 50 kcal/mol^2. An additional 18 distance restraints were included to enforce Watson–Crick hydrogen bonding throughout the calculations. Three hydrogen bonds were included for each of the six G-C base pairs, where the upper and lower margins were set to 1.7–2.1 Å. No hydrogen-bond restraints were used for the A-T base pairs. Water and counterions were not included in the calculations. Helix parameters were calculated with the program CURVES 3.1 (Lavery and Sklenar, 1988,1989).

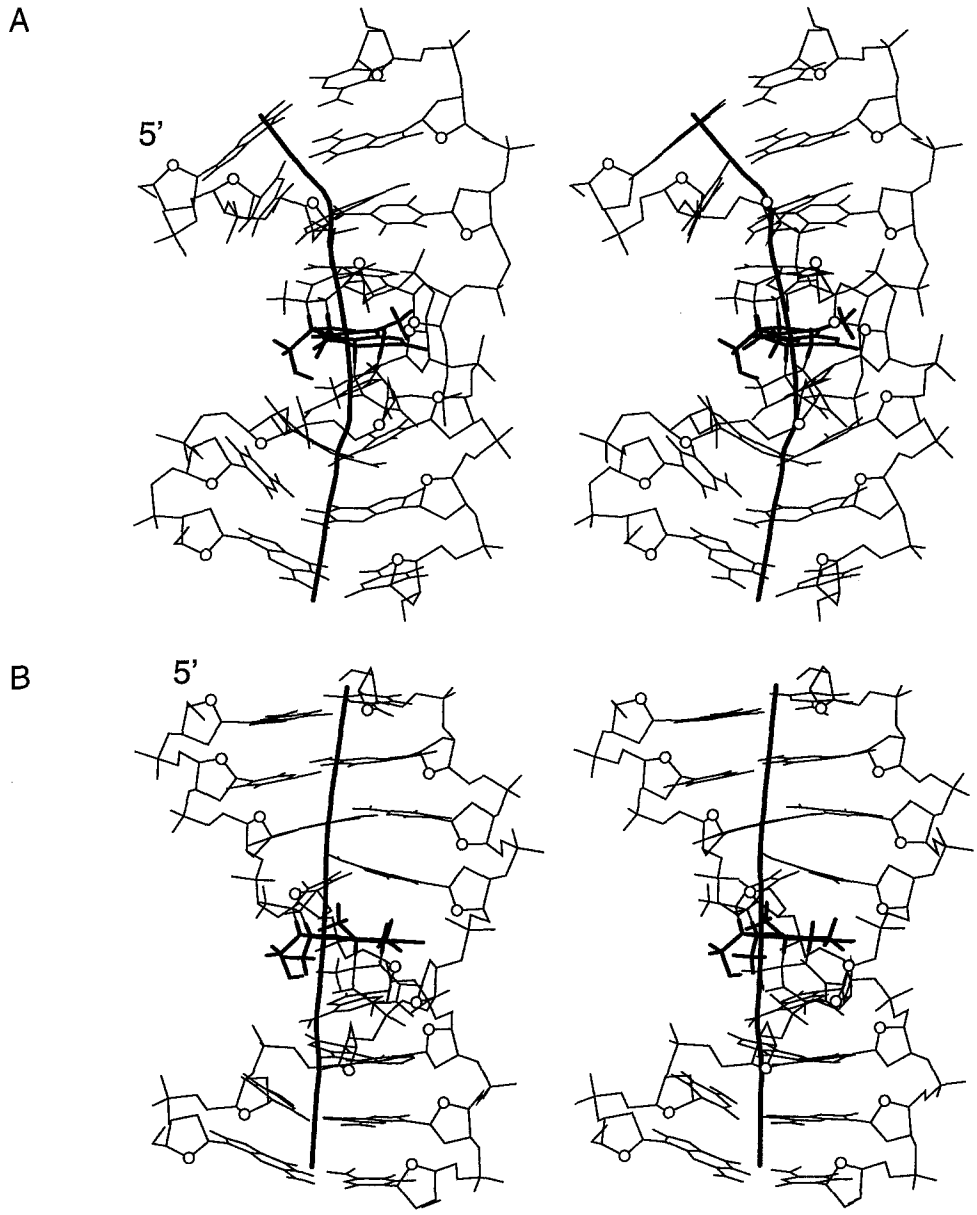

Fig. 6. The XL structures determined using RANDMARDI restraints (B) show that there is considerable local structural distortion, induced by intercalation of psoralen into the DNA, but the DNA returns to a B-form structure within three base pairs of the damage site. There is no significant bend in the helix axis. This agrees with the results of gel shift data. The structure determined using ISPA restraints (A) shows significant bending of the helix and highly strained distortions of the psoralen and thymidine bases.

The XL structures were displayed using the MIDASPlus program (Computer Graphics Laboratory, UCSF, San Francisco, CA). Figure 6B shows the interstrand cross-link structures determined using RANDMARDI. There is a considerable local structural distortion, induced by intercalation of psoralen into DNA, but the DNA returns to the B-form structure within three base pairs of the damage site. There is no significant bend in the helix axis, which agrees with the results of gel shift data (Sinden and Hagerman, 1984; Haran and Crothers, 1988). The structure shown in Fig. 6A was determined using ISPA restraints. It can be seen, in contrast to the experimental evidence, that the helix axis in the ISPA structure is significantly bent, and the psoralen and thymidine bases are highly distorted.

Discussion and Conclusions

Experimental and integration errors in NOE intensities have a large impact on the accuracy of the interproton distances, particularly large distances, determined from NOE intensities. The distance corresponding to a weak peak is not only the victim of errors in its intensity, but also of errors in intensities of all peaks coupled by spin diffusion. The model calculations show that long distances subjected to serious spin diffusion effects are most

sensitive to random errors. In such a case, those distances are typically underestimated because random errors, which increase cross-peak intensity, result in too short distances, while random errors, which decrease peak intensity, preclude distance determination due to the eigenvalue problem. The bias of large distances towards smaller values is a result of errors in the input NOE intensities. However, short distances, corresponding to strong NOE intensities, are slightly overestimated due to the sixth power relation between NOE intensities and distances. We have shown that by properly taking into account random errors in NOE intensities, RANDMARDI can generate accurate distance bounds that reflect the real uncertainty in the calculated distances due to experimental errors and spin diffusion. By randomly varying the input intensities within given noise and integration error levels, distance upper and lower bounds can be calculated from the distance distributions due to input errors. The model calculations show that 30 random variations within realistic error levels yield accurate distance bounds that bracket the true distance values. The upper bounds may require additional correction if the eigenvalue problem is severe, i.e., if the number of times a distance can be calculated is much smaller than the number of trials with random errors. In general, distances corresponding to weak NOE intensities have broad bounds, and strong intensities, on the other hand, produce tight distance bounds.

The distances and bounds of the psoralen–DNA complex formed between oligomer d(5'-GCGTACGC-3')$_2$ and 4'-hydroxymethyl-4,5',8-trimethylpsoralen (HMT) furanoside were calculated from experimental NOESY intensities using RANDMARDI. The fixed distances in the psoralen and thymidine bases of this complex provide an excellent opportunity for the evaluation of the distance bounds calculated by RANDMARDI. It has been shown that the distance bounds obtained for the fixed distances have an average violation of 0.05 Å, and the maximum violation was 0.26 Å (without additional correction of the upper bounds). The lower bounds are, however, well below the actual distances. RANDMARDI distance bounds for large distances are relatively broad and are asymmetric about the actual values. In the majority of situations, the actual distances are much closer to the upper bounds than to the lower bounds. The precision (i.e., the tightness of the bounds) of a calculated distance depends on the peak intensity and spin diffusion involved. The psoralen–DNA interstrand cross-link structure determined using RANDMARDI restraints is consistent with the structural characteristics defined by gel shifts and other non-NMR techniques. In contrast, ISPA restraints yielded an incorrectly bent structure.

In conclusion, we wish to emphasize that the RANDMARDI approach does not remove experimental errors from the integrated intensities. Rather, it adjusts the *precision* of the calculated distance bounds in order to preserve their *accuracy*. We have shown that this can be done if the account of experimental errors was estimated appropriately. The results of this work show that *accurate* distance bounds (even though *imprecise* in unfavorable situations) are important for structure determinations. RANDMARDI is a rational approach to determine NOE-derived distance bounds based on estimated experimental and integration errors.

Acknowledgements

This work was funded in part by the National Institutes of Health, Grants RR-01081 (T.L.J.), GM-39247 (T.L.J.), GM-43219 (D.E.W.), postdoctoral fellowship GM-14966 (H.P.S.), and through instrumentation grants from the U.S. Department of Energy, DE FG05-86ER-75281, and the National Science Foundation, DMB 86-09305 and BBS 87-20134, and by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098 (D.E.W.). The authors acknowledge use of the Computer Graphics Laboratory at UCSF.

References

Beutel, B.A. and Gold, L. (1992) *J. Mol. Biol.*, **228**, 803–812.

Boelens, R., Koning, T.M.G. and Kaptein, R. (1988) *J. Mol. Struct.*, **173**, 299–311.

Borgias, B.A. and James, T.L. (1988) *J. Magn. Reson.*, **79**, 493–512.

Borgias, B.A. and James, T.L. (1990) *J. Magn. Reson.*, **87**, 475–487.

Gorin, A.A., Zhurkin, V.B. and Olson, W.K. (1995) *J. Mol. Biol.*, **247**, 34–48.

Haran, T.E. and Crothers, D.M. (1988) *Biochemistry*, **27**, 6967–6971.

Landy, S.B. and Rao, B.D.N. (1989) *J. Magn. Reson.*, **83**, 29–43.

Landy, S.B. and Rao, B.D.N. (1993) *J. Magn. Reson. Ser. B*, **102**, 209–213.

Lavery, R. and Sklenar, H. (1988) *J. Biomol. Struct. Dyn.*, **6**, 63–91.

Lavery, R. and Sklenar, H. (1989) *J. Biomol. Struct. Dyn.*, **6**, 655–667.

Liu, H., Thomas, P.D. and James, T.L. (1991) *J. Magn. Reson.*, **98**, 163–175.

Madrid, M., Llinas, E. and Llinas, M. (1991) *J. Magn. Reson.*, **93**, 329–346.

Post, C.B., Meadows, R.R. and Gorenstein, D.G. (1990) *J. Am. Chem. Soc.*, **112**, 6796–6803.

Schmitz, U., Ulyanov, N.B., Kumar, A. and James, T.L. (1993) *J. Mol. Biol.*, **234**, 373–389.

Sinden, R.R. and Hagerman, P.J. (1984) *Biochemistry*, **23**, 6299–6303.

Spielmann, H.P., Dwyer, T.J., Sastry, S.S., Hearst, J.E. and Wemmer, D.E. (1995a) *Proc. Natl. Acad. Sci. USA*, **92**, 2345–2349.

Spielmann, H.P., Dwyer, T.J., Hearst, J.E. and Wemmer, D.E. (1995b) *Biochemistry*, **34**, 12937–12953.

Tropp, J. (1980) *J. Chem. Phys.*, **72**, 6035–6043.

Ulyanov, N.B. and James, T.L. (1994) *Appl. Magn. Reson.*, **7**, 21–42.

Ulyanov, N.B., Schmitz, U., Kumar, A. and James, T.L. (1995) *Biophys. J.*, **68**, 13–24.

Weisz, K., Shafer, R.H., Egan, W. and James, T.L. (1994) *Biochemistry*, **33**, 354–366.