charge distribution at the TS. The HF method overestimates the ionic character and yields large negative charges on N and X. In a later paper, we will compare the relative change in charge distribution from reactants to the TS and from the TS to the products.

Acknowledgment. A Killam postgraduate scholarship (to Z.S.) from Dalhousie University and the financial assistance of the Natural Sciences and Engineering Research Council of Canada in the form of an operating grant (to R.J.B.) are gratefully acknowledged. We are grateful to Dalhousie University Computing and Information Services for a generous allocation of computer time.

Registry No. CH_3H, 74-82-8; CH_3F, 593-53-3; CH_3Cl, 74-87-3; CH_3CN, 75-05-8; CH_3NC, 593-75-9; CH_3CCH, 74-99-7; CH_3OH, 67-56-1; CH_3SH, 74-93-1; CH_3NH_2, 74-89-5; H^-, 12184-88-2; F^-, 16984-48-8; Cl^-, 16887-00-6; OH, 14280-30-9; NH_2, 17655-31-1; CN, 57-12-5.

Supplementary Material Available: Tables of total energies at the HF, MP2', and MP2 levels of all species described herein (2 pages). Ordering information is given on any current masthead page.

On the Evaluation of Interproton Distances for Three-Dimensional Structure Determination by NMR Using a Relaxation Rate Matrix Analysis

Carol Beth Post,*[†,⊥] Robert P. Meadows,[‡] and David G. Gorenstein*[‡]

Contribution from the Department of Biological Sciences and Department of Chemistry, Purdue University, West Lafayette, Indiana 47907. *Received October 6, 1989*

Abstract: The accuracy of interproton distances obtained from two-dimensional nuclear Overhauser effect (NOESY) data using a relaxation rate matrix approach is examined by theoretical simulation studies. Interproton distances, the basis for three-dimensional structure determination of macromolecules by NMR, are most often evaluated from NOESY data by using a two-spin approximation or by grouping according to strong, medium, and weak intensities. A more rigorous analysis considers interactions within the full multispin system as specified by the relaxation rate matrix. With this matrix, distances are evaluated directly from measured NOESY volumes at a single mixing time taking into account indirect relaxation effects. However, numerical errors and mathematical difficulties can arise when solving such a matrix equation. Therefore the practicality of the matrix approach including experimental limitations on the input NOESY volumes was investigated. NOESY data were generated over a range of mixing times taking into account random noise, overlapping peak volumes, and the finite sensitivity for measuring cross-peak volumes by using proton coordinates from the crystal structure of lysozyme and a DNA dodecamer. A rigid molecule with a single overall correlation time was assumed. Comparison of the cross-relaxation rates, or interproton distances, obtained from the multispin matrix solution with the actual values indicates that there are errors in the matrix solution, but the errors are smaller than those obtained with the two-spin approximation under many but not all conditions of imperfect data.

I. Introduction

The ability to determine the three-dimensional structure of proteins and nucleic acid oligomers by using interproton distances measured by NMR has been successfully demonstrated.[1-6] In the NMR method, interproton distances evaluated from two-dimensional nuclear Overhauser effect (NOESY) data are used as restraints in a conformational search carried out by molecular dynamics,[7-9] including simulated annealing protocols,[10,11] distance geometry,[12,13] or a minimization method which employs dihedral angles as independent variables.[14] Applications of the NMR method for structure determination have relied on obtaining a large number of distance restraints, in addition to restraints on dihedral angles and hydrogen bonds.[15] In most cases, interproton distances are measured approximately based on an isolated two-spin relationship for cross-relaxation[16,17] or only qualitatively from strong, medium, and weak NOESY cross-peak intensities.[18,19] Lack of quantitative distance measurements necessitates loose restraints in the conformational search; restraints are classified into distance ranges or groups, where lower bounds may be specified by van der Waals radii.[17,20] Although the NMR method is established, there remains a need for increasing the precision of the structures obtained.[21,22] One aspect of the NMR method which could improve the structural solution is more accurate quantification of the interproton distances.

[†] Department of Biological Sciences.
[‡] Department of Chemistry.
[⊥] Present address: Department of Medicinal Chemistry, Purdue University.

(1) Kaptein, R.; Boelens, R.; Scheek, R. M.; van Gunsteren, W. F. *Biochemistry* **1988**, *27*, 5389-5395.
(2) Wuthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons: New York, 1986.
(3) Wuthrich, K. *Science* **1989**, *243*, 45-50.
(4) Clore, G. M.; Gronenborn, A. M. *Protein Eng.* **1987**, *1*, 275-288.
(5) Braun, W. *Q. Rev. Biophys.* **1987**, *19*, 115-157.
(6) Wright, P. E. *TIBS* **1989**, *14*, 255-260.
(7) Kaptein, R.; Zuiderweg, E. R. P.; Scheek, R. M.; Boelens, R.; van Gunsteren, W. F. *J. Mol. Biol.* **1985**, *182*, 179-182.
(8) Nilsson, L.; Clore, G. M.; Gronenborn, A. M.; Brunger, A.; Karplus, M. *J. Mol. Biol.* **1986**, *188*, 455.
(9) Clore, G. M.; Brünger, A. T.; Karplus, M.; Gronenborn, A. M. *J. Mol. Biol.* **1986**, *191*, 523-551.
(10) Nilges, M.; Clore, G. M.; Gronenborn, A. M. *FEBS Lett.* **1988**, *229*, 317-324.
(11) Nilges, M.; Gronenborn, A. M.; Brünger, A. T.; Clore, G. M. *Protein Eng.* **1988**, *2*, 27-38.
(12) Pardi, A.; Hare, D. R.; Wang, C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 8785-8789.
(13) Havel, T. F.; Crippen, G. M.; Kuntz, I. D. *Biopolymers* **1979**, *18*, 73-81.
(14) Braun, W.; Gō, N. *J. Mol. Biol.* **1985**, *186*, 611-626.
(15) Nerdal, W.; Hare, D. R.; Reid, B. R. *J. Mol. Biol.* **1988**, *201*, 717-739.
(16) Hare, D.; Shapiro, L.; Patel, D. J. *Biochemistry* **1986**, *25*, 7456-7464.
(17) Kline, A. D.; Braun, W.; Wuthrich, K. *J. Mol. Biol.* **1988**, *204*, 675-724.

0002-7863/90/1512-6796$02.50/0

Several workers have characterized the degree of error in a two-spin analysis of NOE data measured from a multispin system. The dependence of the error on the spatial arrangement of the magnetic nuclei has been elucidated with a three-spin model,[23,24] and investigations of multispin effects in larger systems have also been made.[22,25-28] In all studies it was found that indirect relaxation contributions to the observed NOE can be significant, particularly at long mixing times, for large molecules with long rotational correlation times or when the two protons under consideration are far apart. Errors in the distance obtained by the two-spin approximation can be as large as ±0.5 Å at shorter mixing times and >±1.0 Å at longer mixing times.[22,26,27]

In view of the error in distance values obtained by a two-spin analysis, it is worthwhile to consider the practical use of the full set of coupled differential equations describing relaxation in a multispin system. The multispin analysis takes into account all possible pairwise relaxation pathways rather than assuming the observed NOESY cross-peak intensity arises only from direct relaxation between the two protons defining the cross-peak. The most direct approach for evaluating interproton distances is to determine the cross-relaxation rates from the NOESY intensities.[26,27,29] [Some authors[29,30] use the term back transformation when referring to the direct calculation of relaxation rates from NOESY intensities, which is in contrast to the calculation of NOESY intensities from a molecular structure either by the inverse matrix operation[23,25,31,32] or by numerical integration procedures.[22,28,33,34]] Application of the direct matrix approach to experimental data has been limited. The structures of small peptides and an antibiotic have been determined by Olejniczak and co-workers[26,35,36] by using distances determined directly from NOESY volumes. For larger molecules, where measuring a complete data set can be difficult, Kaptein and co-workers[30,37] have developed a method involving a NOESY volume matrix which comprises both experimental intensities and theoretical intensities. The direct calculation of distances was also considered by Borgias and James[27] in a simulation study of a DNA octamer at two mixing times.

Analysis involving the relaxation rate matrix is more rigorous than a two-spin analysis; however, difficulties solving the matrix equation can arise; matrix analysis reduces to finding the solution to an eigenvalue problem and hence to matrix diagonalization. The problems associated with solving eigensystems are well-known: "You have probably gathered by now that the solution of eigensystems is a fairly complicated business".[38] Difficulties can arise from failure of the diagonalization algorithms to converge or from numerical errors in the solution as a result of experimental noise and incomplete data.

Given the importance of the NMR method for determining solution structures of macromolecules, a systematic study was undertaken to better define the strength and limitations of the direct matrix approach for evaluating interproton distances. We report here the results of simulation studies examining the convergence behavior and accuracy of the matrix solution as a function of mixing time, including effects from experimental limitations in measuring NOESY data. In the first section of this paper we describe the simulation and analysis of NOESY data with the program MORASS (Multispin Overhauser Relaxation Analysis and Simulation) for a multispin system. The reliability of the eigenvalue solution is examined in Results, first with noise free data. Effects on the accuracy due to experimental uncertainty in the NOESY intensities, including random noise, overlapping peaks, and sensitivity, are then considered. We compare the accuracy of the distance determined from a multispin analysis and eigenvalue solution with that of the two-spin approximation.

II. Theory

NOESY Intensities. The nuclear Overhauser experiment measures cross-relaxation rates, or the rates at which magnetization is transferred between nuclei. For protons, cross-relaxation occurs by through-space dipolar interactions and is a function of the interproton distance. A set of coupled differential equations describe the kinetics of magnetic relaxation in a multispin system.[31,32,39-41] For a two-dimensional NOESY experiment, the evolution of the spectral intensities as a function of the mixing time t_m is described by the simultaneous Bloch equations for relaxation which in matrix form[26,40,42] is

$$\frac{d}{dt}\mathbf{V}(t_m) = -\mathbf{\Gamma}\mathbf{V}(t_m) \tag{1}$$

The elements of the matrix $\mathbf{V}(t_m)$ are the peak volumes from the NOESY spectrum, and $\mathbf{\Gamma}$ is the symmetrical relaxation rate matrix

$$\mathbf{\Gamma} = \begin{bmatrix} \rho_1 & \sigma_{12} & \sigma_{13} & \sigma_{14} & \cdots \\ \sigma_{21} & \rho_2 & \sigma_{23} & \sigma_{24} & \cdots \\ \sigma_{31} & \sigma_{32} & \rho_3 & \sigma_{34} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

The off-diagonal elements, σ_{ij}, are cross-relaxation rates from which the distance between spins i and j is obtained, and the diagonal elements, ρ_i, are the direct relaxation rates for each spin i. (An expression identical to eq 1 involving an asymmetrical rate matrix with chemical kinetic rate constants applies to a multisite system in chemical exchange.[40,43-46])

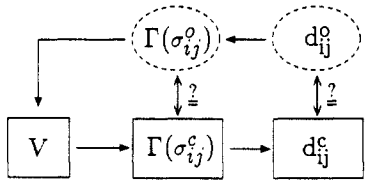
The solution to the differential eq 1 is

$$\mathbf{V}(t_m) = \exp(-\mathbf{\Gamma}t_m)\mathbf{V}_0 \tag{2}$$

The elements v_0 of the diagonal matrix $\mathbf{V}_0$ equal the volume of the diagonal peaks at $t_m = 0$. Rearrangement of eq 2 gives an

(18) Clore, G. M.; Gronenborn, A. M.; Brunger, A.; Karplus, M. *J. Mol. Biol.* **1985**, *186*, 435-455.
(19) Clore, G. M.; Gronenborn, A. M.; Kjaer, M.; Poulsen, F. M. *Protein Eng.* **1987**, *1*, 305-311.
(20) Lee, M. S.; Gippert, G. P.; Soman, K. V.; Case, D. A.; Wright, P. E. *Science* **1989**, *245*, 635-637.
(21) Güntert, P.; Braun, W.; Billeter, M.; Wüthrich, K. *J. Am. Chem. Soc.* **1989**, *111*, 3997-4004.
(22) Clore, G. M.; Gronenborn, A. M. *J. Magn. Reson.* **1989**, *84*, 398-409.
(23) Keepers, J. W.; James, T. L. *J. Magn. Reson.* **1984**, *57*, 404-426.
(24) Landy, S. B.; Rao, B. D. N. *J. Magn. Reson.* **1989**, *83*, 29-43.
(25) Dobson, C. M.; Olejniczak, E. T.; Poulson, F. M.; Ratcliffe, R. G. *J. Magn. Reson.* **1982**, *48*, 97-110.
(26) Olejniczak, E. T.; Gampe, R., Jr.; Fesik, S. *J. Magn. Reson.* **1986**, *67*, 28-41.
(27) Borgias, B. A.; James, T. L. *J. Magn. Reson.* **1988**, *79*, 493-512.
(28) Madrid, M.; Mace, J. E.; Jardetzky, O. *J. Magn. Reson.* **1989**, *83*, 267-278.
(29) Bremer, J.; Mendz, G. L.; Moore, W. J. *J. Am. Chem. Soc.* **1984**, *106*, 4691-4696.
(30) Boelens, R.; Koning, T. M. G.; Kaptein, R. *J. Mol. Struct.* **1988**, *173*, 299-311.
(31) Bothner-by, A. A.; Noggle, J. H. *J. Am. Chem. Soc.* **1979**, *101*, 5152-5155.
(32) Macura, S.; Ernst, R. R. *Mol. Phys.* **1980**, *41*, 95-117.
(33) Banks, K. M.; Hare, D. R.; Reid, B. R. *Biochemistry* **1989**, *28*, 6996-7010.
(34) Lefèvre, J.; Lane, A. N.; Jardetzky, O. *Biochemistry* **1987**, *26*, 5076-5090.
(35) Fesik, S. W.; O'Donnell, T. J.; Gampe, R. T., Jr.; Olejniczak, E. T. *J. Am. Chem. Soc.* **1986**, *108*, 3165-3170.
(36) Fesik, S. W.; Bolis, G.; Sham, H. L.; Olejniczak, E. T. *Biochemistry* **1987**, *26*, 1851-1859.
(37) Boelens, R.; Koning, T. M. G.; van der Marel, G. A.; van Boom, J. H.; Kaptein, R. *J. Magn. Reson.* **1989**, *82*, 290-308.
(38) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes, The Art of Scientific Computing*; Cambridge University Press: New York, 1986; p 340.
(39) Bodenhausen, G.; Ernst, R. R. *J. Am. Chem. Soc.* **1982**, *104*, 1304-1309.
(40) Abel, E. W.; Coston, T. P. J.; Orrell, K. G.; Sik, V.; Stephenson, D. *J. Magn. Reson.* **1986**, *70*, 34-53.
(41) Grassi, M.; Mann, B. E.; Pickup, B. T.; Spencer, C. M. *J. Magn. Reson.* **1986**, *69*, 92-99.
(42) Massefski, W., Jr.; Bolton, P. H. *J. Magn. Reson.* **1985**, *65*, 526-530.
(43) Perrin, C. L.; Gipe, R. K. *J. Am. Chem. Soc.* **1984**, *106*, 4036-4038.
(44) Muhandiram, D. R.; McClung, R. E. *J. Magn. Reson.* **1987**, *71*, 187-192.
(45) Post, C. B.; Ray, W. J., Jr.; Gorenstein, D. G. *Biochemistry* **1989**, *28*, 548-558.
(46) Johnston, E. R.; Dellwo, M. J.; Hendrix, J. *J. Magn. Reson.* **1986**, *66*, 399-409.

Scheme I



expression for the relaxation rates in terms of the NOESY volumes:[26,29,43]

$$-\Gamma = \frac{1}{t_m} \ln [V(V_0)^{-1}] \quad (3)$$

The logarithm in eq 3 is readily evaluated by substitution with the eigenvalues and eigenvectors of $V(V_0)^{-1}$:

$$-\Gamma = \frac{1}{t_m} K(\ln [L])K^T \quad (4)$$

Matrix **K** and the diagonal matrix **L** contain the eigenvectors and eigenvalues of $V(V_0)^{-1}$, respectively. The evaluation of the cross-relaxation rates therefore reduces to the eigenvalue problem of determining **K** and **L**.

We note that the solution for Γ does not require assignment of the NOESY volumes to specific protons; all that is necessary is to correctly associate diagonal and off-diagonal volumes for identification of elements in **V**. Further, the direct solution for Γ provides values for σ_{ij}, without any assumptions about the relaxation mechanism or parameters such as motional correlation times.[26] Of course, the interpretation of σ_{ij} in terms of interproton distances does require such assumptions.

The present work investigates by simulation techniques the reliability with which cross-relaxation rate constants can be determined from NOESY data by using the eigenvalue solution, eq 4. NOESY data were generated (eq 2) by using σ_{ij} and ρ_i values obtained from expressions for dipolar relaxation in an n-spin system[47]

$$\sigma_{ij} = \frac{\gamma^4\hbar^2}{10\ \langle d_{ij}^3\rangle^2}[6J(2\omega) - J(0)] \quad (5)$$

$$\rho_i = \sum_{j\neq i}^{n} \frac{\gamma^4\hbar^2}{10\ \langle d_{ij}^3\rangle^2}[J(0) + 3J(\omega) + 6J(2\omega)]$$

$$J(\omega) = \left(\frac{\tau_c}{1 + \omega^2\tau_c^2}\right)$$

ω and γ are the resonance frequency and the gyromagnetic ratio for ^1H nuclei, respectively. The spectral densities, $J(\omega)$, specifying the transition probability are defined with a single, overall rotational correlation time, τ_c, and interproton distances, d_{ij}, are obtained from crystallographic or model structures. The only internal motion considered in the simulation is rotation of methyl groups. As methyl rotation occurs on a time scale of 0.01–0.2 ns,[48,49] significantly faster than the overall molecular rotation of proteins and DNA oligomers, an interproton distance involving a methyl proton was considered to be an average over the three protons in the methyl group. That is, adopting a three-site jump model, relaxation between a methyl and non-methyl proton pair or between two methyl groups is proportional to the interproton distance[48,50] averaged over methyl protons:

$$\frac{1}{m}\sum_{j=1}^{m}\frac{1}{\langle d_{ij}^3\rangle^2}$$

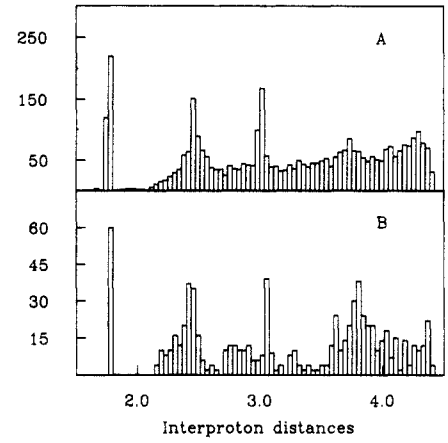


Figure 1. Interproton distance distribution for $d_{ij} < 4.4$ Å for (A) lysozyme (696 nonexchangeable protons, 3685 distances) and (B) the DNA dodecamer (228 nonexchangeable protons, 778 distances).

For a methyl–non-methyl pair, $m = 3$, and for an intermethyl pair, $m = 9$. The effects due to angular averaging for the spectral density $J(\omega)$ as a result of methyl rotation[48,50] were neglected. However, this simplification is not of consequence for defining the accuracy of the matrix solution.

A direct measure of the error in the matrix solution is the deviation in σ_{ij} values. Nevertheless, it is useful to relate the simulation results to the distances needed for structure determination, therefore the σ_{ij} values obtained by the eigenvalue solution were converted to d_{ij} values (eq 5) and the accuracy reported in some cases as the deviations in d_{ij}^c.

Structure Models. Scheme I outlines the protocol of this simulation study implemented with the program MORASS. The volumes of diagonal and off-diagonal peaks were generated for a given t_m (eq 2) and the relaxation rate matrix $\Gamma(\sigma_{ij}^o)$ calculated from a set of protons and a single correlation time (eq 5). $\Gamma(\sigma_{ij}^c)$ was then calculated (eq 4) from the unaltered simulated NOESY intensities **V** (noise-free data) or after altering elements of **V** to mimic various types of experimental noise, and the agreement was checked between the calculated σ_{ij}^c and the actual σ_{ij}^o values or the corresponding d_{ij}^c and d_{ij}^o values.

NOESY data were generated for both the protein lysozyme (129 residues) and an oligonucleotide, $d(CGCGAATTCGCG)_2$. Lysozyme proton positions were built from geometric considerations and the heavy atom coordinates of the crystallographic structure (Handoll, H.; Phillips, D. C. personal communications). The full coordinate set was optimized by energy minimization with the program CHARMM.[51] Atomic positions for all atoms of the oligonucleotide were built with standard B-DNA geometry by using the program AMBER,[52] followed by energy minimization. Except as noted above for methyl rotation, the macromolecule is assumed rigid so that molecular motions are described with a single rotational correlation time. Because the effects from multispin relaxation depend on the geometric arrangement of the protons, the distribution for each molecule of interproton distances less than 4.4 Å is shown in Figure 1. In addition to the peak in both distributions at 1.75 Å corresponding to geminal protons, the lysozyme distribution (Figure 1A, 696 nonexchangeable protons, 3685 distances) is nearly uniform from 2.5 to 4.4 Å but with a somewhat larger number of distances near 2.5 and 3.0 Å, while the oligonucleotide distribution (Figure 1A, 228 nonexchangeable protons, 778 distances) is trimodal with peaks at 2.3, 3.0, and 3.8 Å. Inclusion of exchangeable protons does not change the distributions significantly. Since a distribution similar to Figure 1A was found with the smaller protein crambin, and since distributions from several different DNA sequences were similar to Figure 1B, the distributions in Figure 1 appear to represent proteins and DNA in general.

(47) Noggle, J. H.; Schirmer, R. E. *The Nuclear Overhauser Effect*; Academic Press: New York, 1971.
(48) Olejniczak, E. T.; Dobson, C. M.; Karplus, M.; Levy, R. M. *J. Am. Chem. Soc.* **1984**, *106*, 1923–1930.
(49) Torchia, D. A. *Ann. Rev. Biophys. Bioeng.* **1984**, *13*, 125–144.
(50) Tropp, J. *J. Chem. Phys.* **1980**, *72*, 6035–6043.
(51) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
(52) Weiner, P.; Kollman, P. *J. Comput. Chem.* **1981**, *2*, 287.

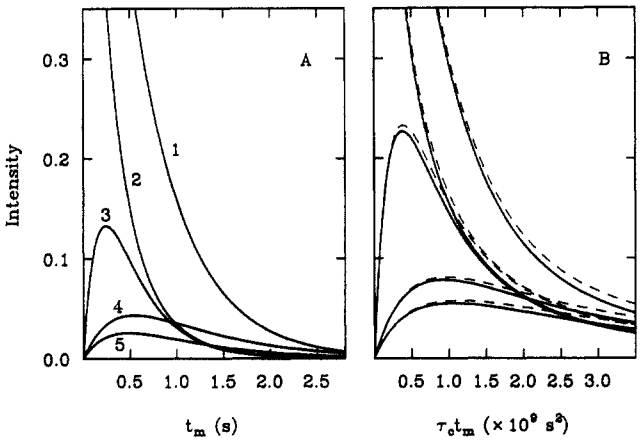


Figure 2. Time development of NOESY peaks calculated from the lysozyme coordinates for the diagonal peaks (1) Ser91 Hα and (2) Leu56 Hβ1 and the cross-peaks (3) Leu56 Hβ1-Hβ2, (4) Phe Hδ-Hϵ, and (5) Ser91 Hα-Hβ1. A rigid molecule rotation time was assumed: (A) τ_c = 1 ns and (B) τ_c = 4.5 ns (solid) and 9.0 ns (dashed). The abscissa of (B) is the generalized parameter $\tau_c t_m$.

The accuracy of the eigenvalue solution was determined as a function of mixing time for a field strength of 500 MHz and rotational correlation times equal to 1.0, 4.5, and 9.0 ns, typical for the size of molecules studied by NMR. Either lysozyme or the dodecamer proton coordinates were used for all values of τ_c studied. A single model for a protein or nucleic acid while varying τ_c had the advantage of a constant proton set with respect to spatial distribution. Given the imposed variation in τ_c without explicitly altering the macromolecular size, it is appropriate to consider only interior protons and not include surface protons. As such, the simulation of NOESY data and calculation of Γ included protons falling within a certain size sphere placed near the center of the molecule, while the errors in σ_{ij} and d_{ij} were evaluated for the protons within a smaller sphere. To facilitate the large number of matrix evaluations required for this study (computer processing times are given below), the size of the spheres was chosen to include fewer than 200 spins. For lysozyme, a 10-Å sphere centered at Leu56 Hα gave 194 protons for the data simulation and Γ calculation, and a 9-Å sphere gave 125 protons in the error analysis. For the oligonucleotide, a 19.2- and 15.4-Å sphere centered at A6 H1′ gave 168 and 125 protons, respectively.

The simulation and analysis of NOESY data were carried out with MORASS. Matrix diagonalization routines from the IMSL and EISPACK libraries were used and found to have comparable convergence properties. For the analysis involving noise-free data, double-precision operations allowed convergence of the matrix diagonalization in some instances where single-precision operations failed. However, for the more realistic case of imperfect data, double-precision operations are not warranted. The processing time on a MicroVax III workstation required for a 194-proton system was 4 min, while 496 protons took 1.9 h.

III. Results

We first examine the reliability of the eigenvalue solution as a function of t_m and τ_c with noise-free data simulated from lysozyme and the DNA dodecamer. It is demonstrated that in the large molecule limit ($\omega\tau_c \gg 1$) the results obtained with a given τ_c value can be related to other τ_c by the appropriate scaling of t_m. Since this limit is nearly valid for most molecules to which the NMR method for structure determination is applied, the results from simulations including experimental error described in the remainder of Results are reported for a single value of τ_c = 4.5 ns, from which the behavior for other τ_c can be obtained. In addition, test calculations on the DNA oligomer showed essentially identical behavior as lysozyme, thus results using imperfect data are shown for lysozyme only.

Noise-Free Data. The accuracy with which Γ can be evaluated by a multispin matrix approach with noise-free NOESY data was determined by using cross-peak volumes obtained directly from

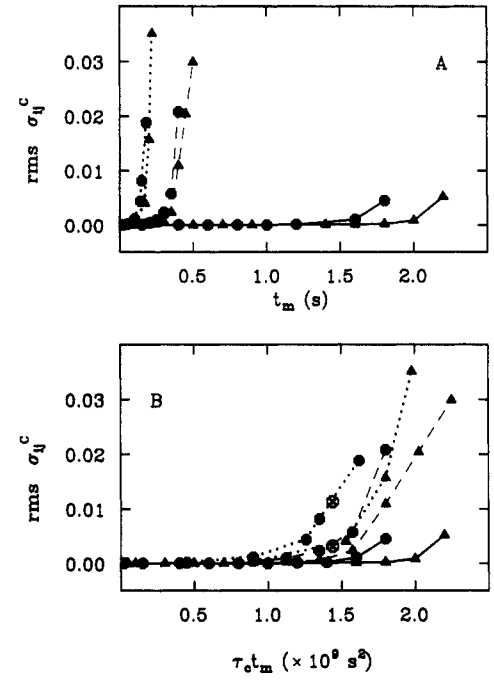


Figure 3. The rms error in σ_{ij}^c for the matrix method with noise-free data is plotted as a function of t_m (A) and $\tau_c t_m$ (B). Simulations were done with τ_c = 1.0 (solid), 4.5 (dashed), and 9.0 ms (dotted). The model systems are lysozyme (●) and the dodecamer (▲).

σ_{ij} and ρ_i calculated from the structures. An illustration of the time development of diagonal and off-diagonal peak volumes generated from the coordinates of lysozyme for τ_c = 1 ns is shown in Figure 2A and for τ_c = 4.5 and 9.0 ns in Figure 2B as a function of $\tau_c t_m$. The motivation for $\tau_c t_m$ as the independent parameter comes from the fact that both σ_{ij} and ρ_i are proportional to τ_c in the large molecule limit of $\omega\tau_c \gg 1$, i.e., consider eq 2 in the case where $\Gamma \propto \tau_c$. In a 500-MHz field, a τ_c value of 4.5 ns approaches the large molecule limit and the curves plotted as a function of $\tau_c t_m$ in Figure 2B are nearly superimposable. For $\tau_c > 9$ ns, the build-up intensities as a function of $\tau_c t_m$ are identical with the dashed curves (results not shown).

With v_{ij} elements simulated at a given t_m for lysozyme and the dodecamer, cross-relaxation rates were evaluated from eq 4. The rms deviations between the calculated σ_{ij}^c and the actual σ_{ij}^o used to simulate the NOESY data are shown in Figure 3A for τ_c = 1.0 (solid), 4.5 (dashed), and 9.0 ns (dotted). The matrix solution applied to noise-free data is quite accurate up to a certain t_m; there is no error in the calculated σ_{ij}^c values at short t_m. Only for t_m values beyond the maximum in the cross-peak intensities and where the diagonal intensities are greatly reduced is there significant error in σ_{ij}^c. Eventually as t_m increases, the diagonalization algorithm fails to converge, even for double-precision operations. The mixing time at which the σ_{ij} errors become significant occurs at shorter t_m for longer correlation times and is similar for proteins and DNA (Figure 3A). Given the limiting behavior for long τ_c in the build-up curves (see Figure 2B), a better parameter for indicating the onset of ill behavior in the eigenvalue solution is $\tau_c t_m$. The deviations in σ_{ij}^c are therefore replotted in Figure 3B, which shows that the critical value is $\tau_c t_m \sim 1.5 \times 10^{-9}$ s^2, or $t_m \sim 0.33$ s for τ_c = 4.5 ns, or $t_m \sim 0.17$ s for τ_c = 9.0 ns. These mixing times are substantially longer than the regime where the two-spin approximation is valid.

Random Noise. Random noise from a Gaussian distribution was added to all peak volumes in V to simulate both a constant low-level thermal noise and a peak-integration error. A level of constant noise plus noise proportional to the individual volume required generating two random numbers for each v_{ij} element. One random number specified a value ±0.1 to ±0.5% of v_0 added to all v_{ij} independent of its magnitude, while a second random number specified a value ±1 to ±4% of the particular v_{ij}. Several sets of random noise (typically 10) were added to a given V(t_m),

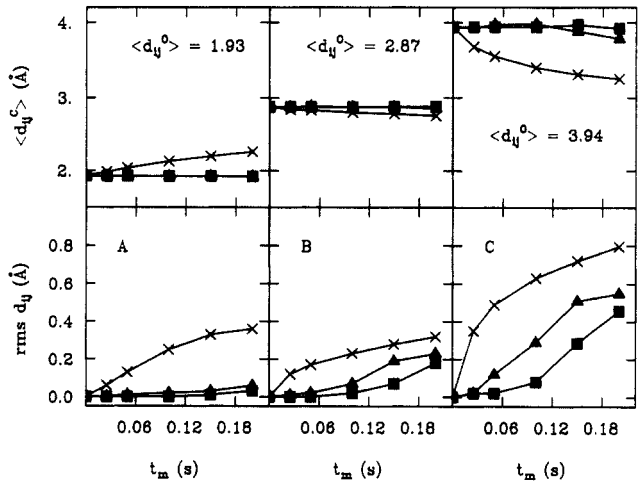


Figure 4. The average $\langle d_{ij}^c \rangle$ and the rms deviation between the calculated d_{ij}^c and the actual d_{ij}^0 values when random noise is added to the simulated data from the eigenvalue method (■, ▲) (eq 4) or the two-spin ratio approximation (×) (eq 6). Distances are grouped as (A) $d_{ij} < 2.4$, (B) $2.4 \le d_{ij} < 3.4$, and (C) $3.4 \le d_{ij} < 4.4$. Random noise was added to the simulated NOESY data (τ_c = 4.5 ns) at a level of 2% (■) or 4% v_{ij} (▲), simulating integration error. A level of 0.1% v_0, simulating thermal noise, was also added to all v_{ij}. Only one set of curves is shown for the two-spin values since the results with both noise levels are nearly identical.

and the rms deviations in σ_{ij} and d_{ij} calculated from each set were averaged to obtain statistically converged rms values.

The accuracy of the eigenvalue solution for Γ calculated from NOESY data containing random noise is shown in Figure 4 using interproton distances obtained from σ_{ij}^c (eq 5). The rms deviation between the actual distances d_{ij}^0 and the calculated distances d_{ij}^c as well as the calculated average distance $\langle d_{ij}^c \rangle$ are shown for three groups of interproton distances in lysozyme defined by the crystallographic structure. Different noise levels, indicated in the figure caption, were examined. For comparison, d_{ij} values are also determined from the same volume elements by the linear two-spin approximation:

$$\frac{v_{ij}}{v_{kl}} = \frac{\sigma_{ij}}{\sigma_{kl}} = \frac{d_{kl}^6}{d_{ij}^6} \qquad (6)$$

where the distance d_{kl} is assumed known. The reference distance was 2.52 Å for Leu56 Hα-Hβ of lysozyme. [In practice, the reference distance should be from a fixed-distance proton pair, not a variable distance pair such as Hα-Hβ. For the simulation study it is inconsequential since all distances are rigidly fixed. The value 2.52 Å is approximately equal to the fixed distance of a Tyr or Phe Hδ-Hε pair.] The average $\langle d_{ij}^c \rangle$ and rms deviation for the two-spin distances are essentially independent of noise level, thus the error with the two-spin method is represented by one curve in Figure 4.

At low noise levels the accuracy of the matrix method is high but diminishes as the noise level increases. With overall noise levels greater than 10%, the matrix diagonalization fails to converge. The eigenvalue method is more sensitive to the constant thermal noise than to the integration error; when the constant noise level is near 1% v_0, poor results are obtained with the eigenvalue method. Nonetheless, with a noise level of 3–4% v_{ij} integration error and 0.5% v_0 thermal noise, the eigenvalue solution gives lower rms deviations in interproton distances than the two-spin ratio method. [The rms error in d_{ij} with the two-spin method depends on the reference distance. As reported by others,[53] the error is smallest for distances nearly equal to that of the reference proton pair. As such, choosing a reference proton pair of shorter distance would decrease the two-spin rms error for small d_{ij} values but increase the error at large d_{ij}.] Moreover, the eigenvalue solution gives an accurate average interproton distance, as expected when the error in the solution is random. In contrast, the two-spin results show a bias toward long d_{ij} for distances less than the reference distance, and toward short d_{ij} for distances greater than the

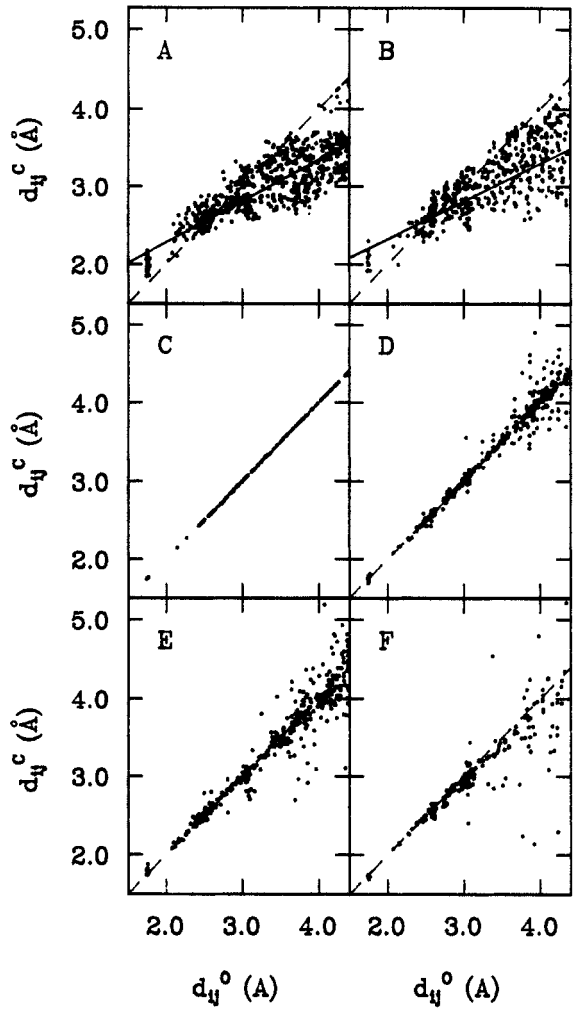


Figure 5. Calculated interproton distances, d_{ij}^c (Å), plotted against the actual value, d_{ij}^0. Distances were calculated from σ_{ij}^c values obtained with the eigenvalue method (C to F) or from a two-spin analysis employing ratios of the cross-peak volumes (A and B). NOESY intensities were simulated as noise-free (A and C), data with 3% random noise plus 0.1% thermal noise added (B and D), data with overlapping peaks where 25% of the cross-peaks had 12% noise error (E), and incomplete data from which intensities less than 1% v_0 (signal to noise = 100) were eliminated (F). All distributions are from t_m = 200 ms data. The best fit line (solid) of each distribution is shown. The dashed line is unity slope.

reference distance.[53] This bias is evident in a plot of d_{ij}^c against d_{ij}^0 (Figure 5); the slope of the best fit line for the distribution from the two-spin analysis is approximately 0.5 (Figure 5A,B), significantly less than the expected value of unity, while that from the eigenvalue analysis equals 1.0 (Figure 5C,D). In Figure 5 panels A and C were obtained with noise-free data and B and D with 3% random noise at t_m = 200 ms, a mixing time with large rms deviations. Similar plots at t_m = 50 ms are shown in Figure 6. Even with noise-free data the two-spin approximation has a substantial error and bias in $\langle d_{ij}^c \rangle$ (Figures 5A and 6A).

Peak Overlap. A major experimental difficulty in measuring NOESY intensities from large molecules is that there is substantial overlap of peaks, particularly for diagonal peaks. The problem in quantifying volumes of overlapping peaks was simulated in a fashion analogous to low-level random noise except that a higher level of noise was imposed on a fraction of the NOESY intensities calculated from lysozyme. The effect of peak overlap on the accuracy of the eigenvalue solution was studied for diagonal and off-diagonal peaks independently and in combination.

Overlap of diagonal peaks was examined by replacing the simulated value of v_{ii} with either the average diagonal volume, $\langle v_{ii} \rangle$, or by adding a large random noise to each v_{ii}. The former

(53) Clore, G. M.; Gronenborn, A. M. *J. Magn. Reson.* **1985**, *61*, 158–164.

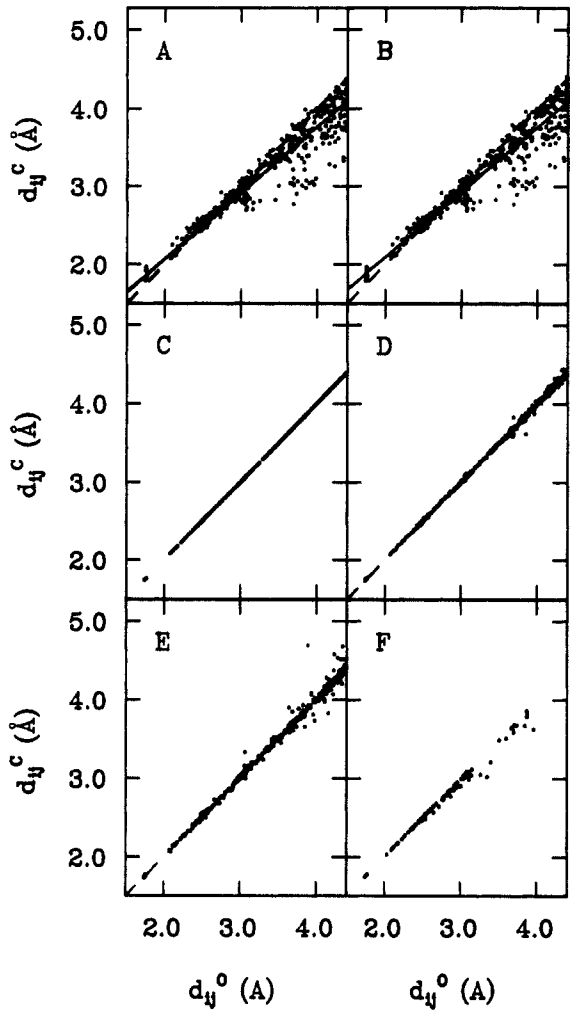


Figure 6. Same as Figure 5 except for t_m = 50 ms data.

case could apply to studies of proteins and higher molecular weight oligonucleotides for which few diagonal resonances are resolved, while the latter case is relevant to smaller molecules for which a greater number of diagonal peaks are resolved. The rms deviations in d_{ij} for the three groups of distances when v_{ii} is replaced by $\langle v_{ii} \rangle$ or when random noise of ±30% v_{ii} is added to individual v_{ii} are plotted in the upper panels of Figure 7. Even with this large uncertainty in diagonal peak volumes, the rms deviations are comparable or smaller than those of the two-spin approximation.

To mimic cross-peak overlap, noise of ±12% v_{ij} was added to 25% of the off-diagonal volumes chosen at random from V as well as ±3% random noise added to all other v_{ij}. Scatter plots of d_{ij}^c vs d_{ij}^0 are shown for t_m = 200 and 50 ms in Figures 5E and 6E, respectively. At these noise levels the rms deviations in d_{ij} are smaller for the matrix solution (Figure 7, ♦) than for the two-spin approximation (Figure 7, ×). Similar accuracy is obtained with the matrix solution when a larger fraction of peaks are overlapped if the error in v_{ij} is smaller; if 50% of the peaks include ±6% error and the lower ±3% error is added to the remainder of the peaks, the deviations are similar to those in Figure 7 (♦). When random noise was added at a higher level or to a greater percentage of the cross-peak volumes, the matrix diagonalization failed to converge or often returned with negative eigenvalues, physically unreal values, and the accuracy of the solution was poor.

Finally, the combined effect of uncertainity due to overlap of diagonal and off-diagonal peak intensities is also shown in Figure 7 (■). Diagonal peaks of V were replaced by $\langle v_{ii} \rangle$, and the same noise was added to v_{ij} as for offdiagonal overlap alone (♦). The results for the rms deviations show that errors upon combining the two uncertainties are not additive; the deviations with overlap error in both v_{ii} and v_{ij} are only slightly larger than those found with noise in v_{ij} alone for longer d_{ij}. At shorter d_{ij}, the deviations

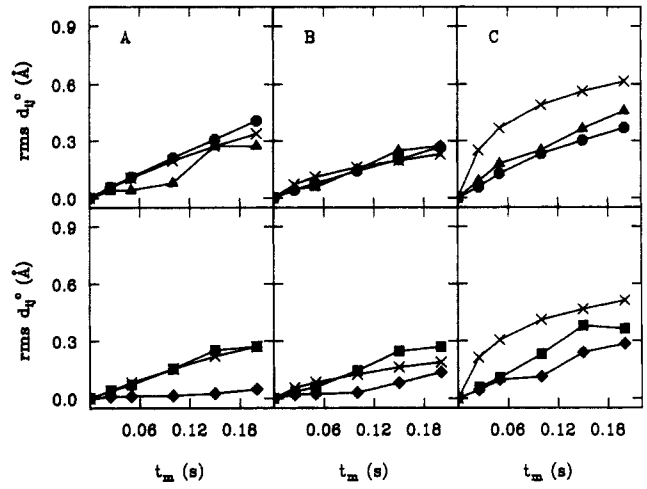


Figure 7. Rms deviations between the calculated d_{ij}^c and the actual d_{ij}^0 when accounting for overlapping peaks for distance groups (A) d_{ij} < 2.4, (B) 2.4 ≤ d_{ij} < 3.4, and (C) 3.4 ≤ d_{ij} < 4.4. The results from the two-spin ratio (×) are compared with those of the eigenvalue solution: upper panels, diagonal intensities replaced with $\langle v_{ii} \rangle$ (●) or ±30% v_{ii} random noise added (▲); lower panels, off-diagonal intensities overlap simulated by adding high level random noise (±12% v_{ij}) to 25% of the off-diagonal intensities and low-level noise (±3.0% v_{ij}) plus 0.1% v_0 thermal noise to other intensities (♦) and combined effects when the diagonal is replaced with $\langle v_{ii} \rangle$ plus the same off-diagonal overlap noise (■).

are the result of diagonal peak uncertainty alone.

Incomplete NOESY Data. In practice, it is possible to obtain only a partial volume matrix as defined by the inherent sensitivity of the NMR spectrometer; small cross-peak volumes corresponding to distances greater than 4.5–5.0 Å (depending on τ_c) cannot be accurately measured. To examine the effects of incomplete data on the reliability of the eigenvalue method, elements of V less than a specified cutoff volume were replaced with a defined volume, v_r. In most instances v_r was set equal to 0.0. However, better convergence behavior was sometimes found when v_r was set equal to a small value determined empirically and typically less than 0.2% V_0. Hence, using a low level background in V by artificially setting all weak cross-peak volumes to a small value, even though many of these v_{ij} elements are actually near zero, improved the convergence behavior, while artificially setting all small v_{ij} to zero sometimes lead to ill-conditioned eigensystems. Moreover, when a small value was used for v_r, the solution for Γ generally improved by a 0.1- to 0.2-Å decrease in the rms deviation of d_{ij}. The apparent optimum value for v_r is approximately one-half the cutoff volume, i.e., one-half the noise level.

The error in d_{ij}^c values from incomplete V matrices was determined as a function of mixing time for several cutoff values or signal-to-noise levels. The results shown in Table I and Figure 8 are the rms deviations in d_{ij} when v_{ij} < 0.2, 0.5, 1.0, or 2% of the equilibrium magnetization v_0 were replaced with v_r. NOESY data measured at a 500-MHz field strength illustrating cross-peak volumes of 1% have been reported.[17,26] Thus the range in cutoff values simulates sensitivity limits which are achievable in practice. The completeness of the NOESY data is indicated in Table I by the percentage of the total number of v_{ij} elements that were greater than the cutoff value. Only distances corresponding to these v_{ij} elements were included in the error analysis.

The trend of increasing error with t_m is also found with incomplete V as demonstrated by Figure 8. For all levels of sensitivity examined, the deviations in the eigenvalue solution for cross-relaxation rates for proton pairs separated by <3.4 Å is less than 0.35 Å, while the two-spin rms deviations for this group (Figures 8A,B) vary from 0.07 to 0.4 Å. Even when only 50% of the intensities for distances between 2.4 and 3.4 Å are included, as for short t_m with cutoff volumes of 1.0 and 2.0% v_0 (signal to noise of v_0 equal to 100 and 50, see Table I), the eigenvalue solution has a small rms error, substantially less than that from the two-spin ratio method. In the distribution of d_{ij}^c against d_{ij}^0

Table I. Rms Error in d_{ij}^{x} Determined from Incomplete NOESY Data by the Matrix Method[a]

interproton distance		mixing time 25	50	100	150	200
		Cutoff = 0.2%, S/N = 500				
0–2.4	rms	0.00	0.00	0.00	0.00	0.01
	% v_{ij}[b]	100	100	100	100	100
2.4–3.4	rms	0.02	0.01	0.00	0.01	0.06
	% v_{ij}	100	100	100	100	98
3.4–4.4	rms	0.07	0.13	0.20	0.22	0.44
	% v_{ij}	58	91	99	94	93
		Cutoff = 0.5%, S/N = 200				
0–2.4	rms	0.00	0.00	0.01	0.01	0.02
	% v_{ij}	100	100	100	100	100
2.4–3.4	rms	0.03	0.07	0.08	0.09	0.14
	% v_{ij}	86	100	100	98	98
3.4–4.4	rms	0.09	0.21	0.37	0.62	0.51
	% v_{ij}	6	44	80	86	85
		Cutoff = 1.0%, S/N = 100				
0–2.4	rms	0.01	0.01	0.02	0.03	0.03
	% v_{ij}	100	100	100	100	100
2.4–3.4	rms	0.03	0.09	0.19	0.23	0.22
	% v_{ij}	50	81	95	94	98
3.4–4.4	rms		0.26	0.48	0.63	0.73
	% v_{ij}	0	9	41	51	63
		Cutoff = 2.0%, S/N = 50				
0–2.4	rms	0.01	0.04	0.09	0.30[c]	0.48[c]
	% v_{ij}	100	100	100	94	90
2.4–3.4	rms	0.03	0.09	0.27	0.33[c]	0.36[c]
	% v_{ij}	20	46	76	73	74
3.4–4.4	rms			0.65	0.99[c]	1.08[c]
	% v_{ij}	0	0	12	16	20

[a] V was simulated with τ_c = 4.5 ns from lysozyme proton coordinates. Volume elements smaller than the cutoff value, expressed as a percentage of v_o, were replaced with v_r (see text). When all volume elements are included the rms error is 0.0 for all groups for $t_m \leq 250$ ms. [b] Percentage of distances with v_{ij} intensity greater than the cutoff volume. Total number of interproton distances for $d_{ij} < 2.4$ Å is 91, for $2.4 \leq d_{ij} < 3.4$ Å is 253 and for $3.4 \leq d_{ij} < 4.4$ Å is 349. [c] Matrix diagonalization resulted in negative eigenvalues.

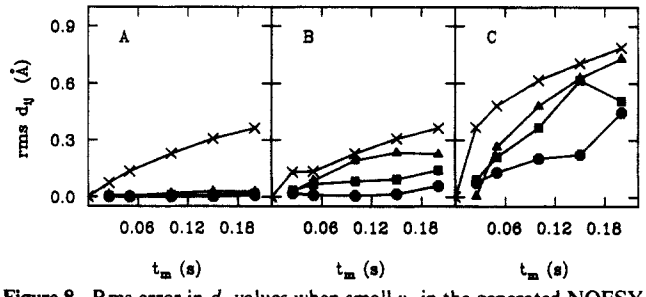


Figure 8. Rms error in d_{ij} values when small v_{ij} in the generated NOESY data are replaced with a constant v_r value, simulating incomplete experimental data. The cutoffs for v_{ij} are 1.0% (▲), 0.5% (■), and 0.2% (●) of the diagonal volume at t_m = 0, corresponding to a signal to noise of 100, 200, and 500, respectively. Distances are grouped as (A) $d_{ij} <$ 2.4, (B) $2.4 \leq d_{ij} < 3.4$, and (C) $3.4 \leq d_{ij} < 4.4$. The values for a two-spin analysis (×) are shown for comparison.

(Figure 5F and 6F) there is a high density of points along the line of unity slope with the eigenvalue solution although a small number of points lie far from the line in the case of the longer t_m (Figure 5F). With regard to a structure determination, such large deviations could have an adverse effect. Large errors of this nature should become evident however from a comparison of σ_{ij} and d_{ij} values evaluated at distinctly different t_m.

IV. Conclusions

As interproton distances are the fundamental basis of three-dimensional structure determination by NMR, it is important to consider methods which improve the accuracy of their evaluation.[6,21] Due to the error[22-28] in the two-spin approximation procedure for evaluating interproton distances from NOESY intensities, we have studied the feasibility of a rate matrix approach for this task. In this paper we define the practical limitations in the matrix method due to numerical error in diagonalization of a matrix of NOESY intensities when effects from experimental limitations on their measurement are present. The simulation studies show that the matrix solution is accurate when the input spectral intensities are well determined. With good quantification of NOESY volumes and complete data sets including volumes equal to 1.0% of the diagonal intensity at t_m = 0, the distances calculated with the matrix solution are more accurate than those obtained assuming the two-spin model. A relatively larger error in the quantification of a fraction of the peaks, as in overlapping resonances, can be tolerated. The need for a complete data set in the matrix analysis means that all possible peak volumes should be measured. Complete resonance assignments are not required. (Of course, the assignments as well as a dynamic model for τ_c are needed to relate the calculated σ_{ij} values to the restraints required in the structure determination.)

More accurate distances justify the use of tighter constraints in the conformational search by molecular dynamics or distance geometry. An advantage of tighter constraints is to achieve greater precision in the set of structures[12,54] obtained by the NMR method. [Although the finding of greater precision among structures when tighter restraints are used was obtained with studies on oligonucleotides in ref 12 and 54, it is also likely to hold for proteins since positional fluctuations are reduced in general with stronger force constants.] It is reasonable that better determined interproton distances would lead as well to improved accuracy in the three-dimensional structure, although such an improvement has not yet been rigorously demonstrated. Studies so directed are in progress in our laboratories. One study has indicated that the structure determination process behaves better when the distances result from a multispin analysis rather than from a two-spin analysis; Nikonowicz et al.[55,56] have used NMR distances to derive a family of structures for the extra-helical adenosine tridecamer d(CGCAGAATTCGCG)$_2$. By using distances from the hybrid relaxation matrix methodology in combination with restrained molecular dynamics and starting from two quite different initial structures, it was possible to iteratively refine both to a common family of structures consistent with the NOESY-derived distances. On the other hand, differentiating between structures in which the extra-helical adenosine stacks between the third and fourth base pairs forms H-bonds to a G–C base pair to form a triplet, or even stacks between the second and third base pairs was not possible by using distances derived from the two-spin approximation. Only the theoretical NOESY spectra derived from the hybrid matrix/restrained MD structures are consistent with the experimental NOESY spectra.

Most of our simulation studies were carried out with a matrix of 194 spins from the interior of lysozyme. Similar dependences on mixing time and sensitivity to imperfect NOESY data in the accuracy of the matrix solution were found for the DNA oligomer. Because of the limiting behavior when τ is large (see Results section on noise-free data) our results as a function of t_m at τ_c = 4.5 ns can be extrapolated to other τ_c values. The larger matrix size needed to include the full set of 696 lysozyme protons (approximately a 20-Å diameter sphere) would require significantly more computer time and would have poorer convergence behavior (data not shown). Therefore to evaluate distances for a molecule with greater than ≈200 protons, a procedure that divides the molecule into overlapping regions and uses more than one matrix to evaluate the distances might be useful.

The matrix method will benefit from increased sensitivity and signal to noise of high field spectrometers and from improvements in software for volume integration of NOESY peaks.[57,58] With

(54) Gronenborn, A. M.; Clore, G. M. *Biochemistry* **1989**, *28*, 5978–5984.
(55) Nikonowicz, E.; Meadows, R. P.; Gorenstein, D. G. *Bull. Magn. Reson.* **1989**, *11*, 226–229.
(56) Nikonowicz, E.; Meadows, R. P.; Gorenstein, D. G. *Biochemistry* **1990**, *29*, 4193–4204.
(57) Denk, W.; Baumann, R.; Wagner, G. *J. Magn. Reson.* **1986**, *67*, 386–390.

noise-free data, essentially exact solutions are obtained when $\tau_c t_m < 1.5 \times 10^{-9}$ s^2. In contrast, the accuracy of d_{ij} values obtained by using a two-spin approximation is not increased by improvements in the data because of the inherent error in neglecting multispin relaxation effects.

Procedures involving a matrix analysis are being actively investigated. Although the requirement for nearly complete data with well-determined intensities may prohibit a multispin analysis of interproton distances in the first steps of structure determination by NMR, such an analysis is beneficial in refinement of the structural solution. The need for nearly complete NOESY data sets can be circumvented in the first stages of structure determination by combining experimental data with cross-peak intensities calculated from an initial model.[30,37,55,59] An iterative procedure could also serve to improve the quantification of NOESY intensities for instances such as overlapping resonances. Furthermore, a novel approach utilizing NOESY data directly in the conformational search, in contrast to the interproton distances interpreted from them, was recently reported.[60] At the least, the final structure should be evaluated by comparing simulated NOESY cross-peak volumes calculated from the structural solution with the experimentally observed NOESY volumes, perhaps in a fashion analogous to the crystallographic *R*-factor[22,61,62] or as an average percent difference,[59] with some provision for possible discrepancies resulting from internal motions.

Acknowledgment. We gratefully acknowledge many useful discussions with Drs. E. T. Olejniczak and C. R. Jones and the help of J. Fish in the initial stages of the development of MORASS. The work was supported in part by the National Institutes of Health (Grant RR01077) and the National Science Foundation (Grant BBS 8614177).

Registry No. d(CGCGAATTCGCG), 128496-41-3; lysozyme, 9001-63-2.

(58) Olejniczak, E. T.; Gampe, R. T., Jr.; Fesik, S. W. *J. Magn. Reson.* **1989**, *81*, 178-185.
(59) Gorenstein, D. G. G.; Meadows, R. P.; Metz, J. T.; Nikonowicz, E.; Post, C. B. In *Advances in Biophysical Chemistry*; Bush, A., Ed.; J.A.I. Press: In press.
(60) Yip, P.; Case, D. A. *J. Magn. Reson.* **1989**, *83*, 643-648.
(61) Summers, M. F.; South, T. L.; Kim, B.; Hare, D. R. *Biochemistry* **1990**, *29*, 329-340.
(62) Baleja, J. D.; Moult, J.; Sykes, B. D. *J. Magn. Reson.* **1990**, *87*, 375-384.

Nature of the Carbon-Phosphorus Double Bond and the Carbon-Phosphorus Triple Bond As Studied by Solid-State NMR

James C. Duchamp,[1a] Marek Pakulski,[1b] Alan H. Cowley,[1b] and Kurt W. Zilm*,[1a]

Contribution from the Department of Chemistry, Yale University, New Haven, Connecticut 06511, and Department of Chemistry, University of Texas at Austin, Austin, Texas 78712.
Received January 11, 1990

Abstract: The nature of the carbon-phosphorus double bond in 2,4,6-*t*-$Bu_3C_6H_2P{=}C(SiMe_3)_2$ and the carbon-phosphorus triple bond in 2,4,6-*t*-$Bu_3C_6H_2C{\equiv}P$ has been studied by ^{13}C and ^{31}P solid-state NMR. Magic angle spinning and static cross-polarization experiments have been used to determine the principal elements of the ^{13}C and ^{31}P shielding tensors. In the ^{13}C spectra, the presence of a dipolar coupling to the ^{31}P nucleus permits assignment of the orientation of the ^{13}C shielding tensors in the molecular frame. These shift tensors are compared to previous work on diphosphenes, disilenes, alkenes, and alkynes. It is found that the shift anisotropies for ^{31}P and ^{13}C in these multiply bonded environments are quite similar when the larger intrinsic chemical shift range for ^{31}P is taken into account.

Introduction

Advances in synthetic chemistry over the last two decades have resulted in the successful synthesis and isolation of a variety of main-group compounds containing multiply bonded functional groups.[2] Examples of room temperature stable species with Si=Si,[3] P=P,[4] and Sn=Sn[5] double bonds have been realized, as well as compounds with multiple bonds between main-group and transition elements. A number of mixed-main-group systems with, e.g., C=Si,[6] C=P,[7] and C≡P,[7] multiple bonds are also now known. In this paper we report the first use of solid-state NMR methods to investigate the nature of the C=P double bond in 2-(2,4,6-tri-*tert*-butylphenyl)-1,1-bis(trimethylsilyl)phosphaethene (**1**) and the C≡P triple bond in 2-(2,4,6-tri-*tert*-butylphenyl)-phosphaacetylene (**2**). Previous efforts to characterize these novel compounds have involved studies of their chemical reactivity or used structural methods in addition to theoretical treatments. Reactivity of **1**, **2**, and related compounds has been found to compare favorably with that of similar alkenes and alkynes.[7,8] X-ray crystallography finds the C=P bond length in **1** to be 1.665

(1) (a) Yale University. (b) University of Texas at Austin.
(2) (a) Cowley, A. H. *Polyhedron* **1984**, *3*, 389-432. (b) Cowley, A. H. *Acc. Chem. Res.* **1984**, *17*, 386-392. (c) West, R. *Pure Appl. Chem.* **1984**, *56*, 163. (d) Raabe, G.; Michl, J. *Chem. Rev.* **1985**, *85*, 419. (e) Cowley, A. H.; Norman, N. C. *Prog. Inorg. Chem.* **1986**, *34*, 1. (f) Regitz, M.; Binger, P. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 1484.
(3) West, R.; Fink, M. J.; Michl, J. *Science (Washington, D.C.)* **1981**, *214*, 1343.
(4) Yoshifuju, M.; Shima, I.; Inamoto, N. *J. Am. Chem. Soc.* **1981**, *103*, 4587, 4589.
(5) Davidson, P. J.; Harris, D. H.; Lappert, M. F. *J. Chem. Soc., Dalton Trans.* **1976**, 2286.
(6) Brook, A. G.; Abdesaken, F.; Gutekunst, G.; Kallury, R. K. *J. Chem. Soc., Chem. Commun.* **1981**, 191.
(7) (a) Appel, R.; Knoll, F.; Ruppert, I. *Angew. Chem., Int. Ed. Engl.* **1981**, *20*, 731-744. (b) Kroto, H. W. *Chem. Soc. Rev.* **1982**, *11*, 435.
(8) (a) Schäfer, A.; Weidenbruch, M.; Saak, W.; Pohl, S. *Angew. Chem., Lett. Ed. Engl.* **1987**, *26*, 776-777. (b) Annen, U.; Regitz, M. *Tetrahedron Lett.* **1987**, *28*, 5141-5144. (c) Quin, L. D.; Hughes, A. N.; Pete, B. *Tetrahedron Lett.* **1987**, *28*, 5783-5786. (d) Xie, Z.; Wislan-Neilson, P.; Neilson, R. *Organometallics* **1985**, *4*, 339-344. (e) Becker, G.; Becker, W.; Knebl, R.; Schmidt, H.; Weeber, U.; Westerhausen, M. *Nova Acta Leopold.* **1985**, *264*, 55-67.

0002-7863/90/1512-6803$02.50/0