JOURNAL OF MAGNETIC RESONANCE **124,** 177–188 (1997)
ARTICLE NO. MN961033

# Relax, a Flexible Program for the Back Calculation of NOESY Spectra Based on Complete-Relaxation-Matrix Formalism

ADRIAN GÖRLER AND HANS ROBERT KALBITZER*

*Department of Biophysics, Max-Planck-Institute for Medical Research, Jahnstrasse 29, D-69120 Heidelberg, Federal Republic of Germany*

Received September 13, 1996; revised October 9, 1996

**RELAX is a flexible program for the quantitative analysis of NOESY spectra. It allows the simultaneous application of different models describing the internal and overall motion of the molecule under investigation for individual spin pairs or groups of spins. A correction for anisotropy effects due to the deviation of the molecule from a spherical shape is calculated automatically from the trial structure. The program can deal with completely relaxed spectra as well as spectra recorded with a short relaxation delay. An execution-time-controlled splitting of the relaxation matrix reduces the computation time significantly without any loss of accuracy. This is especially important for large molecules or medium distance cutoffs.** © 1997 Academic Press

## INTRODUCTION

The quantitative analysis of nuclear Overhauser effect spectra allows the determination of interatomic distances. In order to obtain a highly resolved NMR structure, a large number of distance estimates with high accuracy is essential (*1*). Commonly distances are calculated using the isolated-spin-pair approximation (ISPA). For short mixing times it relates the cross-peak volumes $V_{ij}$ to the distances $r_{ij}$ between two atoms $i$ and $j$ by (*2*)

$$V_{ij} = \alpha r_{ij}^{-6}. \quad [1]$$

Usually, the constant $\alpha$ is calculated from the cross-peak volume of some spins with known distance. For longer mixing times, distances are often calculated from the initial slope of the build-up curve of the NOE (*3*). The validity of Eq. [1] is based on three fundamental assumptions: spin pairs can be treated as isolated, the molecule is rigid, and the rotational diffusion of the molecule is isotropic. This approximation was sufficient in the beginning of the structure determination of macromolecules from NMR data. However, with the availability of sensitive high-field spectrometers and the accompanying general improvement of spectral quality, Eq. [1] is a too-crude approximation. A much superior method for evaluating NOE data is the complete-relaxation-matrix analysis. In this approach, all dipolar interacting spins are treated as a network and the volumes of the NOE cross peaks are calculated by exponentiating the matrix of cross-relaxation rates—the relaxation matrix. This approach not only allows the calculation of NOE volumes for long mixing times but additionally makes it possible to describe the internal motions of the molecule in a detailed way.

During the past years, a number of groups have presented computer programs [CORMA (*4*), BCKCALC (*5*), X-PLOR (*6*), and BIRDER (*7*)] that are capable of calculating NOESY cross-peak volumes by complete-relaxation-matrix analysis. The first calculations have been performed mostly for theoretical reasons (*8, 9*). Other programs, such as MARDIGRAS (*10, 11*), NO2DI (*12*), and a program from Kim and Reid (*13*), use relaxation-matrix analysis to calculate interatomic distances from NOESY cross-peak volumes. In IRMA (*14, 15*) and DINOSAUR (*16*), a relaxation-matrix analysis is applied iteratively with molecular-dynamics calculations for the refinement of the structure. The direct comparison of simulated and experimental NOE values allows the determination of a measure for the quality of a structure analogous to the X-ray $R$ factor which is in principle superior to the often-applied rmsd analysis (*17, 18*) of the set of structures obtained. Such a measure can be used to replace distance constraints in molecular-dynamics or distance-geometry calculations (*19–21*) by incorporating it as a pseudo-energy. A current development is the use of calculated NOESY cross peaks for the automatic search and assignment of experimental peaks, as done in the programs ASNO (*22*) and NOAH (*23*).

The programs mentioned above have different restrictions and limitations. Some (BIRDER) treat the molecule under investigation as rigid. Others allow the description of aromatic ring flips and the rotation of methyl groups by distance averaging (X-PLOR) or discrete jump models (IRMA) including a free rotor model (MARDIGRAS). In X-PLOR, internal motions can additionally be described according to the so-called model-free approach presented by Lipari and Szabo (*24, 25*). Like BIRDER, a recent version of CORMA

* To whom correspondence should be addressed.



1090-7807/97 $25.00

is capable of calculating NOE cross peaks with a finite relaxation delay. A program called SYMM (*26*) can be applied to remove the effects of a finite relaxation delay from experimental NOE cross-peak volumes. BIRDER can additionally treat the molecule as a rigid top with two different correlation times for the rotation about its symmetry axis and the orthogonal rotation. To our knowledge, X-PLOR so far is the only program which allows restricting the calculation to a subset of cross peaks, which is much faster. Most of the programs use matrix diagonalization to calculate NOESY cross peaks while BCKCALC, for example, applies numerical integration.

Because of the growing importance of relaxation-matrix analysis, we present here a newly developed set of routines that allow the simulation of NOE cross peaks in a fast, flexible, and easy-to-handle way. They make it possible to calculate cross peaks by the application of a dynamic model which can be adapted to the molecule under investigation by combining various spectral densities. Calculations can be performed assuming a finite or infinite relaxation delay. The computation can be restricted to any subset of cross peaks. As a first basic application of these routines, the program RELAX is presented.

## MATERIAL AND METHODS

For best portability, the program has been written in ANSI-C. Currently, full-featured stand-alone versions of RELAX exist for 80 x 86, Bruker Aspect 1, SGI, and Convex C220. They are available from the authors upon request. For comfortable display of the results a simplified version of this program has already been implemented in the program AURELIA (*27*), which is available from Bruker (Karlsruhe).

Calculations have been performed for the 74-residue protein tendamistat from *Streptomyces tendae*, for the 87-residue protein HPr from *Staphylococcus aureus*, and for fragments of actin from rabbit skeletal muscle with sizes increasing in steps of 25 residues from 25 up to 372. The results presented in this paper were obtained on a SGI Indigo$^2$ Workstation (MIPS R4400, 200 MHz, 64 Mbyte RAM).

## THEORETICAL CONSIDERATIONS

In the following, those results of the theory of the relaxation-matrix formalism which are the basis of the algorithms presented in this paper will be recalled. It is assumed that spins that are magnetically equivalent are combined into groups $i$, $j$, ... with the corresponding number of group members $n_i$, $n_j$, .... The evolution of the deviation $\Delta\mathbf{M}_z = \mathbf{M}_z - \mathbf{M}_0$ of the longitudinal magnetization of the equilibrium magnetization $\mathbf{M}_0$ in an NOESY experiment can be described by the generalized Solomon equation (*28*)

$$\frac{d}{dt}\,\Delta\mathbf{M}_z(t) = -\mathbf{D}\Delta\mathbf{M}_z(t), \qquad [2]$$

where $\mathbf{D}$ is the dynamic matrix given by $\mathbf{D} = \mathbf{N}\cdot\mathbf{R}' + \mathbf{K}$. The kinetic matrix $\mathbf{K}$ describes conformational or chemical exchange (*21*) and $\mathbf{R}'$ is the symmetrized relaxation matrix (*29–31*). The diagonal matrix $\mathbf{N}$ contains the numbers of group members $n_i$ of each group $i$ as its elements. For dipolar homonuclear ($\gamma_i = \gamma_j$) relaxation and spin $I = \frac{1}{2}$, the auto-relaxation rates $R'_{ii}$ and the cross-relaxation rates $R'_{ij}$ are given by

$$\begin{aligned} R'_{ii} &= q(1 - 1/n_i)[3J^1_{ii}(\omega) + 12J^2_{ii}(2\omega)] \\ &\quad + q\sum_{j\neq i} n_j/n_i[J^0_{ij}(0) + 3J^1_{ij}(\omega) + 6J^2_{ij}(2\omega)], \\ R'_{ij} &= q[6J^2_{ij}(2\omega) - J^0_{ij}(0)]. \end{aligned} \qquad [3]$$

$J^n_{ij}$ ($n = 0, 1, 2$) are the spectral densities for $n$-quantum transitions characterizing the motion of spins $i$ and $j$, and $q = (1/10)\gamma_i^2\gamma_j^2\hbar^2(\mu_0/4\pi)^2$ is the dipolar interaction constant. In the current version of RELAX, chemical-exchange effects are neglected. Under this simplification Eq. [2] has the solution

$$\Delta\mathbf{M}_z(t) = \Delta\mathbf{M}_z(0)\exp(-t\mathbf{N}\cdot\mathbf{R}'). \qquad [4]$$

The NOE cross-peak volumes $\mathbf{V}(t)$ are (*29, 30*)

$$V_{ij}(t) = \alpha M_{z,j}(0)[\exp(-t\mathbf{N}\cdot\mathbf{R}')]_{ij} \qquad [5]$$

with an arbitrary scaling factor $\alpha$. Usually it is assumed that the system is completely relaxed prior to the first pulse of the NOESY sequence. With $M_{z,j}(0) = n_j/\alpha$, the normalized symmetric matrix of NOE intensities $\mathbf{V}(t)$ is

$$\mathbf{V}(t) = \mathbf{N}^{1/2}\cdot\mathbf{A}(t)\cdot\mathbf{N}^{1/2} \qquad [6]$$

with

$$\mathbf{A}(t) = \exp(-t\mathbf{N}^{1/2}\cdot\mathbf{R}'\cdot\mathbf{N}^{1/2}). \qquad [7]$$

However, fully relaxed NOESY spectra are hardly ever recorded. A much better signal-to-noise ratio can be achieved if more experiments are accumulated with a shortened relaxation delay $t_d$ which is also used for water suppression by selective saturation. In this case, the longitudinal magnetization $\mathbf{M}_{z,j}$ recovers only partly during the recovery time $t_r$ and is no longer proportional to $n_j$ prior to the first pulse. The spectral quality can be improved and $t_1$-dependent echo formation can be suppressed by destroying all longitudinal and transverse magnetization by the application of two orthogonal spin-lock pulses prior to the delay $t_d$. Since all magnetization is destroyed by the $\mathbf{B}_1$-field inhomogeneity, it can be assumed that $\mathbf{M}_z(0) = 0$ after the application of the spin-lock pulses and recovers only during the relaxation

delay $t_r = t_d$. In the absence of spin-lock pulses, the magnetization at the beginning of the mixing time is dependent on the $t_1$ increment. The two cases, however, are formally equivalent since under ideal conditions $\mathbf{M}_z(0) = 0$ is fulfilled just after the 90° detection pulse at $t_1 = 0$. For the cross-peak volumes only the magnetization at $t_1, t_2 = 0$ is important (as a general property of the Fourier transformation). Therefore, in the absence of the spin-lock pulses, there is an additional contribution from the acquisition time $t_{ac}$ to the total recovery time $t_r$ (*7*, *26*, *32*): $t_r = t_{ac} + t_d$.

The $z$ magnetization $\mathbf{M}_z(t_d)$ at the beginning of the mixing time ($t_1$, $t_2 = 0$) can be written as

$$M_{z,j}(t_r) = \frac{1}{\alpha} \sum_k [1 - \exp(-t_r \mathbf{N} \cdot \mathbf{R}')]_{jk} n_k \qquad [8]$$

and the matrix $\mathbf{V}(t)$ takes the form

$$V_{ij}(t_r, t) = \sum_{j,k} [\exp(-t\mathbf{N} \cdot \mathbf{R}')]_{ij} [1 - \exp(-t_r \mathbf{N} \cdot \mathbf{R}')]_{jk} n_k \qquad [9]$$

or in matrix notation,

$$\mathbf{V}(t_r, t) = \mathbf{N}^{1/2} \cdot \mathbf{A}(t) \cdot [\mathbf{A}(t_r)] \mathbf{N}^{1/2}. \qquad [10]$$

Note that the matrix $\mathbf{V}(t_r, t)$ and the corresponding NOESY spectrum are now asymmetric (*7*).

*Spectral Densities*

Equation [3] contains the spectral densities of $n$-quantum transitions $J^n_{ij}(\omega)$ of the motion of the vector $\mathbf{r}_{ij}$ that connects the nuclei $i$ and $j$ relative to the $\mathbf{B}_0$ field. Ideally, the spectral densities should be obtained from an analysis of this motion which is generally not known in detail. Fortunately, there are some special cases for which analytical expressions for the spectral-density functions exist. Those which are implemented in RELAX will be presented briefly in the following along with the mnemonic names used in the program RELAX.

In most applications of the relaxation-matrix formalism, the molecule is assumed to be rigid and diffusing isotropically in the solvent with an overall correlation time $\tau_c$. This case is described by the well-known spectral-density function (*33*)

RIGID:

$$J^n_{ij}(\omega) = \frac{1}{r^6_{ij}} \frac{\tau_c}{1 + \omega^2 \tau_c^2}. \qquad [11]$$

The assumption of a rigid molecule is usually not justified. While, for example, regions in the backbone of a peptide with well-defined secondary structure are relatively rigid, methyl groups or aromatic rings undergo rapid internal motions. The spectral density of the movement of the vector connecting spins which are involved in such a type of motion can be described by jump models representing a jump of a group of spins $i$ between $N_i$ equilibrium sites (*34–37*). Two special cases of jumps are easy to treat, for which the formulas are given in the most general case of two interacting jumping groups of spins $i$ and $j$. If the correlation time $\tau_{1,ij}$ of the jump motion is much slower than the overall rotational correlation time $\tau_c$ of the molecule, a slow-jump approximation can be made and the spectral density results in

SLOW_JUMP:

$$J^n_{ij}(\omega) = \frac{1}{N_i N_j} \left( \frac{\tau_c}{1 + \omega^2 \tau_c^2} \right) \sum_{\mu=1}^{N_i} \sum_{\nu=1}^{N_j} \frac{1}{r^6_{i_\mu j_\nu}}, \qquad [12]$$

which is often referred to as $r^{-6}$ averaging. In the case of fast jumps where $\tau_1 \gg \tau_c$, no simple averaging can be applied. In this approximation the equation for the spectral density is

FAST_JUMP:

$$J^n_{ij}(\omega) = \frac{1}{2N_i^2 N_j^2} \frac{\tau_c}{1 + \omega^2 \tau_c^2} \sum_{\mu,\nu=1}^{N_i} \sum_{o,\pi=1}^{N_j} \frac{1}{r^5_{i_\mu j_\nu} r^5_{i_o j_\pi}} [3(\mathbf{r}_{i_\mu j_\nu} \cdot \mathbf{r}_{i_o j_\pi})^2 - r^2_{i_\mu j_\nu} r^2_{i_o j_\pi}]. \qquad [13]$$

The so-called $r^{-3}$ averaging

AVERAGE_3:

$$J^n_{ij}(\omega) = \frac{1}{N_i^2 N_j^2} \frac{\tau_c}{1 + \omega^2 \tau_c^2} \left| \sum_{\mu=1}^{N_i} \sum_{\nu=1}^{N_j} \frac{1}{r^3_{i_\mu j_\nu}} \right|^2 \qquad [14]$$

is only a good approximation for Eq. [13] in the case of the group $i$ of spins being far apart from the group $j$. Although the general application of Eq. [14] for the treatment of fast jumps is not recommended, it is still offered and might be useful in special cases. The spectral-density function of auto-relaxation $J^n_{ii}(\omega)$ is given by

$$J^n_{ii} = \frac{1}{r^6_{\mu\nu}} \left[ \frac{1}{4} \frac{\tau_c}{1 + \omega^2 \tau_c^2} + \frac{3}{4} \frac{\tau_{e,ij}}{1 + \omega^2 \tau^2_{e,ij}} \right] \qquad [15]$$

for fast as well as for slow jumps, having $\tau^{-1}_{e,ij} = \tau_c^{-1} + \tau^{-1}_{1,ij}$.

These jump models can be regarded as methods for the calculation of the order parameter $S_{ij}$ in the model-free approach of Lipari and Szabo. In this approach the motions of the molecule are assumed to be a superposition of a slow

overall motion with a correlation time $\tau_c$ and rapid internal motions with correlation times $\tau_{1,ij}$ and generalized order parameters $S_{ij}$ which are a measure for their local restriction. Lipari and Szabo show (*24, 25*) that in this case

LIPARI:

$$J^n_{ij}(\omega) = \frac{1}{r^6_{ij}}\left(\frac{S^2_{ij}\tau_c}{1 + \omega^2\tau^2_c} + \frac{(1 - S^2_{ij})\tau_{e,ij}}{1 + \omega^2\tau^2_{e,ij}}\right) \qquad [16]$$

with $\tau^{-1}_{e,ij} = \tau^{-1}_c + \tau^{-1}_{1,ij}$. Due to the assumption that $\tau_{1,ij} \ll \tau_c$ the second term in Eq. [16] is often neglected, simplifying Eq. [16] to

LIPARI_1:

$$J^n_{ij}(\omega) = \frac{1}{r^6_{ij}}\frac{S^2_{ij}\tau_c}{1 + \omega^2\tau^2_c}\,. \qquad [17]$$

*Anisotropy Correction*

By a suitable combination of the spectral densities presented above, it is possible to set up a detailed model for the internal and overall motions of the molecule. Such a model could describe these motions as a superposition of a slow overall rotational diffusion of the molecule with a rotational correlation time $\tau_c$ and fast internal motions which may vary from spin pair to spin pair in the molecule. Usually, it is assumed that this overall diffusion is isotropic. This implies that the shape of the molecule can be approximated by a sphere. Then, the correlation time $\tau_c$ can be estimated by applying the Stokes–Einstein relation

$$\tau_c = \frac{MV'\eta}{kT} = \frac{4\pi\eta d^3}{3kT}\,, \qquad [18]$$

where $\eta$ is the viscosity of the solvent, $k$ the Boltzmann constant, $M$ the molecular mass of the molecule, $T$ the temperature, $d$ the radius of the molecule, and $V'$ the partial volume, which can be assumed to have a constant value for all globular proteins (*38*).

Especially for longish molecules such as DNA or coiled–coiled helices, the assumption of a spherical molecule is not justified. The motion of a molecule with arbitrary shape can be described by three independent rotational diffusion times for the rotation among its three main axes of inertia. The ab initio calculation of these rotational diffusion times from the molecular structure is extremely difficult. However, the problem can be tackled by a special anisotropy correction implemented in RELAX that can be applied if the overall shape of the molecule resembles more an ellipsoid than a sphere. The correction derived in the following is of pure geometric nature. It corrects for the changes in the relaxation rates that would arise from a fictitious deformation of the molecule from a sphere to its actual shape, keeping the volume of the molecule constant.

To approximate the rotational properties of the given trial structure, the ellipsoid is chosen to have an equal volume and equal moments of inertia as the molecule under consideration. With this assumption its semiaxes $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are uniquely defined and can be calculated numerically from the structure. They are later used to calculate the coefficients of a correcting polynomial.

The dipolar relaxation of a rigid molecule diffusing anisotropically with three independent rotational-diffusion coefficients $R_1$, $R_2$, and $R_3$ can be treated according to the theory presented by Woessner (*39*), of which the essentials are quoted in the following. In the most general, completely anisotropic, case the spectral-density function is

$$J^n_{ij}(\omega) = \frac{1}{2r^6_{ij}}(\alpha_{0,0} + \alpha_{1,0} + \alpha_{0,1}m^2_{ij} + \alpha_{2,0}l^4_{ij} + \alpha_{0,2}m^4_{ij} + \alpha_{1,1}l^2_{ij}m^2_{ij}), \qquad [19]$$

whereby the coefficients $\alpha_{l,m}$ are

$$\begin{aligned}
\alpha_{0,0} &= T_\Sigma + 1/6(\delta_1 + \delta_2 - 2\delta_3)T_\Delta,\\
\alpha_{1,0} &= 6T_2 - 3T_\Sigma - (\delta_2 - \delta_3)T_\Delta,\\
\alpha_{0,1} &= 6T_1 - 3T_\Sigma - (\delta_1 - \delta_3)T_\Delta,\\
\alpha_{2,0} &= 3T_\Sigma - 1/2(\delta_1 - 2\delta_2 + \delta_3)T_\Delta - 6T_2,\\
\alpha_{0,2} &= 3T_\Sigma - 1/2(\delta_2 - 2\delta_1 + \delta_3)T_\Delta - 6T_1,\\
\alpha_{1,1} &= 3T_\Sigma + (\delta_1 + \delta_2 - 2\delta_3)T_\Delta - 6(T_1 + T_2 - T_3).
\end{aligned} \qquad [20]$$

The $\delta_i$ and $T_i$ for $i \in \{1, 2, 3, +, -\}$ are given by

$$\begin{aligned}
\delta_i &= (R_i - \bar{R})/\sqrt{\bar{R}^2 - L^2},\\
T_\Delta &= T_+ - T_-,\\
T_\Sigma &= T_+ + T_-,
\end{aligned}$$

and

$$T_i = \tau_i/(1 + \omega^2\tau^2_i), \qquad [21]$$

with

$$\begin{aligned}
1/\tau_\pm &= 6(\bar{R} \pm \sqrt{\bar{R}^2 - L^2})\\
1/\tau_i &= 3(R_i + \bar{R}).
\end{aligned} \qquad [22]$$

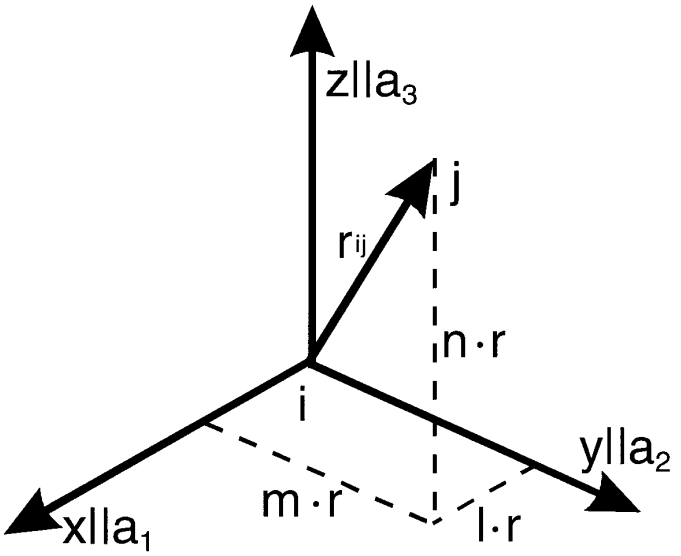



**FIG. 1.** A correction for the anisotropy of the diffusion can be applied in RELAX. The overall shape of the molecule is approximated by an ellipsoid with equal volume and equal moments of inertia. Correction factors are calculated which depend only on the direction cosines $l_{ij}$, $m_{ij}$, and $n_{ij}$ of the vector $\mathbf{r}_{ij}$ connecting two spins $i$ and $j$ with the semiaxes $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ of the ellipsoid.

$\bar{R}$ and $L^2$ only depend on the rotational-diffusion coefficients $R_i$:

$$\bar{R} = (R_1 + R_2 + R_3)/3$$
$$L^2 = (R_1R_2 + R_1R_3 + R_2R_3)/3. \qquad [23]$$

The ratio of the isotropic Eq. [11] to the anisotropic Eq. [19] spectral-density function depends only on the squares of the direction cosines $l_{ij}$ and $m_{ij}$ of the vector $\mathbf{r}_{ij}$, which connects a pair of spins $i$ and $j$ with the semiaxes $\mathbf{a}_1$ and $\mathbf{a}_2$ of the ellipsoid (Fig. 1). It is independent of the spatial separation of $i$ and $j$ $r_{ij}$. Thus it is possible to express this ratio as a polynomial of second degree in $l_{ij}^2$ and $m_{ij}^2$

$$C_{\text{anis}}(l_{ij}^2, m_{ij}^2) = J^n_{\text{aniso}}(\omega)/J^n_{\text{iso}}(\omega)$$
$$= \frac{1 + \tau_c^2\omega^2}{2\tau_c}(\alpha_{0,0} + \alpha_{1,0}\, l_{ij}^2 + \alpha_{0,1}m_{ij}^2 + \alpha_{2,0}l_{ij}^4$$
$$+ \alpha_{0,2}m_{ij}^4 + \alpha_{1,1}l_{ij}^2m_{ij}^2) \qquad [24]$$

with the coefficients $\alpha_{l,m}$ given by Eq. [20].

In general the rotational-diffusion coefficients $R_i$ are hard to determine but if, as assumed, the overall shape of the molecule resembles an ellipsoid with the semiaxes $a_1$, $a_2$, and $a_3$, Stokes hydrodynamic theory yields

$$R_i = \frac{kT}{16\pi\eta}\frac{3(a_1^2P_1 + a_2^2P_2 + a_3^2P_3 - a_i^2P_i)}{a_1^2 + a_2^2 + a_3^2 - a_i^2}. \qquad [25]$$

The $P_i$ are defined by the elliptical integrals

$$P_i = \int_0^\infty \frac{ds}{(a_i^2 + s)\sqrt{(a_1^2 + s)(a_2^2 + s)(a_3^2 + s)}}. \qquad [26]$$

In the special case of a sphere with radius $d = a_1 = a_2 = a_3$ these equations lead to Eq. [18], allowing one to calculate an isotropic rotational correlation time $\tau_c$.

Woessner's theory can now be used to calculate the wanted geometric anisotropy correction. Inserting Eq. [18] into Eq. [25] and using the relation

$$P_1 + P_2 + P_3 = 2/(a_1a_2a_3) = 2/d^3, \qquad [27]$$

one can express the diffusion coefficients $R_i$ of any ellipsoid with volume $\frac{4}{3}\pi\, a_1a_2a_3 = \frac{4}{3}\pi\, d^3$ in terms of $\tau_c$, $a_1$, $a_2$, and $a_3$:

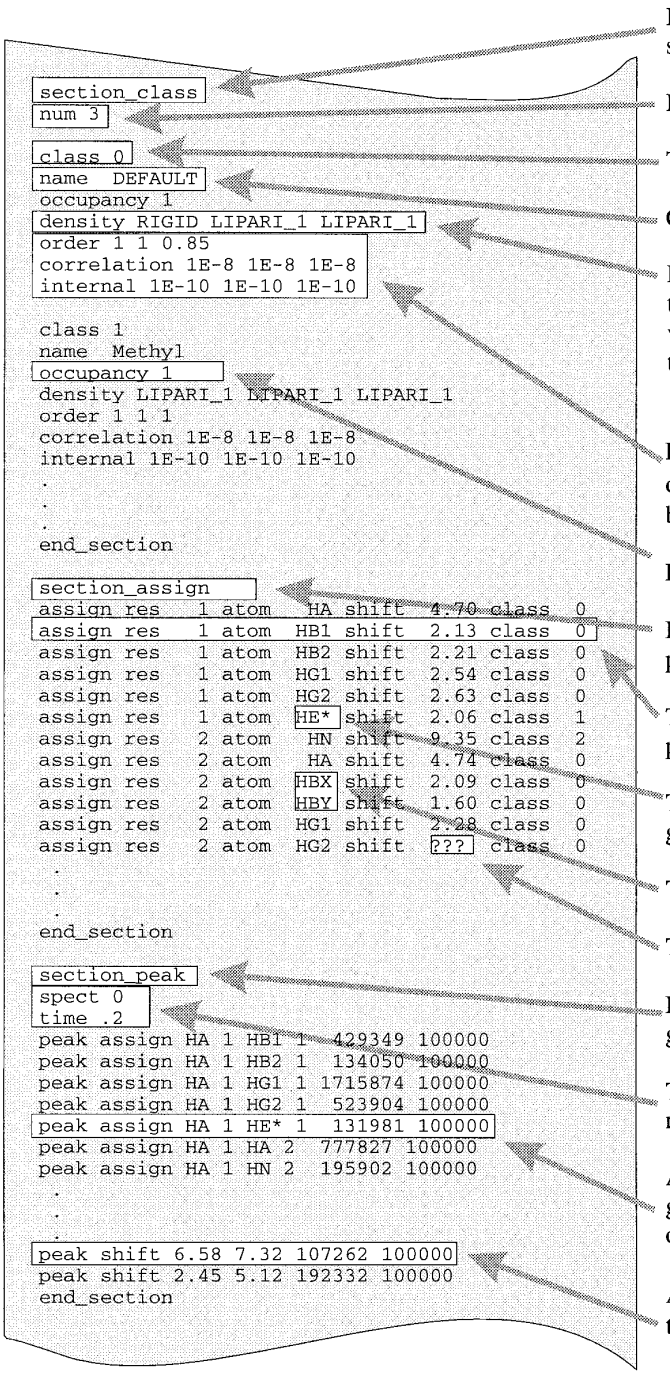
$$R_i = \frac{1}{2\tau_c(P_1 + P_2 + P_3)}\frac{a_1^2P_1 + a_2^2P_2 + a_3^2P_3 - a_i^2P_i}{a_1^2 + a_2^2 + a_3^2 - a_i^2}. \qquad [28]$$

These expressions for the rotational-diffusion coefficients $R_i$ only depend on the ratios of the semiaxes of the ellipsoid. Inserting Eq. [28] into Eq. [23], it can now be used to calculate the coefficients $\alpha_{l,m}$ Eq. [20] of correcting polynomial $C_{\text{anis}}(l_{ij}^2, m_{ij}^2)$ Eq. [24]. Although the calculation of $\alpha_{l,m}$ is computationally very expensive, involving the numerical determination of the tensor of inertia and the numerical integration of elliptical integrals, it must be performed for every trial structure only once. The factors correcting for the anisotropy can be computed for every spin pair simply by calculating the direction cosines $l_{ij}$ and $m_{ij}$ and substituting them in Eq. [24]. Thus the anisotropy correction has only minimal influence on the overall computation time.

## PRACTICAL IMPLEMENTATION

### *The SPT File*

To calculate volumes of NOE cross peaks, it is necessary to know the three-dimensional structure of the molecule. This information is supplied to RELAX by a pdb file which contains the Cartesian coordinates of all atoms. To describe the additional information that is related to the NMR-specific parameters a new file type—the spt file—has been developed and defined (Fig. 2). It consists of three sections describing the parameters of the applied dynamic model, resonance line assignments, and, if available, experimental peak volumes. The spt file can be set up manually by the aid of any spread-sheet program, or the program AURELIA (*27*) can be used to generate a raw frame of an spt file which should be edited by the user to best suit the properties of the molecule under investigation.




In this section all classes are listed and their properties are specified.

In this example there are three classes.

The following properties apply to class 0.

Classes can be given names.

For the calculation of a relaxation rate between two groups that are members of class 0 the spectral density model RIGID will be used. For the calculation of a rate between one group that belongs to class 0 and another group that belongs to class 1 the model LIPARI_1 will be used.

For all pairs of classes the overall correlation time $\tau_c$, the order parameter $S^2$ and the internal correlation time $\tau_1$ must be specified.

For each class an optional occupancy can be given.

In this section groups are defined and can be assigned to ppm-values.

The atom "HB1" of residue 1 has a chemical shift of 2.13 ppm and is member of class 0.

The atoms "HE1", "HE2" and "HE3" are combined to one group.

The assignment of "HB1" and "HB2" is not stereo-specific.

The chemical shift of this atom is unknown.

In this optional section volumes of experimental peaks can be given.

The spectrum 0 has been recorded at a mixing time of 200 ms.

A cross peak between the atom "HA" of residue 1 and the group "HE*" of residue 1 has been measured with a volume of 131981 with an absolute error of 100000.

At the position ($\delta_1$=6.58 ppm, $\delta_2$=7.32 ppm) a cross peak with the volume 107262 has been measured which is unassigned.

**FIG. 2.** A sample extract of an spt file is shown. The file consists of three sections. In the first section, classes of protons are defined along with their properties which describe the dynamics of the molecule. In the second section, all groups of protons are listed and assigned to a class as well as, optionally, to a ppm value. Equivalent atoms can be combined into groups. In the optional third section, experimental volumes can be supplied for a comparison with simulated cross peaks or for further calculations.

To reduce the calculation time, magnetically equivalent spins can be combined to groups. In order to achieve maximum flexibility, a classification concept is applied which allows the assignment of each group to a user-defined class. In the first section of the spt file for all classes and pairs of classes a set of properties is defined which is used to set up the relaxation matrix. These are the type of spectral density that is to be applied, the overall correlation time $\tau_c$, the order parameter $S$, and the internal correlation time $\tau_1$. Additionally, an occupation number can be specified which allows

one, for example, to take into account that amide protons have a reduced occupancy if the solvent contains $D_2O$. RELAX allows the simultaneous application of the spectral-density functions described above. To improve readability the symbolic names RIGID, SLOW_JUMP, FAST_JUMP, AVERAGE_3, LIPARI, and LIPARI_1 are used. The anisotropy correction can be activated by the keyword ANIS. In the second section all groups of spins are listed that would be visible in a corresponding experimental spectrum. If known, a ppm value can be assigned to a group of spins. If such an assignment is not known stereospecifically, this can be indicated by replacing the atom index number by ''X'' or ''Y'' in the atom name. An asterisk can be used to indicate atoms that should be combined into a group. The choice of the spectral-density function defines how their coordinates are to be averaged. If back-calculated volumes should be quantitatively compared in some way with experimental data it is possible to give experimental cross-peak volumes in an optional third section. These peaks can be either assigned or unassigned and can be entered for multiple mixing times. If required, specified peaks can be used to scale the simulated to the experimental data.

*Computation-Time-Controlled Submatrix Sizes*

The calculation of NOE volumes involves the solution of a matrix-exponential function which is a problem with order of complexity $N^3$ (*40*). Fortunately, relaxation rates between nuclei that are far apart are zero in very good approximation. Therefore, only such cross-peak volumes $V_{ij}$ between spin $i$ and spin $j$ need to be calculated for which $i$ and $j$ are nearer than an outer cutoff value $c_r$. Furthermore, the fact that the relaxation matrix is sparse makes it possible to neglect, when calculating the cross peak $V_{ij}$, all other spins whose distance from $i$ and $j$ is greater than an inner cutoff value $c_u$. Thus $V_{ij}$ can be calculated by only exponentiating $\mathbf{R}'_{ij}$, a submatrix

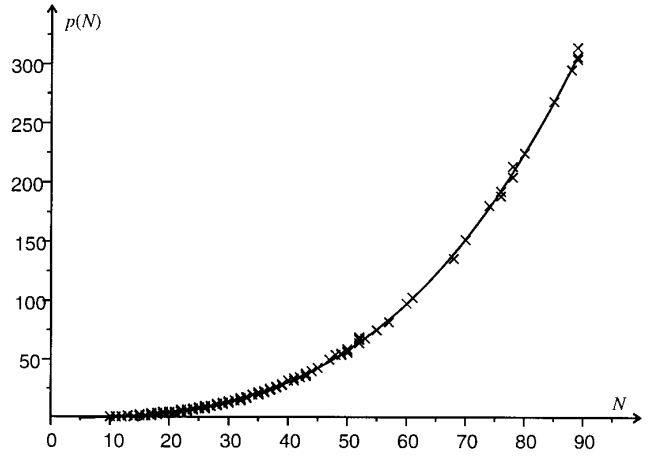



**FIG. 3.** Experimentally determined computation times (×) needed for the calculation of the exponential of a matrix of size $N$ are shown. They can be approximated perfectly by a polynomial of third degree, $p(N) = 0.41 - 0.0043N + 0.0028N^2 + 1.83 \times 10^{-5} N^3$. This polynomial can be used to predict computation times without actually performing the calculation. The data are obtained on a Bruker Aspect 1 workstation.

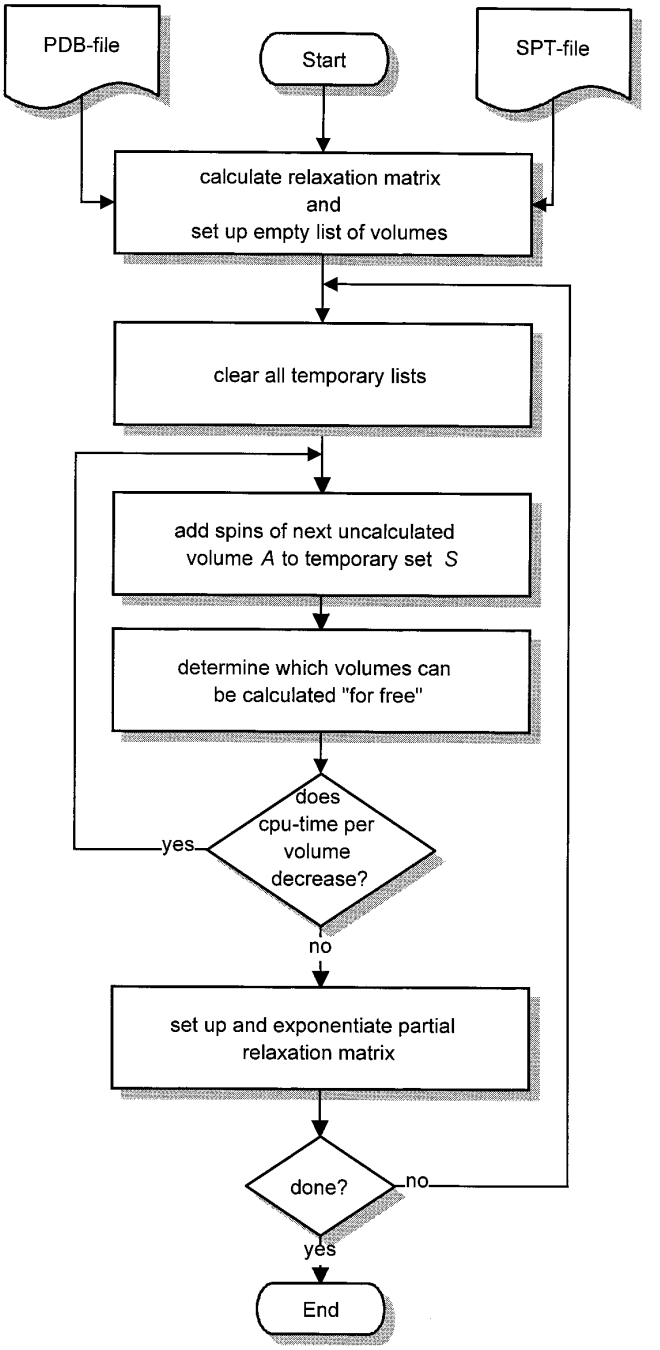



**FIG. 4.** The calculation of the matrix-exponential function of the relaxation matrix $\mathbf{R}'$ is performed by exponentiating small submatrices of $\mathbf{R}'$. If it reduces the overall computation time, submatrices are combined and exponentiated together. To determine the optimal composition of these submatrices, the computation time is estimated for various arrangements prior to the actual calculations.

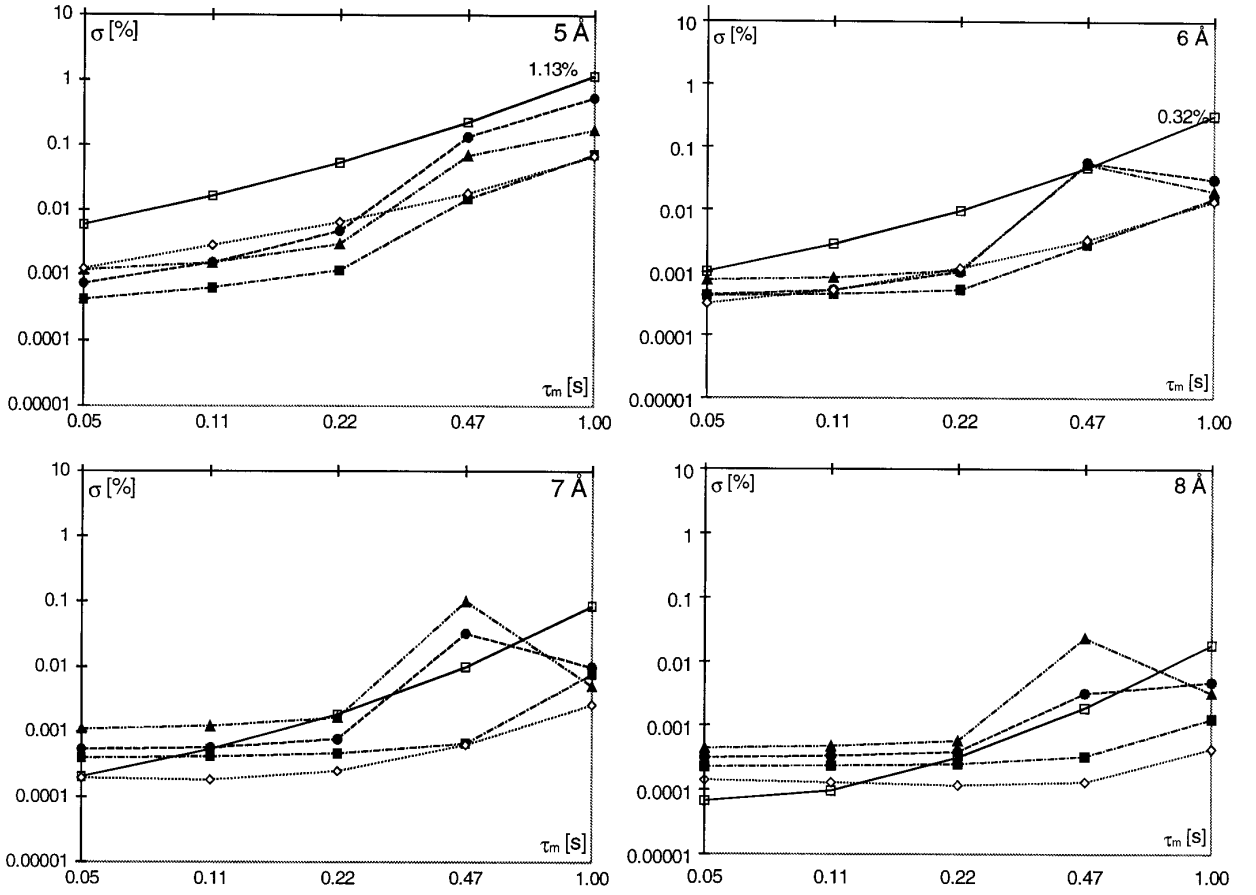



**FIG. 5.** The rmsd values of the average relative deviations of calculations performed with the inner cutoff values $c_u$ = (5, 6, 7, and 8 Å) from calculations performed with $c_u$ = 9 Å are shown. As correlation times, $\tau_c$ = [42 ps (■), 76 ps (●), 240 ps (▲), 760 ps (◇), and 2.4 ns (□)] were chosen. The mixing time was set to $t$ = (50, 110, 220, and 470 ms and 1 s). Only in the worst case of $t$ = 1 s and $c_u$ = 5 Å does the rmsd value exceed 1%. This peak value of 1.32% corresponds to average errors in the distances of 0.22%.

of $\mathbf{R}'$ that only contains relaxation rates between spins that are neighbors of either $i$ or $j$. Because of the $N^3$ dependence of the matrix-diagonalization time, calculating the exponentials of separate submatrices for all peaks can be computationally more efficient than exponentiating the relaxation matrix as a whole. It is possible that a pair of spins $k$ and $l$ have some spatial neighbors in common with spins $i$ and $j$. In such a case the overlap of the submatrices $\mathbf{R}'_{ij}$ and $\mathbf{R}'_{kl}$ can make it more efficient to combine them to $\mathbf{R}'_{ijkl}$ and calculate $V_{ij}$ and $V_{kl}$ by exponentiating $\mathbf{R}'_{ijkl}$. Thus it is not optimal to exponentiate submatrices for all peaks separately.

In RELAX, a dynamic determination of submatrix sizes is applied to minimize computation time. It uses the fact, mentioned above, that the dependence of the diagonalization time of a matrix on its size can be approximated by a polynomial of third degree. If its coefficients are known, this polynomial allows the estimation of the computational effort for the diagonalization of any possible submatrix without actually computing it. The coefficients can be obtained, as shown in Fig. 3, by a numerical fit to experimentally determined computation times $p(N)$ for different submatrix sizes $N$. The actual coefficients vary little from one computer system to another, and the general behavior of the polynomial does not change. Therefore, their choice is not critical for a minimal computation time.

The algorithm implemented in RELAX is explained by Fig. 4. First the relaxation matrix and a list of all peaks volumes that are to be calculated are set up. In a bit matrix **BM**, a bit set at position $(i, j)$ states that the distance of at least one member of group $i$ from one member of group $j$ is smaller than the cutoff value $c_u$. The binary union of row $i$ and row $j$ of **BM** represents the set of spins that must be regarded when calculating the cross peak $V_{ij}$. If, for example, the next uncalculated cross peak is $V_{kl}$, a calculation of both peaks by exponentiating a common submatrix would involve all spins that are represented by the union of the rows $i$, $j$, $k$, and $l$ of **BM**. This larger submatrix allows in addition to $V_{ij}$ and $V_{kl}$, the calculation of $V_{ik}$, $V_{il}$, $V_{jk}$, and $V_{jl}$. Depending on which of these peaks still are to be calculated, it can be decided by the knowledge of $p(N)$ which combination re-

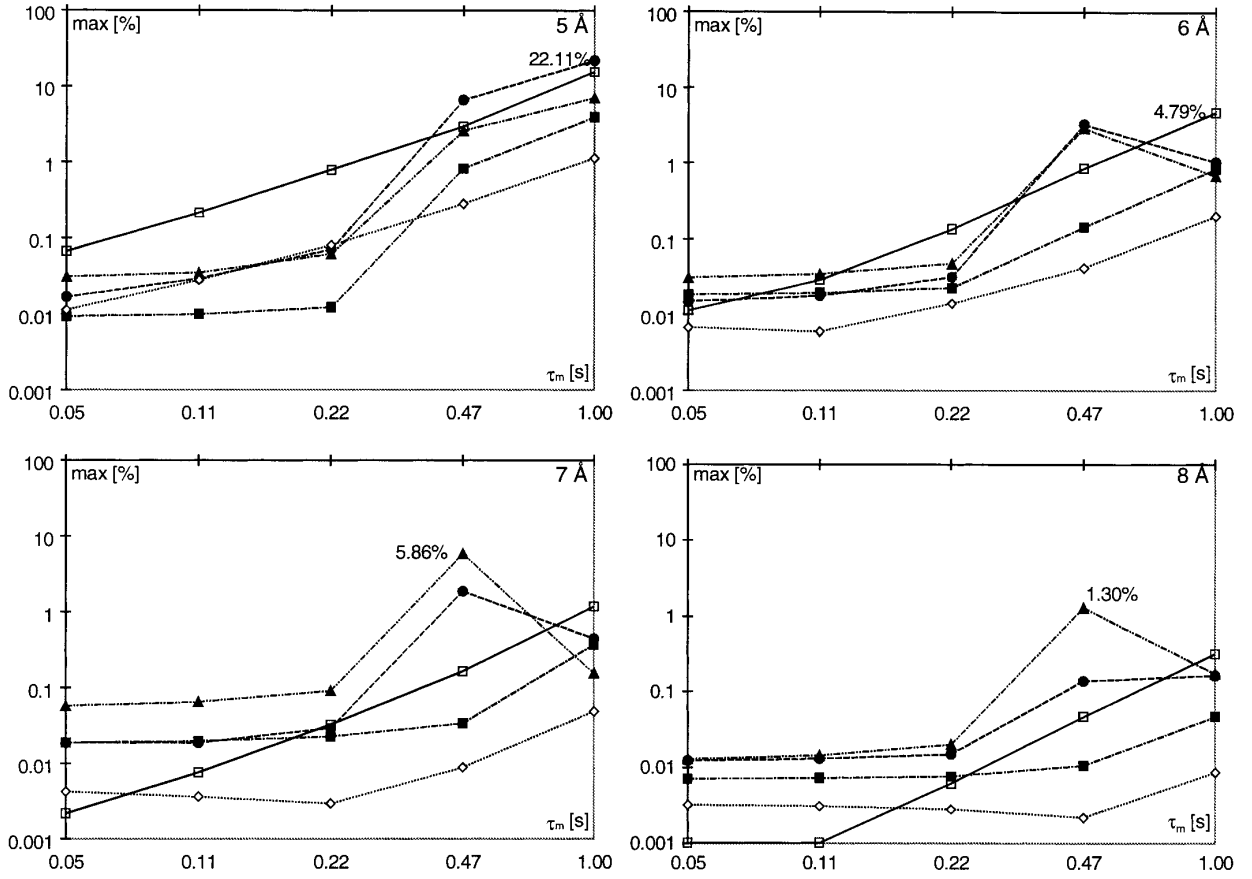



**FIG. 6.** The maximum relative deviations are shown. Values were calculated with the same parameters and are displayed by the same symbols as in Fig. 5. Only in two extreme cases do the maximum relative deviations exceed 10%.

quires less computation time per peak. If it is favorable to calculate both peaks separately the calculation is performed immediately; otherwise it is postponed and a check is made whether even more peaks can be calculated together.

When a fully relaxed spectrum is simulated, the use of submatrices has an additional advantage: if only a limited number of cross peaks are of actual interest, only those submatrices containing these peaks need to be diagonalized. This can further speed the calculation enormously.

## RESULTS AND DISCUSSION

The algorithm implemented in RELAX depends on two different cutoff values: the user-defined outer cutoff value $c_r$ determines the distance range which is of interest for the user, and the inner cutoff value $c_u$, which is invisible to the user, determines the minimal distance range in the bonds of which neighboring spins are taken into account as spin-diffusion pathways when calculating cross-peak volumes.

**TABLE 1**

**Execution Times for the Calculation of $n$ Cross Peaks with an Outer Cutoff Parameter $c_r$ = 5.0 Å Are Shown for Peptides with Different Numbers of Residues $n_{res}$ and Sizes of the Relaxation Matrix $N$**

| $n_{res}$ | 25 | 50 | 75 | 100 | 125 | 150 | 175 | 200 | 225 | 250 | 275 | 300 | 325 | 350 | 372 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | 1088 | 2706 | 4921 | 6509 | 9073 | 10,961 | 14,487 | 16,280 | 16,431 | 18,108 | 20,544 | 22,708 | 24,912 | 27,849 | 29,791 |
| $N$ | 176 | 356 | 533 | 670 | 909 | 1,070 | 1,225 | 1,412 | 1,592 | 1,751 | 1,926 | 2,107 | 2,276 | 2,469 | 2,635 |
| Whole | 1.7 s | 23.0 s | 97 s | 217 s | 651 s | 1,096 s | 1,837 s | 2,796 s | 4,205 s | 5,712 s | 8,031 s | 10,911 s | 13,996 s | 34,113 s | 43,377 s |
| RELAX | 1.0 s | 4.3 s | 19 s | 41 s | 51 s | 67 s | 82 s | 90 s | 116 s | 126 s | 133 s | 173 s | 182 s | 219 s | 235 s |

*Note.* The rows are labeled as in Table 2.

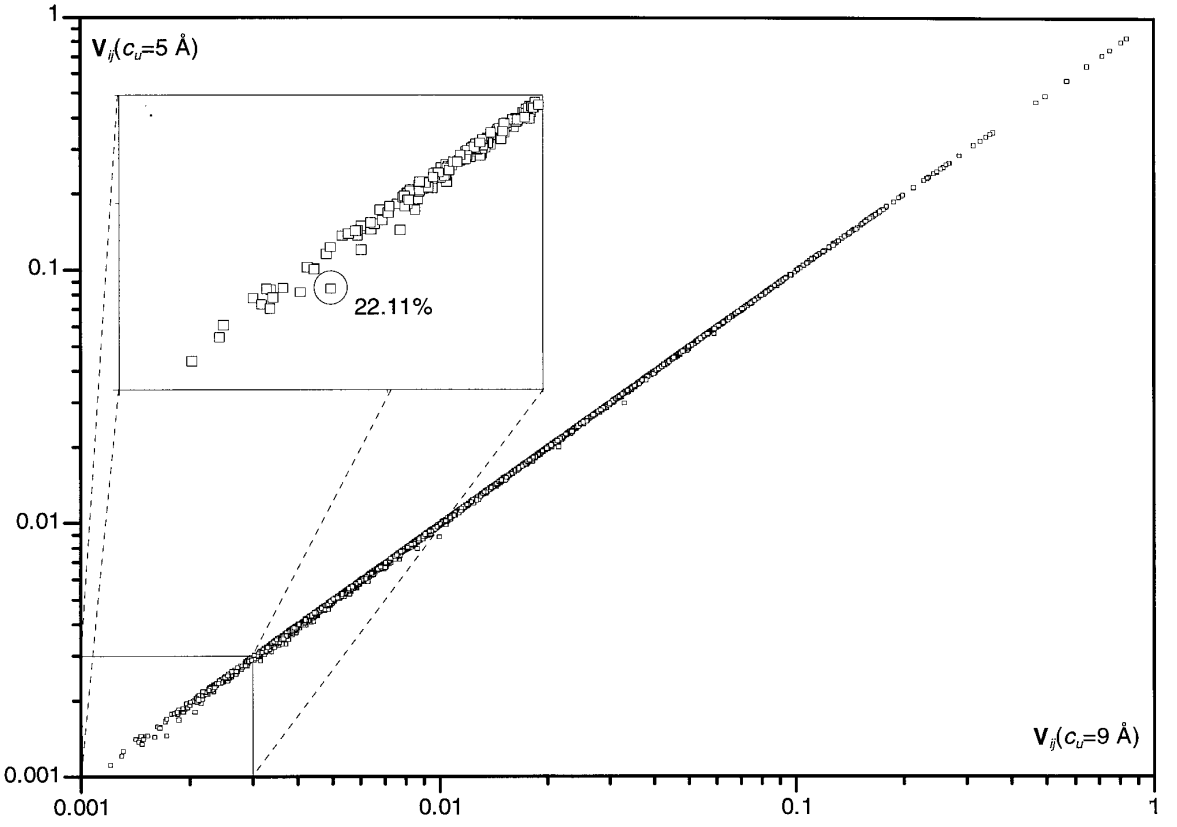



**FIG. 7.** The distribution of deviations in the worst-case calculation. It is evident that large deviations occur only for very small cross-peak volumes. It cannot be decided whether this is due to the algorithm of RELAX or to a general numerical instability of the QL algorithm used for matrix diagonalization.

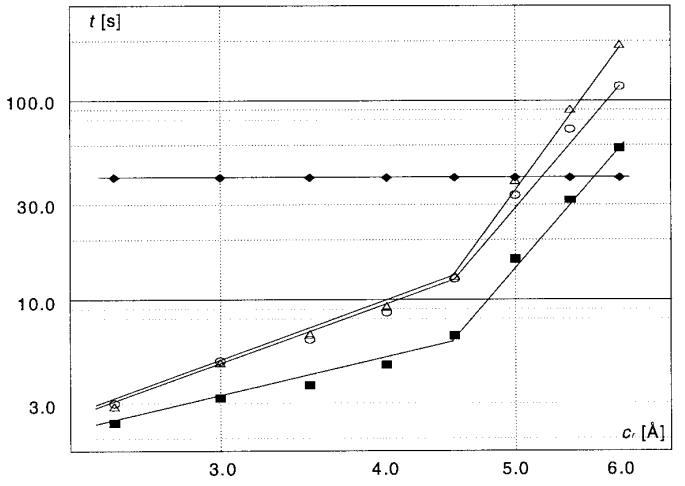



**FIG. 8.** The dependence of computation time on the outer cutoff value $c_r$. The bend in the curves at 4.5 Å results from the fact that the inner cutoff value $c_u$ is kept constant for $c_r < 4.5$ Å, while it is set to $c_u = c_r$ for larger values of $c_r$. The lines show values obtained by exponentiating the matrix as a whole (♦), by setting up separate matrices for every peak (△), by calculating all peaks that have one spin in common (○), and by RELAX (■).

To investigate the dependence of the numerical stability of the algorithm of RELAX on the inner cutoff value $c_u$, calculations have been performed for the trial structure tendamistat from *Streptomyces tendae*. The spectrometer frequency was set to 500 MHz. A dynamic model was set up applying the spectral-density function FAST_JUMP (*13*) for methyl groups, SLOW_JUMP (*37*) for ring protons, and RIGID (*11*) for all other protons. Using Eq. [18], an overall rotational correlation time $\tau_c$ of 2.4 ns was estimated for the molecule in water at 298 K. For all combinations of the correlation times $\tau_c$ = (24, 76, 240, and 760 ps and 2.4 ns), corresponding to $\omega\tau_c$ = (0.075, 0.24, 0.75, 2.4, and 7.5) and mixing times $t_m$ = (50, 110, 220, and 470 ms and 1 s), the 3631 NOE cross-peak volumes that arise from spin pairs that are not further apart than the outer cutoff value $c_r$ of 5 Å have been calculated. Calculations have been performed for different inner cutoff values $c_u$ = (5, 6, 7, 8, and 9 Å). The cross-peak volumes obtained with $c_u$ = (5, 6, 7, and 8 Å) were divided by the values obtained with $c_u$ = 9 Å and the quotients have been analyzed statistically. The rmsd deviations of these quotients from 1 are shown in Fig. 5; only in the worst case of $c_u$ = 5 Å, $\tau_c$ = 2.4 ns, and $t$ = 1 s does the rmsd value $\sigma$ exceed 1%. This corresponds to neglectable

**TABLE 2**
**Execution Times for the Computation of *n* Cross Peaks Are Shown for Different Strategies of Submatrix Construction in Dependence of the Outer Cutoff Parameters $c_r$**

| $c_r$ | 2.5 Å | 3.0 Å | 3.5 Å | 4.0 Å | 4.5 Å | 5.0 Å | 5.5 Å | 6.0 Å |
|---|---|---|---|---|---|---|---|---|
| *n* | 454 | 856 | 1186 | 1590 | 2230 | 3163 | 4055 | 5016 |
| Whole | 41.0 s | 41.0 s | 41.0 s | 41.0 s | 41.0 s | 41.0 s | 41.0 s | 41.0 s |
| Single | 2.9 s | 4.8 s | 6.7 s | 9.2 s | 13.1 s | 39.7 s | 90.4 s | 192.1 s |
| Spins | 3.0 s | 4.9 s | 6.3 s | 8.6 s | 12.7 s | 33.4 s | 72.0 s | 118.0 s |
| RELAX | 2.4 s | 3.2 s | 3.7 s | 4.7 s | 6.6 s | 15.9 s | 31.6 s | 58.0 s |

*Note.* The 74-residue peptide tendamistat served as a trial structure. The values listed under ''whole'' were obtained by diagonalizing the relaxation matrix as a whole ''Single'' represents times that were obtained when diagonalizing an individual submatrix for each peak volume calculated. The execution times labeled ''spins'' were obtained when calculating all peak volumes that have the first spin in common by exponentiating a common submatrix. In the last row, calculation times that were obtained by the computation-time-controlled algorithm of RELAX are listed.

rmsd errors in the distances of less than 0.2%. Figure 6 shows the maximum deviations of the quotients from 1. For $c_u = 5$ Å, $\tau_c = 76$ ps, and $t = 1$ s, a maximum relative deviation of 22% was found. As shown in Fig. 7, this large deviation occurs for a very small cross-peak volume. This deviation of small values (which would still be tolerable for structure calculations) is probably not due to the small size of the inner cutoff value but represents rounding errors of the numerical diagonalization. However, a mixing time of 1 s is not a realistic scenario since the NOE enhancements are almost zero.

According to these results, an accuracy which is sufficient in all practical situations can be achieved if the inner cutoff value $c_u$ is chosen larger than or equal to $c_r$. However, as calculations with small values of $c_u$ showed, a minimum value of 4.5 Å for $c_u$ is required to keep the accuracy high for values of $c_r$ smaller than 4.5 Å. Therefore, in the current implementation of RELAX the inner cutoff value is set according to the formula $c_u := \max(c_r, 4.5\text{ Å})$.

To investigate the efficiency of the applied strategy of submatrix construction, a series of computation-time measurements has been performed. For fragments of actin with sizes increasing from 25 to 372 residues in steps of 25, computation times have been determined with and without the use of submatrices. Possible numerical problems caused by the large matrix sizes have been neglected. As shown in Table 1, the use of submatrices speeds up the calculation drastically. For much larger structures, it can be expected that the computation time per peak of RELAX is independent of the size of the structure, while exponentiating the relaxation matrix as a whole becomes almost untractable.

The influence of the computation-time control implemented in RELAX has been investigated by performing calculations with different strategies of submatrix construction: without any submatrices, with individual submatrices for every cross peak, combining those submatrices that have the first spin in common, and with the computation-time-controlled algorithm of RELAX. The results are shown in Fig. 8 and Table 2. In addition to the reduction of CPU time obtained by the introduction of submatrices, the computation-time control makes the calculations uniformly another 50% faster than they are with other strategies of submatrix construction.

In summary, RELAX represents a very flexible and computationally efficient program, which is suitable for the interactive evaluation of NOESY spectra of large molecules on the basis of the full-relaxation-matrix formalism.

## REFERENCES


1. Y. Liu, D. Zhao, R. Altman, and O. Jardetzky, *J. Biol. NMR* **2,** 373 (1992).
2. D. Neuhaus and M. Williamson, ''The Nuclear Overhauser Effect in Structural and Conformational Analysis,'' VCH, New York/Weinheim/Cambridge, 1989.
3. A. Kumar, G. Wagner, R. R. Ernst, and K. Wüthrich, *J. Am. Chem. Soc.* **103,** 3654 (1981).
4. B. A. Borgias, P. D. Thomas, and T. L. James, ''COmplete Relaxation Matrix Analysis (CORMA),'' University of California, San Francisco, 1987, 1989.
5. K. M. Banks, D. R. Hare, and B. R. Reid, *Biochemistry* **28,** 6996 (1989).
6. A. T. Brünger, ''X-PLOR Version 3.1,'' Yale Univ. Press, New Haven/London, 1992.
7. L. Zhu and B. R. Reid, *J. Magn. Reson. B* **106,** 227 (1995).
8. J. W. Keepers and T. L. James, *J. Magn. Reson.* **57,** 404 (1984).
9. S. Shibata and K. Akasaka, *Magn. Reson. Chem.* **28,** 129 (1990).
10. B. A. Borgias and T. L. James, *J. Magn. Reson.* **87,** 475 (1990).
11. H. Liu, D. L. Banville, V. J. Basus, and T. L. James, *J. Magn. Reson. B* **107,** 51 (1995).
12. F. J. M. v. d. Veen, M. J. J. Blommers, R. E. Schouten, and C. W. Hilbers, *J. Magn. Reson.* **94,** 140 (1991).
13. S.-G. Kim and B. R. Reid, *J. Magn. Reson.* **100,** 382 (1992).
14. R. Boelens, T. M. G. Koning, G. A. van der Marel, J. H. van Boom, and R. Kaptein, *J. Magn. Reson.* **82,** 290 (1989).
15. A. M. J. J. Bonvin, J. A. C. Rullmann, R. M. J. N. Lamerichs, R. Boelens, and R. Kaptein, *Proteins* **15,** 385 (1993).

16. A. M. J. J. Bonvin, R. Boelens, and R. Kaptein, *J. Biomol. NMR* **1,** 305 (1991).
17. C. Gonzalez, J. A. Rullmann, A. M. J. J. Bonvin, R. Boelens, and R. Kaptein, *J. Magn. Reson.* **91,** 659 (1991).
18. A. T. Brünger, G. M. Clore, A. M. Gronenborn, R. Saffrich, and M. Nilges, *Science* **261,** 328 (1993).
19. M. Nilges, J. Habazettl, A. T. Brünger, and T. A. Holak, *J. Mol. Biol.* **219,** 499 (1991).
20. Y. Xu, I. P. Sugár, and N. R. Krishna, *J. Biomol. NMR* **5,** 37 (1995).
21. H. N. B. Moseley, E. V. Curto, and N. R. Krishna, *J. Magn. Reson. B* **108,** 243 (1995).
22. P. Güntert, K. D. Berndt, and K. Wüthrich, *J. Biomol. NMR* **3,** 601 (1993).
23. Ch. Mumenthaler and W. Braun, *J. Mol. Biol.* **254,** 465 (1995).
24. G. Lipari and A. Szabo, *J. Am. Chem. Soc.* **104,** 4546 (1982).
25. G. Lipari and A. Szabo, *J. Am. Chem. Soc.* **104,** 4559 (1982).
26. H. Liu, M. Tonelli, and T. L. James, *J. Magn. Reson. B* **111,** 85 (1996).
27. K.-P. Neidig, M. Geyer, A. Görler, C. Antz, R. Saffrich, W. Beneicke, and H. R. Kalbitzer, *J. Biomol. NMR* **6,** 255 (1995).
28. I. Solomon, *Phys. Rev.* **99,** 2 (1955).
29. J. Jeener, B. H. Meier, P. Bachmann, and R. R. Ernst, *J. Chem. Phys.* **71**(11), 4546 (1979).
30. S. Macura and R. R. Ernst, *Mol. Phys.* **41**(1), 95 (1980).
31. R. R. Ernst, G. Bodenhausen, and A. Wokaun, ''Principles of Nuclear Magnetic Resonance in One and Two Dimensions,'' Clarendon Press, Oxford, 1987.
32. M. Köck and C. Griesinger, *Angew. Chem.* **106,** 338 (1994).
33. R. Kubo and K. Tomita, *J. Phys. Soc. Jpn.* **9,** 888 (1954).
34. J. Tropp, *J. Chem. Phys.* **72,** 6035 (1980).
35. D. E. Woessner, *J. Chem. Phys.* **36,** 1 (1962).
36. D. E. Woessner, *J. Chem. Phys.* **42,** 1855 (1965).
37. P. F. Yip and D. A. Case, ''Computational Aspects of the Study of Biomolecular Macromolecules by NMR,'' Plenum Press, New York, 1991.
38. K. H. Hausser and H. R. Kalbitzer, ''NMR für Mediziner und Biologen,'' Springer, Berlin, 1991.
39. D. E. Woessner, *J. Chem. Phys.* **37,** 647 (1962).
40. C. Moler and C. van Loan, *SIAM Rev.* **20,** 801 (1978).