Journal of Magnetic Resonance **149,** 204–217 (2001)
doi:10.1006/jmre.2001.2295, available online at http://www.idealibrary.com on IDEAL®

Determination of the Rotational Diffusion Tensor of Macromolecules in Solution from NMR Relaxation Data with a Combination of Exact and Approximate Methods—Application to the Determination of Interdomain Orientation in Multidomain Proteins

Ranajeet Ghose, David Fushman,[1] and David Cowburn[2]

Rockefeller University, 1230 York Avenue, New York, New York 10021

Received October 2, 2000; revised January 9, 2001; published online March 22, 2001

In this paper we present a method for determining the rotational diffusion tensor from NMR relaxation data using a combination of approximate and exact methods. The approximate method, which is computationally less intensive, computes values of the principal components of the diffusion tensor and estimates the Euler angles, which relate the principal axis frame of the diffusion tensor to the molecular frame. The approximate values of the principal components are then used as starting points for an exact calculation by a downhill simplex search for the principal components of the tensor over a grid of the space of Euler angles relating the diffusion tensor frame to the molecular frame. The search space of Euler angles is restricted using the tensor orientations calculated using the approximate method. The utility of this approach is demonstrated using both simulated and experimental relaxation data. A quality factor that determines the extent of the agreement between the measured and predicted relaxation data is provided. This approach is then used to estimate the relative orientation of SH3 and SH2 domains in the SH(32) dual-domain construct of Abelson kinase complexed with a consolidated ligand. © 2001 Academic Press

Key Words: **relaxation; rotational diffusion tensor; singular value decomposition; domain orientation.**

INTRODUCTION

Over the past few years, attention has been focused on the determination of several orientation-dependent properties such as the rotational diffusion tensor from NMR relaxation measurements (*1–7*) and the molecular alignment tensor from residual dipolar couplings (*8–10*). These measurements that provide long-range orientational information often complement the lack of long-range NOEs, and are now routinely included in structure refinement protocols (*11*). Both methods mentioned above provide approaches to the determination of interdomain orientation in weakly interacting multidomain systems (*6, 12–14*) where interdomain NOE information is scarce. The accurate determination of the rotational diffusion tensor is also important since it is essential to interpret correctly relaxation data in terms of local dynamics on the fast (ps–ns) timescale (*15–17*).

The influence of the nonisotropic nature of the overall tumbling on NMR relaxation in the solution state has been known since the work of Woessner (*18*). Early experimental applications to macromolecules were by Tjandra *et al.* (*1*), Broadhurst *et al.* (*19*), and Bruschweiler *et al.* (*2*). In the work of Tjandra *et al.*, hydrodynamic calculations were used in conjunction with relaxation measurements to characterize the rotational diffusion tensor in ubiquitin. Bruschweiler *et al.* (*2*) showed that, for a small degree of rotational anisotropy, the diffusion tensor, in any arbitrary frame of reference, can be written in quadratic form, thus simplifying the analysis. This method was applied to the determination of the diffusion tensor in ubiquitin and ribonuclease-H by Lee *et al.* (*3*). Other groups have used the exact equations provided by Woessner (*18*) to determine the tensor in the axially symmetric case (*6, 15–17, 20*).

Blackledge *et al.* (*5*) have shown that, in the presence of a significant deviation from axial symmetry, fitting to an axially symmetric model yields two nearly indistinguishable minima corresponding to diffusion as an oblate or as a prolate ellipsoid. This necessitates, in the case of significant anisotropy and deviation from axial symmetry of the diffusion tensor, use of the exact equations of Woessner (*18*) for a fully anisotropic diffusion tensor. This approach requires a search over a six-dimensional parameter space corresponding to the three principal values of the diffusion tensor and the three Euler angles relating the principal axis frame to the molecular frame. This approach is very computationally intensive and time-consuming. Alternative approaches have recently been suggested which employ simulated annealing (*21*) or Bayesian statistics (*22*). The former approach is completely general though the latter makes assumptions about the nature of the diffusion tensor.

In this paper we propose an alternate, hybrid approach. In the first step, this procedure uses the approximate approach of Bruschweiler (*2*) and Lee (*3*) to obtain starting values for the

[1] Present address: Center for Biomolecular Structure and Organization, Department of Chemistry and Biochemistry, University of Maryland, 1115 Agriculture/Life Sciences Surge Building, College Park, MD 20742-3360.

[2] To whom correspondence should be addressed. Fax: (212) 327-7566. E-mail: cowburn@rockefeller.edu.

1090-7807/01 $35.00
Copyright © 2001 by Academic Press
All rights of reproduction in any form reserved.

principal components for the diffusion tensor as well as to restrict the values of the Euler angles relating the principal axis frame of the diffusion tensor to the molecular frame. The values of the principal components obtained from the approximate analysis are then used, in the second step, as starting input for a three-dimensional simplex search for the tensor values, combined with a three-dimensional search for the Euler angles which relate the principal axis frame of the diffusion tensor to the molecular frame. The Euler angle search is performed on a three-dimensional grid comprising the angular subspace restricted using the results from the approximate method. We also determine a quality factor for the estimated diffusion tensor. The quality factor provides a measure of the agreement between measured relaxation data and that calculated using the estimated diffusion tensor.

THEORY

Heteronuclear Relaxation

For a heteronuclear ^{15}N–^{1}H system, the ^{15}N relaxation rates, $R_1(1/T_1)$, $R_2(1/T_2)$, and steady state heteronuclear NOE are given by

$$R_1 = d^2[J(\omega_H - \omega_N) + 6J(\omega_H + \omega_N)] + 3(c^2 + d^2)J(\omega_N) \quad [1a]$$

$$R_2 = \frac{d^2}{2}[J(\omega_H - \omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + (c^2 + d^2)\left[2J(0) + \frac{3}{2}J(\omega_N)\right] \quad [1b]$$

$$NOE = 1 + \frac{d^2\left(\frac{\gamma_H}{\gamma_N}\right)[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)]}{R_1} \quad [1c]$$

with $d = -(\mu_0/4\pi)(\gamma_H\gamma_N\hbar/2r_{NH}^3)$, $c = \gamma_N B_0 \Delta\sigma/3$, and $\Delta\sigma$ is the chemical shift anisotropy of the ^{15}N nucleus. Equation [1] can be simplified using the reduced spectral density approach (*23, 24*)

$$R_1 = 7d^2 J(0.921\omega_H) + 3(c^2 + d^2)J(\omega_N) \quad [2a]$$

$$R_2 = \frac{13d^2}{2}J(0.955\omega_H) + (c^2 + d^2)\left[2J(0) + \frac{3}{2}J(\omega_N)\right] \quad [2b]$$

$$NOE = 1 + \frac{5d^2\left(\frac{\gamma_H}{\gamma_N}\right)J(0.87\omega_H)}{R_1}. \quad [2c]$$

It is evident from Eq. [2c] that the steady-state NOE is dependent on the high-frequency components of the spectral density function as well as R_1 (which is available from relaxation measurements). The assumption that $J(\omega) \propto \omega^{-2}$ at $\omega \approx \omega_H$ allows the contribution of the high-frequency components to be estimated for Eqs. [2a] and [2b] using Eq. [2c]. Subtracting these contributions, we obtain (*25, 26*)

$$R_1' = R_1 - 7\left(\frac{0.921}{0.87}\right)^2 HF = 3(c^2 + d^2)J(\omega_N) \quad [3a]$$

$$R_2' = R_2 - \frac{13}{2}\left(\frac{0.955}{0.87}\right)^2 HF = (c^2 + d^2)\left[2J(0) + \frac{3}{2}J(\omega_N)\right], \quad [3b]$$

where $HF = d^2 J(0.87\omega_H) = -(\gamma_N/5\gamma_H)(1 - NOE)R_1$. Using Eqs. [3a] and [3b], we define a quantity ρ which depends solely on the ratio of the spectral densities at frequencies zero and ω_N. This quantity, which is independent of local motion, structural parameters such as NH bond length, and CSA values (to first order), is given by (when the effects of exchange on R_2 are absent)

$$\rho = \frac{4}{3}\left(\frac{R_1'}{2R_2' - R_1'}\right) = \frac{J(\omega_N)}{J(0)}. \quad [4]$$

In this paper, we will relate all derived quantities to ρ, so all effects of local motion will be neglected in what follows.

In the most general case of a rigid asymmetric top (no local motion), the spectral density function at a frequency ω is given by (*18*)

$$J(\omega) = \frac{2}{5}\sum_{i=1}^{5}\frac{A_i D_i}{D_i^2 + \omega^2}, \quad [5]$$

where

$$\begin{aligned} D_1 &= (4D_{xx} + D_{yy} + D_{zz}) \\ D_2 &= (D_{xx} + 4D_{yy} + D_{zz}) \\ D_3 &= (D_{xx} + D_{yy} + 4D_{zz}) \\ D_4 &= 6D_{iso} + 6\sqrt{(D_{iso}^2 - D^2)} \\ D_5 &= 6D_{iso} - 6\sqrt{(D_{iso}^2 - D^2)} \end{aligned} \quad [6]$$

with $D_{iso} = (D_{xx} + D_{yy} + D_{zz})/3$ and $D^2 = (D_{xx}D_{yy} + D_{yy}D_{zz} + D_{xx}D_{zz})/3$. D_{xx}, D_{yy}, and D_{zz} are the principal components of the molecular rotational diffusion tensor $\mathbf{D}$ ($D_{xx} \le D_{yy} \le D_{zz}$). The coefficients A_i are given by

$$\begin{aligned} A_1 &= 3y_d^2 z_d^2 \\ A_2 &= 3x_d^2 z_d^2 \\ A_3 &= 3x_d^2 y_d^2 \end{aligned}$$

$$A_4 = \tfrac{1}{4}\left[3\left(x_d^4 + y_d^4 + z_d^4\right) - 1\right] - \tfrac{1}{12}\left[\delta_x\left(3x_d^4 + 6y_d^2z_d^2 - 1\right) + \delta_y\left(3y_d^4 + 6x_d^2z_d^2 - 1\right) + \delta_z\left(3z_d^4 + 6y_d^2x_d^2 - 1\right)\right]$$

$$A_5 = \tfrac{1}{4}\left[3\left(x_d^4 + y_d^4 + z_d^4\right) - 1\right] + \tfrac{1}{12}\left[\delta_x\left(3x_d^4 + 6y_d^2z_d^2 - 1\right) + \delta_y\left(3y_d^4 + 6x_d^2z_d^2 - 1\right) + \delta_z\left(3z_d^4 + 6y_d^2x_d^2 - 1\right)\right], \quad [7]$$

where $\delta_j = (D_{jj} - D_{iso})/\sqrt{D_{iso}^2 - D^2}$ $(j = x, y, z)$, and (x_d, y_d, z_d) denotes coordinates of the NH unit vectors in the principal axis frame of the diffusion tensor. The (x_d, y_d, z_d) are related to the molecular frame unit vectors by the transformation

$$\begin{bmatrix} x_d \\ y_d \\ z_d \end{bmatrix} = R(\alpha, \beta, \gamma) \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad [8]$$

where the transformation matrix $R(\alpha, \beta, \gamma)$ is given by

$$R(\alpha, \beta, \gamma) = \begin{bmatrix} \cos(\alpha)\cos(\beta)\cos(\gamma) - \sin(\alpha)\sin(\gamma) & \sin(\alpha)\cos(\beta)\cos(\gamma) + \cos(\alpha)\sin(\gamma) & -\sin(\beta)\cos(\gamma) \\ -\cos(\alpha)\cos(\beta)\sin(\gamma) - \sin(\alpha)\cos(\gamma) & -\sin(\alpha)\cos(\beta)\sin(\gamma) + \cos(\alpha)\cos(\gamma) & \sin(\beta)\sin(\gamma) \\ \cos(\alpha)\sin(\beta) & \sin(\alpha)\sin(\beta) & \cos(\beta) \end{bmatrix}, \quad [9]$$

where α, β, and γ are the Euler angles relating the principal axis frame of the diffusion tensor to the molecular frame. In the case of an axially symmetric diffusion tensor (symmetric top), we have $D_{zz} = D_\parallel$ and $D_{xx} = D_{yy} = D_\perp$ and Eq. [5] transforms to

$$J(\omega) = \frac{2}{5}\sum_{i=1}^{3} \frac{A_i^{ax} D_i^{ax}}{\left(D_i^{ax}\right)^2 + \omega^2} \quad [10]$$

with

$$\begin{aligned} D_1^{ax} &= (5D_\perp + D_\parallel) \\ D_2^{ax} &= (2D_\perp + 4D_\parallel) \\ D_3^{ax} &= 6D_\perp. \end{aligned} \quad [11]$$

The coefficients A_i^{ax} are given by

$$\begin{aligned} A_1^{ax} &= 3z_d^2\left(1 - z_d^2\right) \\ A_2^{ax} &= \frac{3}{4}\left(1 - z_d^2\right)^2 \\ A_3^{ax} &= \frac{1}{4}\left(3z_d^2 - 1\right)^2. \end{aligned} \quad [12]$$

Effective Diffusion Constant

A simpler, alternative approach is to define an effective diffusion constant for a given NH vector in the following way:

$$D_{eff} = \frac{\omega_N}{6}\sqrt{\frac{\rho}{1-\rho}}. \quad [13]$$

D_{eff} is an apparent isotropic diffusion constant for a given NH vector, and is given by $D_{eff} = 1/15\,J(0)$. Using Eq. [5], D_{eff} can be expressed as

$$D_{eff} = \frac{1}{6\sum_{i=1}^{5} A_i/D_i}. \quad [14]$$

D_{eff} takes the following form when the NH vector in question points along the ith $(i = x, y, z)$ principal axis of the diffusion tensor,

$$D_{eff}^i = \frac{2D^2}{D_{ii} + D_{iso}}, \quad [15]$$

where D^2 and D_{iso} have been defined above. For example, when the NH vector points along the x axis of the diffusion tensor, Eq. [15] transforms to

$$D_{eff}^x = \frac{2(D_{xx}D_{yy} + D_{xx}D_{zz} + D_{yy}D_{zz})}{4D_{xx} + D_{yy} + D_{zz}}. \quad [16]$$

When the anisotropy in the diffusion tensor is small, we may express the principal elements in the following way,

$$\begin{aligned} D_{xx} &= D \\ D_{yy} &= D(1 + \varepsilon_1) \\ D_{zz} &= D(1 + \varepsilon_2), \end{aligned} \quad [17]$$

where ε_1 and ε_2 are small, positive dimensionless quantities, with $\varepsilon_1 < \varepsilon_2$. Using Eq. [17] in Eq. [16] we have

$$D_{eff}^x = 2D\left[\frac{3 + 2(\varepsilon_1 + \varepsilon_2) + \varepsilon_1\varepsilon_2}{6 + (\varepsilon_1 + \varepsilon_2)}\right]. \quad [18]$$

For small anisotropy, we have $\varepsilon_1 + \varepsilon_2 \ll 6$ and $\varepsilon_1\varepsilon_2 \approx 0$. Thus, Eq. [18] transforms to

$$\begin{aligned} D_{eff}^x &= D\left(1 + \frac{\varepsilon_1 + \varepsilon_2}{2}\right) = \frac{D_{yy} + D_{zz}}{2} \\ &= \frac{3D_{iso} - D_{xx}}{2}. \end{aligned} \quad [19]$$

The effective diffusion constant along a given axis is then an average of the principal values along the two orthogonal axes, and we have

$$D^i_{eff} = \frac{3D_{iso} - D_{ii}}{2}, \qquad [20]$$

where $i = x, y, z$. Thus, the D^i_{eff} can be considered to be the components of a second-rank tensor **Q** represented by (*2, 3*)

$$\mathbf{Q} = \frac{3D_{iso}\mathbf{E} - \mathbf{D}}{2}, \qquad [21]$$

where **E** is the identity tensor.

For a given NH vector, the effective diffusion constant can be expressed as (*27*)

$$D_{eff} = D_{iso} + \frac{2}{3}\sum_{ij} Q_{ij}S_{ij} \qquad [22]$$

with $S_{ij} = (3ij - \delta_{ij})/2$ $(i, j = x, y, z)$ and δ_{ij} is the Kronecker delta symbol. This result is obtained from the fact that a second-rank tensor can be decomposed into a scalar (rank 0) component, a pseudo-vector component (rank 1), and a symmetric tensor of rank 2. The scalar D_{eff} can then be obtained taking a scalar product of **Q** and **S**. The components of **S** comprise of the coordinates of the NH unit vector in the molecular frame. The first term on the right-hand side of Eq. [22] comes from the rank 0 (isotropic) component and the second term comes from the contraction of the rank 2 component. The rank 1 component is zero since **D** (and therefore **Q**) is a symmetric tensor (*28*). Simplifying Eq. [22], we obtain (*2, 3*)

$$\begin{aligned} D_{eff} = {} & Q_{xx}x^2 + Q_{yy}y^2 + Q_{zz}z^2 + 2Q_{xz}xz \\ & + 2Q_{xy}xy + 2Q_{yz}yz. \end{aligned} \qquad [23]$$

D_{eff} has transformation properties similar to residual dipolar couplings measured for proteins in orientated systems (*8–10*). Clore *et al.* have proposed previously that the R_2/R_1 ratio (which determines D_{eff}, cf. Eq. [13]) displays a dipolar powder-pattern behavior for a uniformly distributed set of NH vectors (*7*). Though an analytical proof is very difficult in the general case, it may be obtained quite easily in the case of small anisotropy. Equation [23] can be rewritten in polar coordinates as

$$\begin{aligned} D_{eff} = {} & D_{iso} + \frac{D_{yy} - D_{xx}}{4}\sin^2\theta\cos(2\varphi) \\ & - \frac{2D_{zz} - D_{xx} - D_{yy}}{12}(3\cos^2\theta - 1). \end{aligned} \qquad [24]$$

We define three new variables, $y = (D_{zz} + D_{yy} - 2D_{eff})/(D_{zz} - D_{xx})$, $f = (D_{yy} - D_{xx})/(D_{zz} - D_{xx})$ and $z = -\cos(\theta)$. Using these in Eq. [24], we have

$$y = f\sin^2\varphi + (1 - f\sin^2\varphi)z^2 \qquad [25]$$

with $0 \le y \le 1$. Equation [25] has the same form as the normalized frequency for a dipolar lineshape as outlined in Appendix I of Slichter (*29*). Thus, following the same procedure as that outlined by Slichter, one can derive an expression for the distribution function $I(y)$ as a function of y. We will reproduce only the final expression here; for details, the reader is referred to Appendix I of Slichter:

$$\begin{aligned} I(y) &= \frac{4}{\sqrt{y(1-f)}}K\left[\frac{f(1-y)}{(1-f)y}\right] \quad \text{for } y > f \\ I(y) &= \frac{4}{\sqrt{f(1-y)}}K\left[\frac{y(1-f)}{(1-y)f}\right] \quad \text{for } y < f. \end{aligned} \qquad [26]$$

Here K is the elliptic integral of the first kind and is given by

$$K(m) = \int_0^{\frac{\pi}{2}} \frac{dx}{\sqrt{1 - m\sin^2 x}}. \qquad [27]$$

There is a discontinuity at $y = f$ in Eq. [26], which corresponds to the NH vector pointing along the y axis of the diffusion tensor. At this point the effective diffusion constant is given by $D_{eff} = (D_{xx} + D_{zz})/2$. This validates the suggestion of Clore *et al.* (*7*), who proposed that the maximum of the R_2/R_1 distribution function corresponds to the NH vectors pointing along the y axis of the diffusion tensor.

Singular Value Decomposition Approach

In the case where we have determined ρ and hence D_{eff} for n different NH vectors, we may write Eq. [23] in a matrix form as

$$\begin{bmatrix} D^1_{eff} \\ D^2_{eff} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ D^n_{eff} \end{bmatrix} = \begin{bmatrix} x_1^2 & y_1^2 & z_1^2 & 2x_1y_1 & 2y_1z_1 & 2x_1z_1 \\ x_2^2 & y_2^2 & z_2^2 & 2x_2y_2 & 2y_2z_2 & 2x_2z_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_n^2 & y_n^2 & z_n^2 & 2x_ny_n & 2y_nz_n & 2x_nz_n \end{bmatrix} \begin{bmatrix} Q_{xx} \\ Q_{yy} \\ Q_{zz} \\ Q_{xy} \\ Q_{yz} \\ Q_{xz} \end{bmatrix}. \qquad [28]$$

Equation [28] can be written in symbolic form as

$$\vec{\mathbf{D}}_{eff} = \mathbf{A}\vec{\mathbf{Q}}, \qquad [29]$$

where $\vec{\mathbf{D}}_{eff}$ is an n-dimensional column vector, $\mathbf{A}$ is an n-by-6 matrix, and $\vec{\mathbf{Q}}$ is a six-dimensional column vector composed of the elements of the $\mathbf{Q}$ tensor in Eq. [21]. The elements of $\vec{\mathbf{Q}}$ and hence those of $\mathbf{D}$ can be obtained by inversion of Eq. [29]. The pseudo-inverse of $\mathbf{A}$ can be written in the following way (*30*),

$$\mathbf{A}^{-1} = \mathbf{V} \cdot diag\left(\frac{1}{s_j}\right) \cdot \mathbf{U}^T, \qquad [30]$$

where $\mathbf{U}$ is an n-by-6 column-orthogonal matrix, $\mathbf{V}$ is a 6-by-6 orthogonal matrix, and s_j are the singular values of $\mathbf{A}$. In the case where the matrix $\mathbf{A}$ contains redundant data, which may occur when two or more vectors are nearly parallel, it is singular. The singular values, s_j, corresponding to those rows are zero, leading to instabilities in Eq. [30]. This instability is characterized by the so-called condition number of the matrix (*30*), which is the absolute ratio of the largest to the smallest singular value of the matrix. The condition number of a matrix is infinity when a singular value is 0. This problem is solved by setting $1/s_j$ to zero for the problematic rows. This corresponds to removal of the corresponding rows from the $\mathbf{A}$ matrix, which is justified since this amounts to removal of redundant data. In the present case this operation is performed until the condition number of the matrix is less than 5. Using Eq. [30] in Eq. [29], we obtain the elements of $\vec{\mathbf{Q}}$ and we can reconstruct the $\mathbf{Q}$ matrix and hence the $\mathbf{D}$ matrix (see Eq. [21]) which is then diagonalized to yield the eigenvalues of the diffusion tensor and the orientation of the principal axes with respect to the molecular frame.

However, as was mentioned previously, Eq. [23] is not applicable in the general case and holds only for systems with a small degree of anisotropy. Application of Eq. [23] to systems with substantial anisotropy may lead to errors in determination of the diffusion tensor (see below). The anisotropy of the diffusion tensor is characterized by two parameters—the first, which characterizes the anisotropy, is given by

$$\zeta = \frac{2D_{zz}}{D_{xx} + D_{yy}}. \qquad [31a]$$

The second, which characterizes the rhombicity, is given by

$$\eta = \frac{3}{2}\frac{(-1)^n(D_{yy} - D_{jj})}{D_{zz}}\frac{\zeta}{\zeta - 1}, \qquad [31b]$$

where $j = z$ and $n = 1$ for an oblate ellipsoid ($D_{zz} \geq D_{yy} > D_{xx}$); $j = x$ and $n = 2$ for a prolate ellipsoid ($D_{zz} > D_{yy} \geq D_{xx}$); $\eta = 0$ for an axially symmetric diffusion tensor; while $\zeta = 1$, $\eta = 0$ for an isotropic diffusion tensor. In order to characterize the errors in the tensor determination from the approximate method, we define two more parameters. The error in the principal element determination may be characterized by

$$\varepsilon_d = 100\sqrt{\sum_{i=x,y,z}\frac{1}{3}\left[\frac{\left(D_{ii}^{act} - D_{ii}^{calc}\right)}{D_{ii}^{act}}\right]^2}. \qquad [32]$$

The error is 0 when there is perfect match between the actual and calculated values. We can formulate a similar definition for the error in the orientation of the diffusion tensor. This is given by

$$\varepsilon_a = 100\left[1 - \frac{Trace|R^{act}(\alpha, \beta, \gamma)R^{calc}(\alpha, \beta, \gamma)^T|}{3}\right]. \qquad [33]$$

The absolute values of the rotation matrices are used in Eq. [30] because inspection of Eqs. [7] and [12] reveals that the coordinates occur as even powers and hence the ratio in Eq. [4] is invariant to an even number of successive reflections about the x–y, x–z or y–z planes (*31*). The solution space of Euler angles is then fourfold degenerate for a right-handed coordinate system; i.e., (α, β, γ), $(\alpha, \beta, 180 + \gamma)$, $(180 + \alpha, 180 - \beta, 180 - \gamma)$, and $(180 + \alpha, 180 - \beta, 360 - \gamma)$ are all valid solutions. ε_a is 0 when the actual and calculated orientations are the same, and 100 when the two orientations are orthogonal to each other. Figures 1a and 1b show the ε_d and ε_a as functions of ζ and η calculated from simulated relaxation data (R_1, R_2, and NOE) for a set of 50 uniformly distributed vectors with a D_{zz} value of 4.0×10^7 s^{-1} at 600.13 MHz. The Euler angles relating the diffusion tensor frame to the molecular frame are given by $\alpha = \beta = \gamma = 45^\circ$. The angular errors remain very small (much smaller than 1%) even for highly anisotropic systems while the error in the principal components becomes large for highly anisotropic diffusion tensors. Thus the approximate approach is accurate in determining the orientation of the diffusion tensor even for highly anisotropic systems. Also shown in Fig. 1c is $\varepsilon_{D_{xx}}$, which is the percentage error in D_{xx} (the largest in this case). Figure 1d shows the RMS error (ε_{iso}) in the isotropic value of the diffusion tensor as a percentage of its actual value. As is evident, the error in isotropic value is smaller than ε_d due to the self-compensatory nature of the errors in the individual values of the principal elements of the diffusion tensor.

To summarize, the approximate approach solves Eq. [28] for the elements Q_{ij} of the $\mathbf{Q}$ tensor, followed by the reconstruction of the $\mathbf{Q}$ tensor and its diagonalization to obtain the principal elements and orientation. The principal elements of the $\mathbf{D}$ tensor can be obtained from the principal elements of the $\mathbf{Q}$ tensor using Eq. [21]. We observed that the Euler angles obtained using the approximate approach are more accurate than the principal values of the tensor. We believe that the accuracy of the approximate approach in predicting the orientation of the $\mathbf{Q}$ tensor (and therefore the $\mathbf{D}$ tensor) lies in the second step of the procedure, i.e., reconstruction and diagonalization of the $\mathbf{Q}$ tensor. To illustrate this, let us adopt a perturbation theory approach. When the anisotropy is small, both the principal values and the orientation of the $\mathbf{Q}$ tensor are predicted accurately.

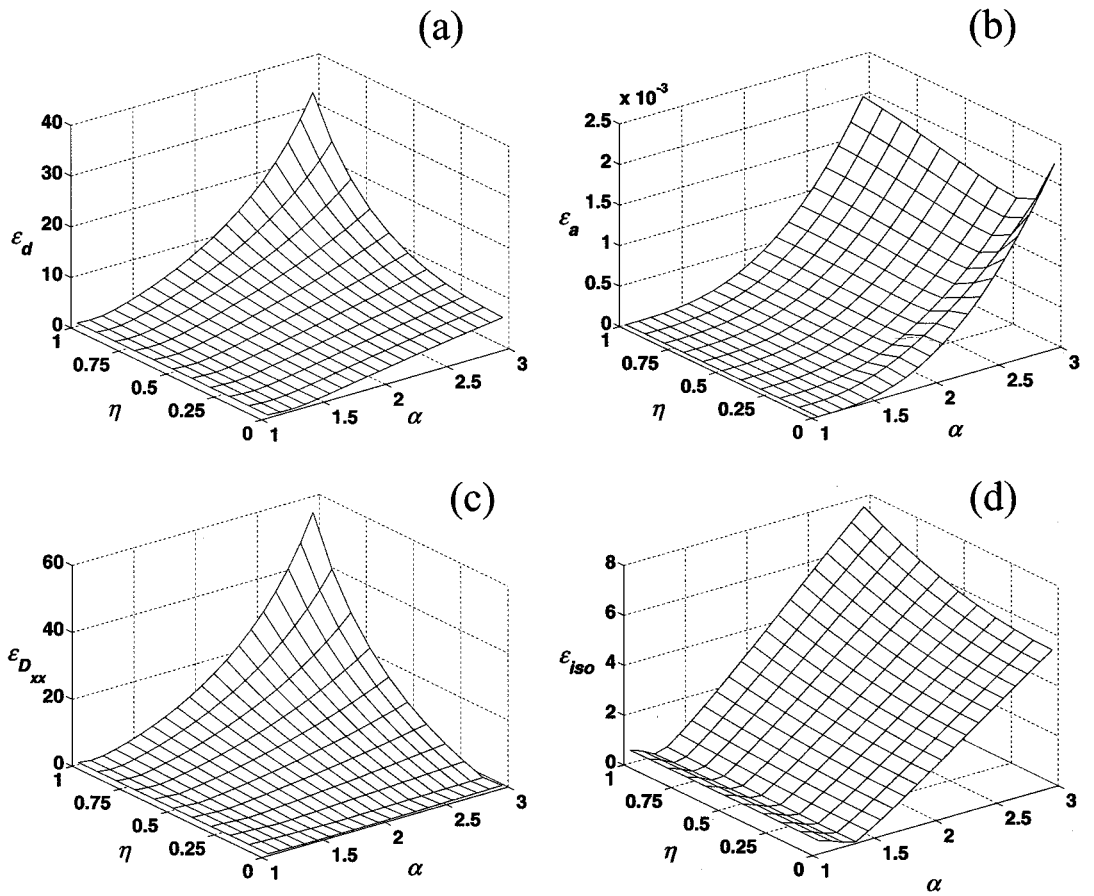

FIG. 1. Plots of the errors in the principal values of the diffusion tensor ε_d (a) and those in the orientation ε_a (b), estimated using the SVD approach, as a function of the anisotropy (ζ) and rhombicity (η) of the diffusion tensor. The errors have been calculated from simulated relaxation data (R_1, R_2, and NOE) for a set of 50 uniformly distributed vectors with a D_{zz} value of 4.0×10^7 s^{-1} at 600.13 MHz. The Euler angles relating the diffusion tensor frame to the molecular frame are given by $\alpha = \beta = \gamma = 45^\circ$. Also shown are the error in the D_{xx} principal component (c) and the isotropic diffusion constant (1/3)Tr(**D**) (d).

As the anisotropy increases, the error in prediction of the correct tensor, we call $\mathbf{Q}^0$, by the approximate method that yields tensor $\mathbf{Q}$ increases. In this case, $\mathbf{Q}$ may be treated as a perturbation on the result obtained from the exact analysis ($\mathbf{Q}^0$). Let the correct tensor be characterized by the eigenvalues Q^0_{ii} and eigenvectors (which represent its orientation) λ^0_i. Thus, $\mathbf{Q} = \mathbf{Q}^0 + \mathbf{Q}'$ and the first-order correction to the eigenvalues is given by $Q^1_{ii} = Q'_{ii}$ which are the diagonal elements of the perturbation $\mathbf{Q}'$. The absolute values of Q^1_{ii} show a monotonic increase as the magnitude of the anisotropy (and the norm of $\mathbf{Q}'$) increases since $\mathbf{Q}$ shows larger and larger deviations from $\mathbf{Q}^0$. This correction to the eigenvalues can be considered to represent the errors in predicting the correct eigenvalues by the approximate method. The eigenfunctions of $\mathbf{Q}$ can be written as a linear combination of the eigenfunctions of $\mathbf{Q}^0$ as

$$\lambda_i = \lambda^0_i + \sum_{n=1}^{\infty} \sum_{j \neq i} a^n_{ij} \lambda^0_j. \qquad [34]$$

The first-order correction, represented by the coefficient a^1_{ij}, is given by

$$a^1_{ij} = \frac{Q'_{ij}}{Q^0_{jj} - Q^0_{ii}}. \qquad [35]$$

As the anisotropy increases, the increase in the numerator is mirrored by the increase in the denominator (as the anisotropy increases, the differences between the eigenvalues of the tensor $\mathbf{Q}^0$ become more pronounced). Thus the λ_{ii} do not show a large deviation from the λ^0_{ii}; i.e., the orientation predicted by the approximate analysis does not deviate much from the actual orientation. This proof is not rigorous, as the perturbation theory approach does not hold when the anisotropy is very large; i.e., the magnitude of $\mathbf{Q}'$ becomes comparable to that of $\mathbf{Q}^0$.

Thus, the singular value decomposition approach can be used to restrict the angular space spanning the angles α, β, and γ in a subsequent, more accurate search for the diffusion tensor using the exact equations represented by Eqs. [4]–[9] (or Eqs. [8]–[12] in the axially symmetric case), thus greatly reducing computational time as opposed to a grid search over the entire

three-dimensional space of Euler angles without compromising the generality of the approach.

Quality Factor for the Diffusion Tensor

In order to estimate the quality of the diffusion tensor calculated from experimental data, it is necessary to provide a measure of how well the relaxation data calculated using the estimated diffusion tensor agree with the experimental data. The effective diffusion constant D_{eff} shown in Eq. [23] can be used as a measure of the quality of the estimated diffusion tensor. Following the definition used by Clore and Garrett (*32*) for the quality factor for residual dipolar couplings (the similarity between D_{eff} and dipolar couplings has been demonstrated above, when the anisotropy is not large), we define a quality factor R by

$$R = \sqrt{\frac{\left\langle \left(D_{eff}^{obs} - D_{eff}^{calc}\right)^2 \right\rangle}{2\left\langle \left(D_{eff}^{obs} - \left\langle D_{eff}^{obs} \right\rangle\right)^2 \right\rangle}}, \qquad [36]$$

where the brackets imply an average over all the available NH vectors. A quality factor defined in this form is expected to be more stable to the completeness of the data set than the χ^2 value obtained from a nonlinear least-squares fit. In the case of an infinitely large set of uniformly distributed NH vectors, $\langle (D_{eff}^{obs})^2 \rangle = \langle (D_{eff}^{calc})^2 \rangle$ and $\langle D_{eff}^{obs} \rangle = \langle D_{eff}^{calc} \rangle$, making $R = 1$. In this case, the standard deviation of the distribution is expressed analytically (for small anisotropy) as

$$\left\langle \left(D_{eff}^{obs} - \left\langle D_{eff}^{obs} \right\rangle\right)^2 \right\rangle = \frac{1}{15}\left[\frac{(2D_{zz} - D_{xx} - D_{yy})^2}{12} + \frac{(D_{yy} - D_{xx})^2}{4}\right]. \qquad [37]$$

For the case of an infinite set of uniformly distributed NH vectors, R assumes an indeterminate form (0/0) for an isotropic diffusion tensor as is evident from Eq. [36] and Eq. [37]. At this limit, the quality factor introduced here is invalid, as is the quality factor for residual dipolar couplings introduced by Clore and Garrett (*32*). A similar quality factor for relaxation data has been suggested recently by de Alba *et al.* (*33*).

APPLICATIONS

Computer Simulation

Table 1 shows the results of the application of the approximate method to two different random distributions of 50 vectors for a fully anisotropic, an axially symmetric, and an isotropic diffusion tensor. Relaxation data were generated for these distributions using different sets of principal values and orientations of the diffusion tensors. The random errors were 1, 2, and 5% in the R_1, R_2, and NOE measurements, respectively. The errors in the values of the principal elements of the diffusion tensor as well as the tensor orientations were obtained from 5000 Monte Carlo simulations using the random error in the relaxation data. In the case of the fully anisotropic diffusion tensor, the orientation of the principal axis frame in the molecular frame is shown in Fig. 2a. In this case, all three Euler angles α, β, and γ are well defined. In the case of the axially symmetric diffusion tensor (Fig. 2b), the angle γ is arbitrary since there is only one unique axis, the z axis, while the x and y axes have no unique orientation. There is no unique orientation for any of the axes for an isotropic diffusion tensor as is shown in Fig. 2c. In the case where the approximate analysis yields uniquely defined x, y, and z directions for the diffusion tensor, the mean and standard deviation of α, β and γ are determined from the distribution using Eq. [9]. The distribution of the Euler angles α, β, and γ corresponding to the orientation of the diffusion tensor axes for Fig. 2a is shown in Fig. 2d. The values of the diffusion tensor

TABLE 1

Comparison of the Input (i) and Calculated (c) Values Using the Described SVD Approach on Simulated Relaxation Data for a Set of 50 Uniformly Distributed NH Vectors[a]

Type	τ_c(i)	D_{xx}(i)	D_{xxa}(c)	D_{yy}(i)	D_{yya}(c)	D_{zz}(i)	D_{zza}(c)	α(i)	α(c)	β(i)	β(c)	γ(i)	γ(c)
1. Full	16.67	0.80	0.81 ± 0.01	1.00	0.99 ± 0.01	1.20	1.19 ± 0.01	45	43.1 ± 2.0	45	46.8 ± 1.5	45	43.5 ± 2.2
2. Axial	16.67	0.90	0.89 ± 0.01	0.90	0.92 ± 0.01	1.20	1.19 ± 0.01	45	47.8 ± 1.8	45	45.7 ± 1.4	45	U[b]
3. Iso	16.67	1.0	0.98 ± 0.01	1.0	1.00 ± 0.01	1.0	1.02 ± 0.01	45	U[b]	45	U[b]	45	U[b]
4. Full	8.33	1.0	0.91 ± 0.02	2.0	2.04 ± 0.02	3.0	2.90 ± 0.02	25	24.4 ± 1.1	45	44.7 ± 1.2	60	60.6 ± 1.0
5. Axial	8.33	1.5	1.46 ± 0.02	1.5	1.52 ± 0.02	3.0	2.91 ± 0.02	25	27.5 ± 1.0	45	45.8 ± 0.7	60	U[b]
6. Iso	8.33	2.0	1.96 ± 0.02	2.0	2.00 ± 0.01	2.0	2.05 ± 0.02	25	U[b]	45	U[b]	60	U[b]

[a] The τ_c(i) values are in ns, and are $1/(2\mathrm{Tr}[\mathbf{D}(\mathrm{i})])$. The τ_c(c) (not shown) can be calculated from the elements of **D**(c). The values of the diffusion tensor elements are in units of 10^7 s^{-1}. The values of the Euler angles (α, β, γ) are in degrees. The (i) and (c) refer to the input and extracted values for the various elements. The errors in the extracted values have been calculated using 5000 Monte Carlo steps using the noise in the relaxation data (1% for R_1, 2% for R_2, and 5% for NOE).

[b] For the axially symmetric diffusion tensor, the x and y axes of the diffusion tensor and hence the angle γ are arbitrary. Thus, since the SVD approach fits a fully anisotropic model to the data, all values of γ in the 0°–360° range are obtained as solutions. This is indicated by the letter U, signifying a uniform spread in the calculated values. In the case of the isotropic diffusion tensor, the relaxation rates are completely independent of the orientation of the NH vector and, hence, all three axes of the diffusion tensor and hence the angles α, β, and γ are arbitrary.

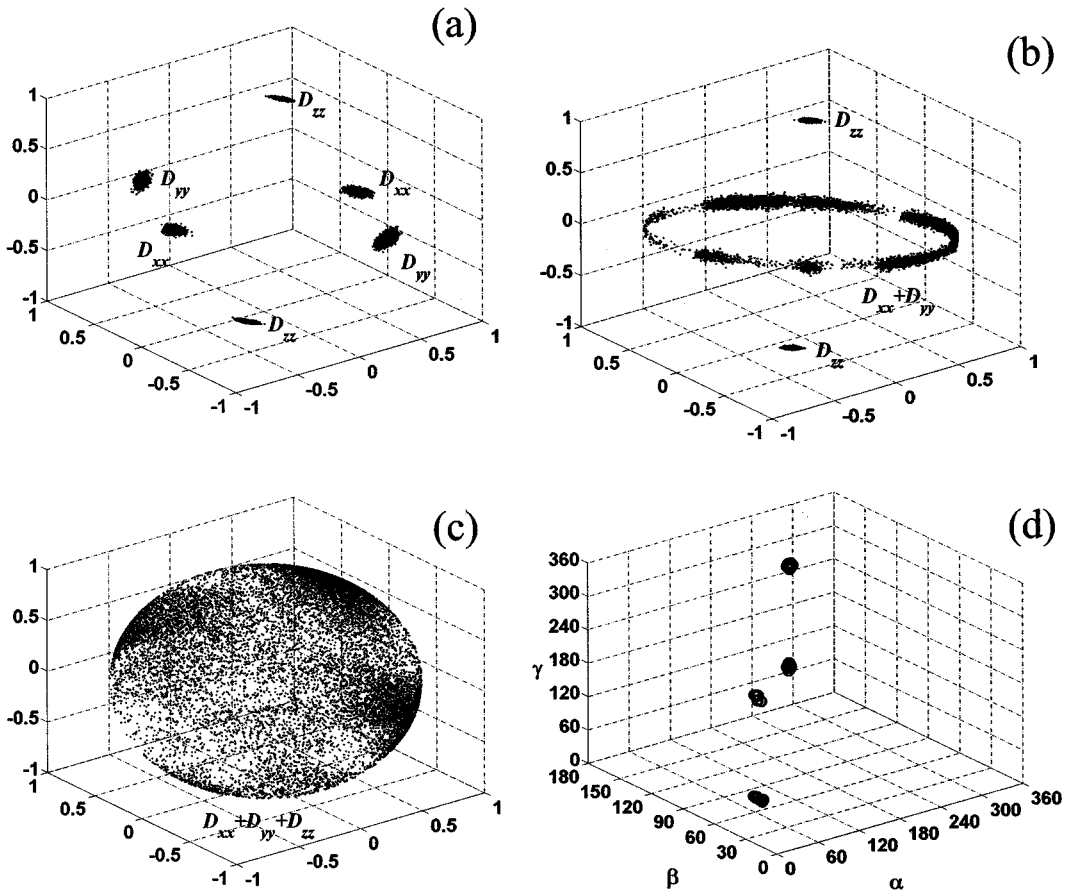

FIG. 2. Distribution of eigenvectors for a fully anisotropic (a), axially symmetric (b), and isotropic (c) diffusion tensor. In the case of the fully anisotropic diffusion tensor, all three axes are well defined. In the axially symmetric case, only the *z* axis is unique, while the *x* and *y* axes of the diffusion tensor are arbitrary. In the isotropic case, all three axes are arbitrary. Also shown (d) is the distribution of Euler angles corresponding to the distribution of (a), which is the fully anisotropic case. It is evident that all three Euler angles α, β, and γ are well defined and there are four degenerate solutions (in a right-handed coordinate system).

obtained by this analysis (D_{xxa}, D_{yya}, D_{zza}) are then used as starting values in a simplex search over a three-dimensional grid of the Euler angles α, β, and γ (1° steps) using the exact equations given by Eqs. [4]–[9]. The angular space is constrained to lie between $[\alpha_{av} \pm K_1\Delta\alpha, \beta_{av} \pm K_2\Delta\beta, \gamma_{av} \pm K_3\Delta\gamma]$, where X_{av} and ΔX ($X = \alpha, \beta, \gamma$) are the average values and standard deviations of the Euler angles obtained from the approximate analysis.

The K_i ($i = 1, 2, 3$) are determined in the following way—we start with initial K_i values of 2.0 and perform the calculation. If the boundary of the restricted space of an Euler angle is hit on two successive steps of a calculation, the K_i value corresponding to that particular Euler angle is increased by 0.5, thus widening the search space for that Euler angle. In all the cases we investigated, we used rather conservative final K_i values that varied from 2.0 to 5.0.

In the case where only the *z* axis is uniquely defined (this is assumed to be the case when the standard deviation of the angle $\gamma > 35°$), the fit to the fully anisotropic diffusion tensor using the exact approach is not performed but fits to the axially symmetric tensor are performed using Eqs. [8]–[12] constraining the angular space to lie between $[\alpha_{av} \pm K_1\Delta\alpha, \beta_{av} \pm K_2\Delta\beta, \gamma = 0]$ where the angles are determined from the last row of Eq. [9] setting $\gamma = 0$ and using $(D_{jja} + D_{kka})/2$ and D_{iia}, $(i, j, k = x, y, z)$ as starting values for the search (where D_{iia} is the unique value obtained from the approximate analysis). In the case where the approximate analysis indicates an isotropic tensor (when the standard deviation of the values of all three angles α, β, and $\gamma > 35°$) fits are performed to both the axially symmetric and the isotropic models with the search extending over the entire space of α and β angles in the case of the former fit using $[(D_{xxa} + D_{yya})/2, D_{zza}]$ as starting values. The above procedure is implemented in a MATLAB (*34*) package DIFFTENS shown schematically in Fig. 3.

It should be restated here that in the case when an axially symmetric model is fit to a diffusion tensor which is fully anisotropic, one obtains two minima which are similar in their χ^2 values—one corresponding to an oblate tensor and the other to a prolate tensor (*5*). We can, however, estimate whether the tensor is indeed closer to an oblate or prolate model by investigating the

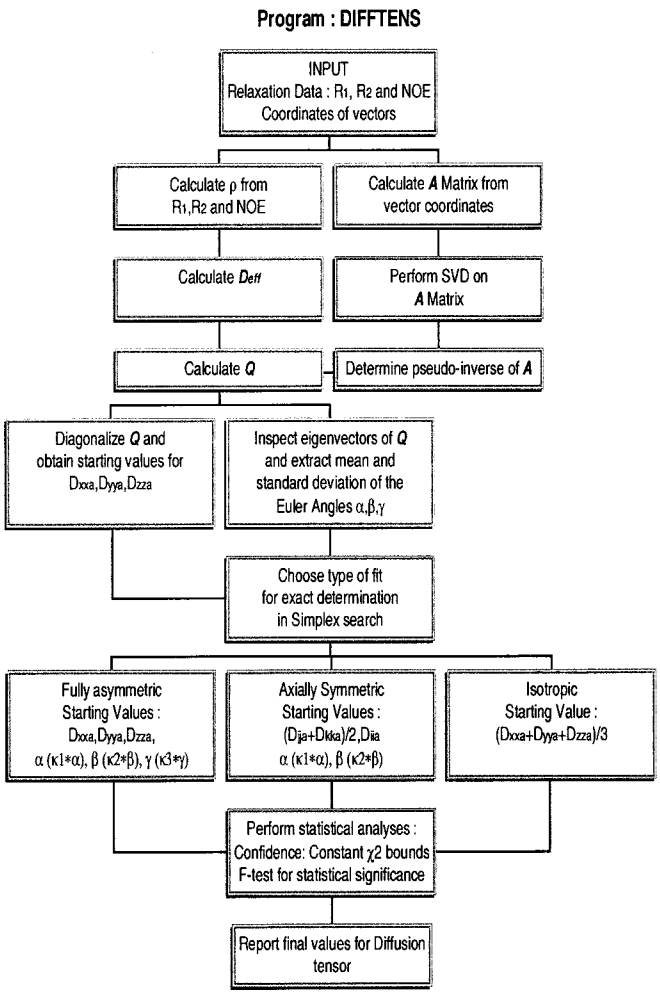

FIG. 3. Schematic representation of the MATLAB program DIFFTENS used to estimate the diffusion tensor. In the case of the fit to the axially symmetric model the starting values for the axially symmetric fit are chosen as D_{iia}, $(D_{jja} + D_{kka})/2$ $(i, j, k = x, y, z)$, where D_{iia} is the unique value and D_{jja} and D_{kka} are closest in value to each other as obtained from the SVD analysis.

principal values obtained from the approximate analysis. The approximate approach considers the fully anisotropic tensor and is thus free from the two-minima problem. Thus, in our subsequent fit to the axially symmetric model, only the relevant minimum is investigated.

The results of the application of the exact approach with the angular space restricted by the results of the SVD approach are shown in Tables 2a and 2b for the fully anisotropic diffusion tensors from Table 1 (Type 1 and 4). For very highly anisotropic systems, the values of the diffusion tensor elements themselves are not accurate, as noted above. In these cases, it might be necessary to use a few different, close values for the initial guesses for the diffusion tensor elements input into the simplex search, so as to avoid obtaining a false minimum. This could be done, for example, by multiplying the values of the tensor elements obtained by the SVD approach by different scaling factors. However, the angular restriction still holds, because the orientation of the diffusion tensor obtained by the SVD approach is still very accurate, as demonstrated above, and herein lies the real power of this hybrid approach.

Application to Real Systems

Application I—Test case: Ubiquitin. Tables 3a and 3b display the results of the above analysis applied to ^{15}N-labeled ubiqutin. The orientations of the NH vectors from the crystal structure (1ubq · pdb) were used. R_1, R_2, and NOE data were those provided by Tjandra *et al.* (*1*). Residues which exhibited significant exchange effects or rapid motion on the fast timescale (NOE < 0.7) were excluded from the analysis. In all, 55 residues were included in the analysis (these were the same as those utilized by Tjandra *et al.* (*1*)). It is interesting to note that the approximate analysis indicated a fully anisotropic tensor with the principal elements given by $D_{xxa} = 3.76 \pm 0.02 \times 10^7$ s^{-1}, $D_{yya} = 3.88 \pm 0.02 \times 10^7$ s^{-1}, and $D_{zza} = 4.46 \pm 0.02 \times 10^7$ s^{-1} ($\zeta = 0.17$ and $\eta = 0.28$), with the Euler angles relating the diffusion tensor frame to the molecular frame given by $\alpha = 48.9 \pm 2.7°$, $\beta = 38.7 \pm 1.6°$, and $\gamma = 79.1 \pm 6.9°$. The errors were determined from 2000 Monte Carlo simulations utilizing the experimental error in the relaxation rates. The exact analysis using the values of D_{xxa}, D_{yya}, and D_{zza} as starting values and constraining the Euler angles α, β, and γ to lie in the range $49 \pm 6°$, $39 \pm 5°$, and $79 \pm 14°$ yielded values of $D_{xx} = 3.76 \pm 0.17 \times 10^7$ s^{-1}, $D_{yy} = 3.88 \pm 0.17 \times 10^7$ s^{-1}, and $D_{zz} = 4.46 \pm 0.14 \times 10^7$ s^{-1} for the principal elements of the diffusion tensor and $\alpha = 48.0° \pm 14.0°$, $\beta = 39° \pm 9.0°$, and $\gamma = 85° \pm 37.0°$ for the Euler angles. The confidence limits for the principal elements were determined using the method

TABLE 2a
Results for Simulated Tensor 1—Final Results[a]

Model	D_{xx}	D_{yy}	D_{zz}	α	β	γ	τ_c	χ^2	F	P (%)	R^b
Full	0.81 ± 0.02	0.98 ± 0.02	1.19 ± 0.02	43	47	43	16.72	42.38	—	—	0.16
Axial	0.90 ± 0.01	0.90 ± 0.01	1.19 ± 0.02	49	47	—	16.77	153.16	57.49	$5.32e^{-11}$	0.31
Iso	0.97 ± 0.01	0.97 ± 0.01	0.97 ± 0.01	—	—	—	17.15	776.49	62.41	$3.02e^{-14}$	0.71

[a] The exact calculation was performed for the simulated tensor 1, shown in the first row of Table 1 using the above starting values for the tensor elements and using the results of the SVD analysis to restrict the space of Euler angles. The restricted angular space was $\alpha = 43 \pm 6°$, $\beta = 47 \pm 5°$, and $\gamma = 43 \pm 7°$.

[b] R has been calculated from Eq. [36].

TABLE 2b
Results for Simulated Tensor 4—Final Results[a]

Model	D_{xx}	D_{yy}	D_{zz}	α	β	γ	τ_c	χ^2	F	P (%)	R^b
Full	0.99 ± 0.05	2.01 ± 0.05	3.05 ± 0.06	24	46	61	8.27	70.87	—	—	0.09
Axial	1.46 ± 0.08	1.46 ± 0.08	2.85 ± 0.16	20	46	—	8.66	1188.67	347.01	$1.14e^{-25}$	0.43
Iso	1.82 ± 0.03	1.82 ± 0.03	1.82 ± 0.03	—	—	—	9.17	3664.06	31.99	$2.60e^{-9}$	0.75

[a] The exact calculation was performed for the simulated tensor 4 shown in the fourth row of Table 1 using the above starting values for the tensor elements and using the results of the SVD analysis to restrict the space of Euler angles. The restricted angular space was $\alpha = 24 \pm 5^\circ$, $\beta = 46 \pm 5^\circ$, and $\gamma = 61 \pm 5^\circ$.

[b] R has been calculated from Eq. [36].

of constant χ^2 boundaries (*30*) (the 68.3% confidence limits are reported). The confidence limits of the estimated Euler angles were obtained from the widths of the contour maps of the target function from the covariance matrix calculated by utilizing the principal elements of the diffusion tensor obtained at the χ^2 minimum. Examples of portions of the α–β projection and the α–γ projection near the minimum of the 3D contour map are depicted in Figs. 4a and 4b. The value of χ^2 at the minimum was found to be 656.11. Performing the fit to the axially symmetric model yielded values of $D_\perp = 3.82 \pm 0.06 \times 10^7\ s^{-1}$ and $D_\parallel = 4.43 \pm 0.11 \times 10^7\ s^{-1}$ ($\zeta = 1.16$) for the principal elements of the diffusion tensor and $\alpha = 48.0^\circ \pm 15.0^\circ$ and $\beta = 39^\circ \pm 9.0^\circ$ for the Euler angles (Fig. 4c depicts a portion of the contour map of the target function at the χ^2 minimum) and there was a marginal increase in the value of χ^2 to 680.85. A statistical analysis of the significance of the results using the F test (*30*) yielded an F value of 0.93 and indicated that there was a 59.9% probability that the marginal change in the χ^2 by going from the axially symmetric to the fully asymmetric model was obtained by chance. However, the decrease in the χ^2 value in going from the isotropic to the axially symmetric model was found to be statistically significant (see Table 3b). This is in complete agreement with previously obtained results (*1, 3*). The entire calculation (excluding the detailed analysis of the errors in the Euler angles) required approximately 1 h on a Pentium III computer operating at 733 MHz. The slowest step in the calculation was, as expected, the grid search over the restricted subspace of Euler angles for the fully anisotropic and axially symmetric models.

Application II—SH(32) dual-domain construct of Abelson kinase complexed with a consolidated ligand. In the dual-domain construct of the SH2 and SH3 domains of Abelson kinase complexed with a consolidated ligand, the relative orientation of the two domains can be obtained by determination of the diffusion tensor individually for each of the two domains separately and then by aligning the two tensor axes (*6*). This approach is valid only for systems in which the two domains are relatively rigid with respect to each other, so that the relative mobility of the individual domains is limited. In an earlier publication, the change in relative orientation of the two domains on binding to a consolidated ligand was investigated assuming an axially symmetric diffusion tensor (*6*). An axially symmetric diffusion tensor, though a valid approximation in the unligated form, is not expected to hold in the ligated case. Here, we apply the fully anisotropic model to the ligated form of the SH(32) dual-domain system of Abelson kinase. As is shown by the results of Tables 4a–4d a fully asymmetric diffusion tensor is indicated for both domains (a P value of less than 5% is considered to be statistically significant). The rhombicity in the tensor is quite large with $\zeta = 1.14$ and $\eta = 0.68$ for the SH2 domain while the corresponding values for the SH3 domain are 1.29 and 0.95, respectively. The magnitude of the rhombicity can be attributed to the presence of the consolidated ligand on one face of the dual-domain construct. The principal components of the diffusion tensor as determined from the individual domains seem to differ slightly, being marginally larger in the SH3 domain. This indicates a small amount of relative motion between the two domains. This motion could be caused by some residual mobility of the two domains in the ligated state, as well as the on/off processes associated with the consolidated ligand binding at one or both sites. We utilized the orientation of the NH bond vectors for the individual domains in the Abl-SH(32) crystal structure (2abl · pdb). Structural studies on other SH2 domains indicate that ligand binding has no significant effect on the protein core (*35, 36*). A crystal structure is available for the Abl-SH3 domain complexed with ligand (1abo · pdb). The NH vector orientations obtained from this (1abo · pdb) structure produced no significant changes in the results for the SH3 domain. In both cases, only the core residues were used to calculate the diffusion tensor. In the case of the SH2 subdomain of the core residues, the residues 161, 178 exhibit conformational exchange and were excluded from the relaxation analysis. Further, the residues 153, 174, 176, 186, 191, 193, and 204 showed large errors in R_1 and R_2 values ($>10\%$) and were excluded (49 residues used). Table 4b shows the results of the analysis. In order to assess the quality of the estimated diffusion tensor, R values (from Eq. [36]) were estimated for the fully anisotropic, axially symmetric, and isotropic diffusion tensors. These values were found to be 0.61, 0.63, and 0.70 for the three models. The R values obtained on inclusion of the residues excluded from the tensor estimation due to large errors were 0.64, 0.66, and 0.70. Though, in this case, the values of the R factors were marginally higher, the fully anisotropic model produced a lower R factor than the two other cases. In

TABLE 3a
Results for Ubiquitin at 600 MHz—SVD Analysis

	D_{xxa}	D_{yya}	D_{zza}	α	β	γ
Results from SVD	3.76 ± 0.02	3.88 ± 0.02	4.46 ± 0.02	48.9 ± 2.7	38.7 ± 1.6	79.2 ± 6.8
Input into simplex	3.76	3.88	4.46	43–55	34–44	65–93

TABLE 3b
Results for Ubiquitin at 600 MHz—Final Results

Model	D_{xx}	D_{yy}	D_{zz}	α^a	β	γ	τ_c	χ^2	F	P (%)	R
Full	3.76 ± 0.17	3.88 ± 0.17	4.46 ± 0.14	48 ± 14	39 ± 9	85 ± 37	4.13	655.91	—	—	0.55
Axial	3.82 ± 0.06	3.82 ± 0.06	4.43 ± 0.11	48 ± 15	39 ± 9	—	4.14	680.86	0.93	59.93	0.55
Iso	4.06 ± 0.16	4.06 ± 0.16	4.06 ± 0.16	—	—	—	4.10	1139.90	11.46	$7.4e^{-03}$	0.71

[a] The definitions of the Euler angles used here are different from but consistent with those of Tjandra *et al.* (*1*).

TABLE 4a
Results for SH2 Domain for SH(32)/Ligand at 600 MHz—SVD Analysis

	D_{xxa}	D_{yya}	D_{zza}	α	β	γ
Results from SVD	1.36 ± 0.02	1.54 ± 0.02	1.66 ± 0.02	168.2 ± 4.2	81.5 ± 10.1	99.4 ± 8.3
Input into simplex	1.36	1.54	1.66	153–183	52–112	74–124

TABLE 4b
Results for SH2 Domain for SH(32)/Ligand at 600 MHz—Final Results

Model	D_{xx}	D_{yy}	D_{zz}	α	β	γ	τ_c	χ^2	F	P (%)	R
Full	1.34 ± 0.08	1.56 ± 0.08	1.65 ± 0.09	156 ± 16	61 ± 30	115 ± 12	10.96	58.33	—	—	0.61
Axial	1.45 ± 0.04	1.45 ± 0.04	1.66 ± 0.08	168 ± 33	91 ± 41	—	10.91	75.94	6.49	0.34	0.63
Iso	1.54 ± 0.01	1.54 ± 0.01	1.54 ± 0.01	—	—	—	10.85	90.63	2.90	4.52	0.71

TABLE 4c
Results for SH3 Domain for SH(32)/Ligand at 600 MHz—SVD Analysis

	D_{xxa}	D_{yya}	D_{zza}	α	β	γ
Results from SVD	1.28 ± 0.02	1.65 ± 0.02	1.89 ± 0.02	187.5 ±6.4	50.0 ± 1.5	178.6 ± 11.0
Input into simplex	1.28	1.65	1.89	168–208	40–60	148–208

TABLE 4d
Results for SH3 Domain for SH(32)/Ligand at 600 MHz—Final Results

Model	D_{xx}	D_{yy}	D_{zz}	α	β	γ	τ_c	χ^2	F	P (%)	R
Full	1.31 ± 0.17	1.62 ± 0.18	1.89 ± 0.12	197 ± 27	47 ± 7	166 ± 26	10.36	25.86	—	—	0.47
Axial	1.46 ± 0.07	1.46 ± 0.07	1.88 ± 0.13	203 ± 24	54 ± 11	—	10.44	36.27	3.83	4.02	0.58
Iso	1.60 ± 0.02	1.60 ± 0.02	1.60 ± 0.02	—	—	—	10.44	57.07	4.01	2.10	0.70

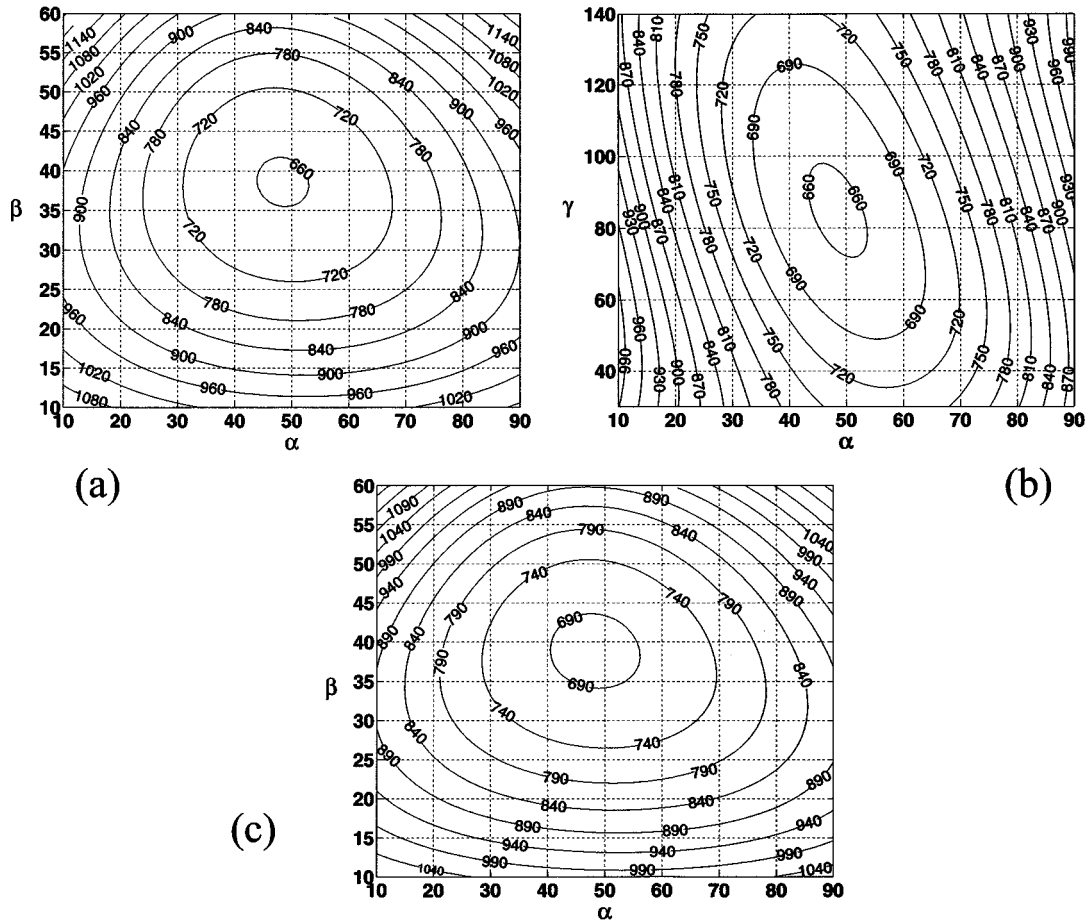

FIG. 4. Representation of a portion of the 3D contour map of the target function for a fully anisotropic diffusion tensor for ubiquitin at the χ^2 minimum. The α–β (a) and α–γ (b) projections are shown. Also shown is a portion of the corresponding 2D contour map (α–β) for the fit to the axially symmetric diffusion tensor (c). Only the region in the vicinity of one of the minima is depicted. There are four symmetry-related minima in the case of the fully anisotropic tensor and two symmetry-related minima in the case of the axially symmetric tensor. The 68.3% confidence limits for the Euler angles have been obtained from these contour maps.

the case of the SH3 subdomain, residues 87, 89, and 99 were excluded from the analysis due to conformational exchange and residues 66, 93, and 99, due to the large error in R_1 and R_2 values (25 residues used). The R factors obtained for the three models were 0.47, 0.58, and 0.70. Inclusion of residues 66, 93, and 99 yielded R values of 0.53, 0.58, and 0.70, respectively. Note that the above-mentioned residues with large experimental errors were included in the previous analysis of the SH(32) data (*6*). The quality factor analysis then justifies their exclusion in the present analysis, which yields results which are largely consistent with those obtained in earlier studies (see below) (*6*). The total computation time required was about $2\frac{1}{2}$ h for the SH2 and $3\frac{1}{2}$ h for the SH3 domain.

The relative orientation between the two domains can be obtained, for example, by transforming both domains into the frame of reference of the diffusion tensor. This can be represented mathematically by

$$r_d = R_{SH2} r_{SH2} = R_{SH3} r_{SH3}, \qquad [38]$$

where r_d, r_{SH2}, and r_{SH3} are the coordinates of an arbitrary vector in the principal axis frame of the diffusion tensor, the pdb frame of SH2, and the pdb frame of SH3, respectively. R_{SH2} and R_{SH3} are the transformation matrices which transform the individual pdb frames of SH2 and SH3 into the principal axis frame of the diffusion tensor. Thus, the Euler angles that transform the pdb frame of SH3 into the pdb frame of SH2 can be obtained from $R_T = R_{SH2}^{-1} R_{SH3}$. There are four valid solutions for the relative orientation of the two domains (in a right-handed coordinate frame) (*37*). However, some of this degeneracy may be resolved using chemical-bonding considerations such as the fact that the N-terminal part of the SH2 domain is linked to the C-terminal part of the SH3 domain. Further, the presence of the consolidated ligand makes certain conformations impossible. In the present case, of the four allowed solutions, only one puts the N-terminus of the SH2 domain close to the C-terminus of the SH3 domain (5.8 Å), allowing chemical ligation to occur with minimal reorientation of the loop region at the C-terminal end of the SH2 domain (as compared to the unligated case as derived in (*6*)). In the three other cases, the corresponding distances vary from 17.4 to 24.3 Å, making these impossible to achieve. Shown in

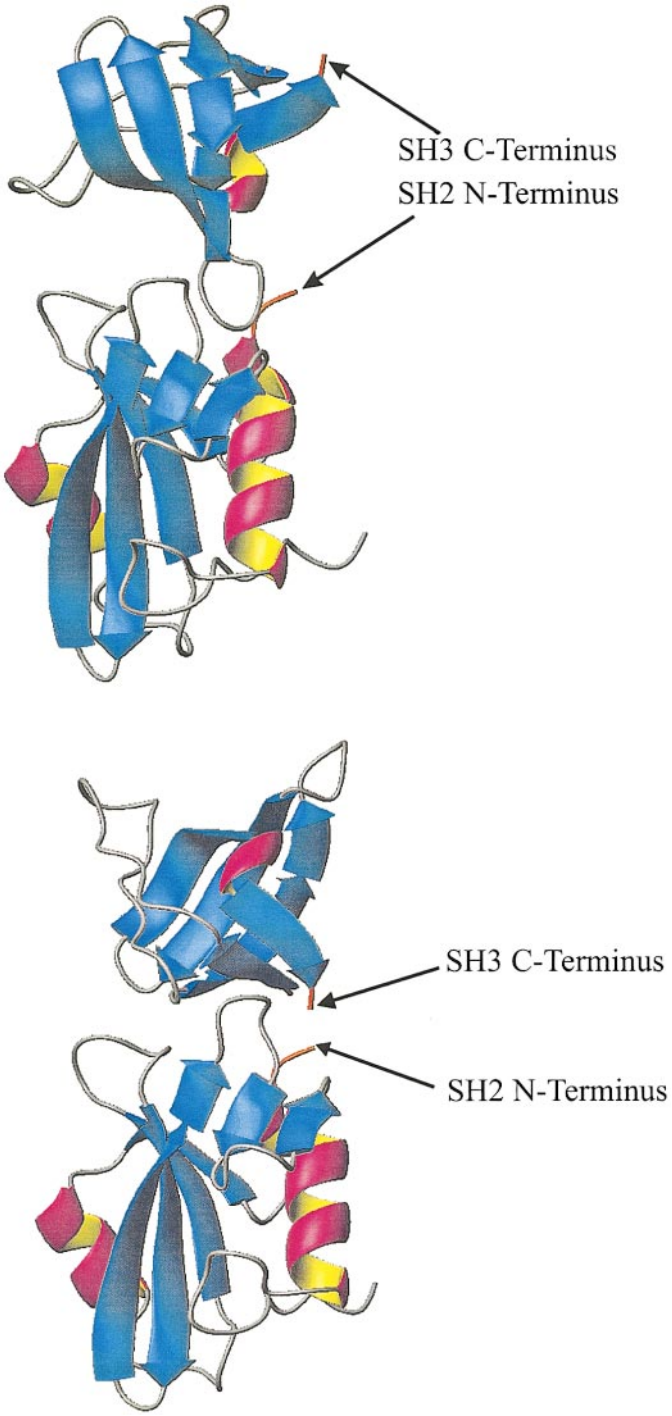

FIG. 5. Two of the possible orientations of the two domains of Abl-SH32. The C-terminus of the SH3 domain and the N-terminus of the SH2 domain, which are bonded, are shown shaded orange. The incorrect solution shown in the top panel has these two termini far (17.4 Å) from each other whereas in the correct solution (bottom panel) these are adjacent (5.8 Å).

Fig. 5 are two of the possible solutions. The top panel in Fig. 5 shows one of the incorrect solutions in which the two relevant termini (shaded orange) are separated by 17.4 Å. The correct solution shown in the bottom panel in Fig. 5 corresponds to a rotation by $\alpha = 316^\circ$, $\beta = 102^\circ$, and $\gamma = 251^\circ$ of the SH3 domain with respect to the SH2 domain (the individual rotations being $\alpha = 156^\circ$, $\beta = 61^\circ$, and $\gamma = 115^\circ$ for the SH2 domain and $\alpha = 17^\circ$, $\beta = 133^\circ$, and $\gamma = 194^\circ$ for the SH3 domain). The relative "bend" between the two domains structure is quite similar, within experimental uncertainity, to that obtained previously by Fushman *et al.* (*6*). However, in that case, the assumption of an axially symmetric diffusion tensor required the interdomain "twist" to be chosen empirically, allowing optimal binding of the consolidated ligand. The orientation of the two domains obtained here brings the binding sites on the two domains into a proper position to accommodate the consolidated ligand.

The moderately large errors in the determination of the Euler angles relating the molecular frame to the principal axis frame of the diffusion tensor for the individual SH3 and SH2 domains reflect on the quality of the relaxation data. This method for the determination of domain orientation is completely general and can yield very accurate results when the quality of relaxation data is high and the anisotropy of the system is large.

CONCLUSIONS

We have provided a simple and efficient way of obtaining the rotational diffusion tensor of proteins of known structure, from relaxation data, using a combination of approximate and exact methods, resulting in a significant reduction in computation time as opposed to a grid search over the entire three-dimensional space of Euler angles relating the molecular frame to the principal axis frame of the diffusion tensor without compromising the generality of the approach. The approach is implemented in a MATLAB package DIFFTENS, which can be obtained by e-mail at the following address: ghoser@mriris.rockefeller.edu. The utility of this approach has been demonstrated in the determination of the interdomain orientation between the SH3 and SH2 domains in a SH(32) dual-domain construct of Abelson kinase complexed with a consolidated ligand.

REFERENCES

1. N. Tjandra, S. E. Feller, R. W. Pastor, and A. Bax, Rotational diffusion anisotropy of human ubiquitin from ^{15}N NMR relaxation, *J. Am. Chem. Soc*. **117,** 12562–12566 (1995).
2. R. Bruschweiler, X. Liao, and P. E. Wright, Long-range motional restrictions in a multidomain zinc-finger protein from anisotropic tumbling, *Science* **268,** 886–889 (1995).
3. L. K. Lee, M. Rance, W. J. Chazin, and A. G. Palmer, III, Rotational diffusion anisotropy of proteins from simultaneous analysis of ^{15}N and ^{13}C alpha nuclear spin relaxation, *J. Biomol. NMR* **9,** 287–298 (1997).
4. V. Copie, Y. Tomita, S. K. Akiyama, S. Aota, K. M. Yamada, R. M. Venable, R. W. Pastor, S. Krueger, and D. A. Torchia, Solution structure and dynamics of linked cell attachment modules of mouse fibronectin containing the RGD and synergy regions: Comparison with the human fibronectin crystal structure, *J. Mol. Biol.* **277,** 663–682 (1998).
5. M. Blackledge, F. Cordier, P. Dosset, and D. Marion, Precision and uncertainty in the characterization of anisotropic rotational diffusion by ^{15}N relaxation, *J. Am. Chem. Soc.* **120,** 4538–4539 (1998).

6. D. Fushman, R. Xu, and D. Cowburn, Direct determination of changes of interdomain orientation on ligation: Use of the orientational dependence of ^{15}N NMR relaxation in Abl SH(32), *Biochemistry* **38,** 10225–10230 (1999).
7. G. M. Clore, A. M. Gronenborn, A. Szabo, and N. Tjandra, Determining the magnitude of the fully asymmetric diffusion tensor from heteronuclear relaxation data in the absence of structural information, *J. Am. Chem. Soc.* **120,** 4889–4890 (1998).
8. J. R. Tolman, J. M. Flanagan, M. A. Kennedy, and J. H. Prestegard, Nuclear magnetic dipole interactions in field-oriented proteins: Information for structure determination in solution, *Proc. Natl. Acad. Sci. USA* **92,** 9279–9283 (1995).
9. A. Bax and N. Tjandra, High-resolution heteronuclear NMR of human ubiquitin in an aqueous liquid crystalline medium, *J. Biomol. NMR* **10,** 289–292 (1997).
10. G. M. Clore, M. R. Starich, and A. M. Gronenborn, Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses, *J. Am. Chem. Soc.* **120,** 10571–10572 (1998).
11. G. M. Clore and A. M. Gronenborn, New methods of structure refinement for macromolecular structure determination by NMR, *Proc. Natl. Acad. Sci. USA* **95,** 5891–5898 (1998).
12. M. W. F. Fischer, J. A. Losonczi, L. J. Weaver, and J. H. Prestegard, Domain orientation and dynamics in multidomain proteins from residual dipolar couplings, *Biochemistry* **38,** 9013–9022 (1999).
13. M. Markus, R. Gerstner, D. Draper, and D. Torchia, Refining the overall structure and subdomain orientation of ribosomal protein S4 Delta 41 with dipolar couplings measured by NMR in uniaxial liquid crystalline phases, *J. Mol. Biol.* **292,** 375–387 (1999).
14. N. Skrynnikov, N. Goto, D. Yang, W. Choy, J. Tolman, G. Mueller, and L. Kay, Orienting domains in proteins using dipolar couplings measured by liquid-state NMR: Differences in solution and crystal forms of maltodextrin binding protein loaded with beta-cyclodextrin, *J. Mol. Biol.* **295,** 1265–1273 (2000).
15. D. Fushman, S. Cahill, and D. Cowburn, The main chain dynamics of the dynamin pleckstrin homology (PH) domain in solution: Analysis of ^{15}N relaxation with monomer/dimer equilibration, *J. Mol. Biol.* **266,** 173–194 (1997).
16. F. Cordier, M. Caffrey, B. Brutscher, M. A. Cusanovich, D. Marion, and M. Blackledge, Solution structure, rotational diffusion anisotropy and local backbone dynamics of *Rhodobacter captulatus* cytochrome c2, *J. Mol. Biol.* **281,** 341–361 (1998).
17. J. M. McDonnell, D. Fushman, C. L. Milliman, S. J. Korsmeyer, and D. Cowburn, Solution structure of the pro-apoptotic molecule BID: A structural basis for apoptotic agonists and antagonists, *Cell* **96,** 625–634 (1999).
18. D. Woessner, Nuclear spin relaxation in ellipsoids undergoing rotational Brownian motion, *J. Chem. Phys.* **37,** 647–654 (1962).
19. R. W. Broadhurst, C. H. Hardman, J. O. Thomas, and E. D. Laue, Backbone dynamics of the A-domain of HMG1 as studied by ^{15}N NMR spectroscopy, *Biochemistry* **34,** 16608–16617 (1995).
20. K. Huang, R. Ghose, J. M. Flanagan, and J. H. Prestegard, Backbone dynamics of the N-terminal domain in *E. coli* DnaJ determined by ^{15}N and ^{13}CO relaxation measurements, *Biochemistry* **38,** 10567 (1999).
21. P. Dosset, J. Hus, M. Blackledge, and D. Marion, Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data, *J. Biomol. NMR* **16,** 23–28 (2000).
22. M. Andrec, K. G. Inman, D. J. Weber, R. Levy, and G. Montelione, A Bayesian statistical method for the detection and quantification of rotational diffusion anisotropy from NMR relaxation data, *J. Magn. Reson.* **146,** 66–80 (2000).
23. N. A. Farrow, O. Zhang, A. Szabo, D. A. Torchia, and L. E. Kay, Spectral density function mapping using ^{15}N relaxation data exclusively, *J. Biomol. NMR* **6,** 153–162 (1995).
24. R. Ishima and K. Nagayama, Protein backbone dynamics revealed by quasi spectral density function analysis of amide N-15 nuclei, *Biochemistry* **34,** 3162–3171 (1995).
25. D. Fushman, N. Tjandra, and D. Cowburn, Direct measurement of ^{15}N chemical shift anisotropy in solution, *J. Am. Chem. Soc.* **120,** 10947–10952 (1998).
26. D. Fushman and D. Cowburn, Model-independent analysis of ^{15}N chemical shift anisotropy from NMR relaxation data. Ubiquitin as a test example, *J. Am. Chem. Soc.* **120,** 7109–7110 (1998).
27. A. W. Joshi, "Matrices and Tensors in Physics," 2nd ed., Wiley Eastern, New Delhi (1991).
28. **D** and **Q** are diagonal in the same reference frame.
29. C. P. Slichter, "Principles of Magnetic Resonance," 3rd ed., Springer Series in Solid State Physics 1, Springer-Verlag, Berlin (1990).
30. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Numerical Recipes in C," Cambridge Univ. Press, New York (1992).
31. This amounts to reversing the signs of a pair of rows of the transformation matrix while preserving the sign of the determinant (+1) and thus retaining its handedness. An odd number of reflections changes a right-handed coordinate frame to a left-handed one.
32. G. M. Clore and D. S. Garrett, R-factor, free R, and complete cross-validation for dipolar coupling refinement of NMR structures, *J. Am. Chem. Soc.* **121,** 9008–9012 (1999).
33. E. de Alba, J. L. Baber, and N. Tjandra, The use of residual dipolar coupling in concert with backbone relaxation rates to identify conformational exchange by NMR, *J. Am. Chem. Soc.* **121,** 4282–4283 (1999).
34. "MATLAB—The Language of Technical Computing Version 5," The MathWorks Inc., Natick (1999).
35. G. Waksman, S. E. Shoelson, N. Pant, D. Cowburn, and J. Kuriyan, Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: Crystal structures of the complexed and peptide-free forms, *Cell* **72,** 779–790 (1993).
36. C.-H. Lee, D. Kominos, S. Jacques, B. Margolis, J. Schlessinger, S. E. Shoelson, and J. Kuriyan, Crystal structures of peptide complexes of the N-terminal SH2 domain of the Syp tyrosine phosphatase, *Structure* **2,** 423–438 (1994).
37. There are 16 solutions in general, 4 corresponding to the Euler angle sets for each of the individual domains. However, this degeneracy can be decreased by a factor of 4 by realizing that only 4 unique solutions correspond to fixing one of the domains while allowing the other domain to assume its 4 allowed (degenerate) orientations. The remaining orientations can be generated from these 4 orientations by rotations of the entire dual-domain system.