Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Magnetic Resonance 168 (2004) 336–345
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Efficient and accurate determination of the overall rotational diffusion tensor of a molecule from ^{15}N relaxation data using computer program ROTDIF

Olivier Walker, Ranjani Varadan, and David Fushman*

Department of Chemistry and Biochemistry, Center for Biomolecular Structure and Organization, University of Maryland, College Park, MD 20742, USA

Received 26 August 2003; revised 19 February 2004
Available online 22 April 2004

Abstract

We present a computer program ROTDIF for efficient determination of a complete rotational diffusion tensor of a molecule from NMR relaxation data. The derivation of the rotational diffusion tensor in the case of a fully anisotropic model is based on a six-dimensional search, which could be very time consuming, particularly if a grid search in the Euler angle space is involved. Here, we use an efficient Levenberg–Marquardt algorithm combined with Monte Carlo generation of initial guesses. The result is a dramatic, up to 50-fold improvement in the computational efficiency over the previous approaches [Biochemistry 38 (1999) 10225; J. Magn. Reson. 149 (2001) 214]. This method is demonstrated on a computer-generated and real protein systems. We also address the issue of sensitivity of the diffusion tensor determination from ^{15}N relaxation measurements to experimental errors in the relaxation rates and discuss possible artifacts from applying higher-symmetry tensor model and how to recognize them.

Keywords: Rotational diffusion tensor; Rotational anisotropy; Orientation dependence of spin-relaxation; Software for diffusion tensor determination

1. Introduction

In the recent years there has been significant interest in the characterization of the overall hydrodynamic properties of biomacromolecules. Because the overall tumbling provides a dominant contribution to spin-relaxation rates, its accurate characterization is essential for the accuracy of the derived picture of protein dynamics [1–4]. In particular, a failure to take the rotational anisotropy into account could result in spurious conformational exchange motions. Other emerging applications include structure characterization in multidomain systems [5–8] and analysis of protein–ligand interactions [5,9] and protein association [10].

The information on the overall rotational diffusion of a molecule is encoded in the rates of nuclear spin relaxation that can be measured in various groups in a protein. Several experimental and analytical methods (reviewed in [9,11]) have been developed to determine the overall rotational diffusion tensor of a protein from nuclear spin-relaxation data [5,12–20]. Models for computer prediction of the diffusion tensor based on atomic structure of a molecule are also available [21,22].

In the most general case, the rotational diffusion tensor is characterized by six parameters: three principal values, D_{xx}, D_{yy}, D_{zz} (we assume $D_{xx} \leqslant D_{yy} \leqslant D_{zz}$), and the Euler angles α, β, and γ relating the principal axes frame of the tensor to the molecular frame (e.g., pdb file coordinates). Full characterization of a rotational diffusion tensor requires determination of all six components of the tensor. The tensor derivation from experimental data (e.g., ^{15}N relaxation) involves an optimization search in a six-dimensional space and could be time consuming. The problem is significantly simplified when a molecule can be approximated by a sphere or by an axially symmetric top. This, however, does not apply to the most general case. For example, our theoretical

* Corresponding author. Fax: 1-301-314-0386.
E-mail address: fushman@wam.umd.edu (D. Fushman).

1090-7807/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2004.03.019

analysis of hydrodynamic properties for a representative set of 878 protein structures suggests that about 90% of monomeric prolate proteins have rotational anisotropy, $\xi = 2D_{zz}/(D_{xx} + D_{yy})$, greater than 1.17 and 60% have the rhombicity, $\eta = 1.5(D_{yy} - D_{xx})/[D_{zz} - 0.5(D_{xx} + D_{yy})]$, greater than 0.2 (Geraghty et al., in preparation). Efficient computational methods are, therefore, needed for the complete tensor determination.

Several computational approaches have been developed, which employed a combination of simplex and grid searches [5,9,19], simulated annealing [20], or Bayesian statistics [23]. In the simplex-based methods [5,19], the optimization was separated into a simplex search in a 3D space of the principal values of the tensor and a grid search in the Euler angle space. This was done to circumvent the problem of implementing periodic boundary conditions for the Euler angles in the simplex algorithm. The angular-grid search turned out to be time consuming, particularly in the case of a fully anisotropic tensor.

Here, we present a software program ROTDIF that uses a different computational strategy to determine the full set of diffusion tensor parameters and Euler angles. A novel feature of the proposed method is that the optimization search is performed in a full six-dimensional space without separating it into the individual subspaces for the principal values and the angles. This is achieved by using a constrained Levenberg–Marquardt algorithm, which leads to a significant, up to 50-fold gain in the computational efficiency of the analysis without any loss in the accuracy. We demonstrate the accuracy and efficiency of the method by applying it to both synthetic data sets and experimental data of a real protein. We also analyze the expected accuracy and precision of the diffusion tensor derivation depending on the experimental noise in relaxation data.

2. Results and discussion

2.1. Method description

The determination of the rotational diffusion tensor of a molecule uses the dependence of the measured spin-relaxation rates (e.g., for ^{15}N considered here) on the orientation of a given group (NH bond) with respect to the principal axes frame of the diffusion tensor (see, e.g. [3]). Given the structure of a molecule and experimental relaxation data (^{15}N R_1 and R_2 and heteronuclear ^{15}N$\{^1$H$\}$ NOE), the determination of the diffusion tensor parameters is achieved here by minimizing the target function

$$\chi^2 = \sum_{i=1}^{N_r} \left(\frac{\rho_i^{\exp} - \rho_i^{\mathrm{calc}}}{\sigma_i} \right)^2, \qquad (1)$$

where N_r is the total number of NH vectors included in the analysis, parameter ρ is defined as a ratio of the transverse and longitudinal relaxation rates,

$$\rho = \left(\frac{2R_2'}{R_1'} - 1 \right)^{-1} = \frac{3}{4} \frac{J(\omega_N)}{J(0)} \qquad (2)$$

and σ_i denotes the experimental error in ρ_i for vector number i. The value of $\rho^{\exp}$ is directly derived from the measured relaxation parameters, while ρ^{calc} is calculated using expressions for the spectral densities $J(\omega)$, depending on the rotational diffusion model (see, e.g. [19,24]).

The primes in R_2' and R_1' in Eq. (2) indicate that these relaxation rates were modified to subtract the contributions from high-frequency components of the spectral density as described in [9,24]. The advantage of using the ratio, R_2'/R_1', instead of the individual values of these parameters is that this ratio is independent, to a first approximation, of the site-specific variations in the strength of ^{1}H–^{15}N dipolar coupling and ^{15}N chemical shift anisotropy. Moreover, in the case of protein core residues the R_2'/R_1' ratio primarily depends on the overall tumbling and is practically insensitive to fast, subnanosecond backbone dynamics, because the order parameters of local motions in the numerator and the denominator of Eq. (2) cancel out [9].

The novelty of the proposed algorithm over the previous methods [5,9,19,25] is that the optimization search that minimizes the target function, Eq. (1), is performed in a full six-dimensional space. Instead of setting the periodic boundary conditions for the Euler angles we impose strict boundary conditions and restrict the allowed angle space to a 3D box (cube) of the $([0°, 180°]$ intervals for all three Euler angles. Due to the symmetry properties of the problem [13], this subspace fully represents all relevant orientations of the tensor frame. Our analysis showed that it is possible to implement this approach in simplex by imposing a severe penalty function for the angle values outside this 3D box. However, the constrained Levenberg–Marquardt method turned out to be robust and more efficient in this case, even when combined with the Monte Carlo screening of the initial guesses (see below).

The Levenberg–Marquardt algorithm works very well in practice and has become the standard of nonlinear least-squares optimization. It yields accurate results when the initial guess is not far from the global minimum. To avoid being trapped in a local minimum, we generate a random set of initial guesses and, for each of them, minimize the target function χ^2. This is performed in ROTDIF automatically, and the software then selects only those values of the fitting parameters that give the lowest value of the target function. Our extensive testing indicates that a set of $\sim$100 random initial guesses is usually sufficient to find at least one (usually more than a dozen) solution corresponding to

the global minimum. The algorithm is so fast that a full 6D search with 100 initial guesses for 50 NH vectors takes approximately 20 s elapsed time (i.e., about 0.2 s per guess) on a 1.5 GHz Pentium 4 XEON processor. It takes about 40 s if error calculation is included.

The uncertainties in the derived diffusion tensor characteristics are calculated using the method of constant-χ^2 boundaries [26] that incorporates covariance matrix analysis and Monte Carlo simulation of the fitting parameters, with the acceptance condition of at least 500 points within the 68.3% confidence region for the χ^2-distribution.

To better assess the agreement between the calculated and the experimental values of ρ, we also included a quality factor similar to that previously defined in [19,27]:

$$R = \sqrt{\frac{\left\langle (\rho^{\exp} - \rho^{\mathrm{calc}})^2 \right\rangle}{2\left\langle (\rho^{\exp} - \langle \rho^{\exp} \rangle)^2 \right\rangle}}, \tag{3}$$

where the brackets represent averaging over all available NH vectors. $R \to 0$ when the calculated and experimental values of ρ are close to each other.

2.2. Basic features of the ROTDIF program

(1) The ROTDIF program is written in Matlab. It can run in a user–computer dialog mode or in the automatic mode. The program automatically performs data analysis for all three models for the overall rotational diffusion: isotropic, axially symmetric, and fully anisotropic, and determines the most appropriate model based on the statistical F test [28]. Because the fit between a model and experimental data generally improves with the number of adjustable parameters in the model function, we use the F test to evaluate whether the decrease in the target function achieved with a larger number of fitting parameters is statistically significant. A large value of F justifies the inclusion of the additional terms in the fit. The results are also represented in terms of the probability P that the observed improvement in the $(N+q)$-parameter fit over the N-parameter fit is obtained by chance. For the F test statistics to be significant at the $(1-\alpha)\%$ confidence level, P should be less than α.

(2) The software extracts NH-vector coordinates from a user-specified PDB file. If hydrogen atoms are not present in the PDB file, these will be automatically built based on the peptide-plane geometry (e.g. [29]). Alternatively, the user can input NH-vector coordinates as a separate input parameter.

(3) All results are being automatically saved to a text file and displayed on screen in the form of text and plots. Figs. 1–4 in this paper represent outputs of the program, adapted for publication purposes.

(4) The software also allows visualization of the principal axes of the derived diffusion tensor. This is done by generating pseudo-atoms positioned in the origin and along the x-, y-, and z-axes of the tensor. Coordinates of these atoms can be saved in the PDB format (e.g., appended to a user-specified PDB file), and can then be viewed using standard protein visualization software, see examples in our papers [4,7,8].

(5) In the current version of ROTDIF, the method is implemented for ^{15}N relaxation data. It can be adapted for analysis of other nuclei, e.g., ^{13}C.

3. Applications

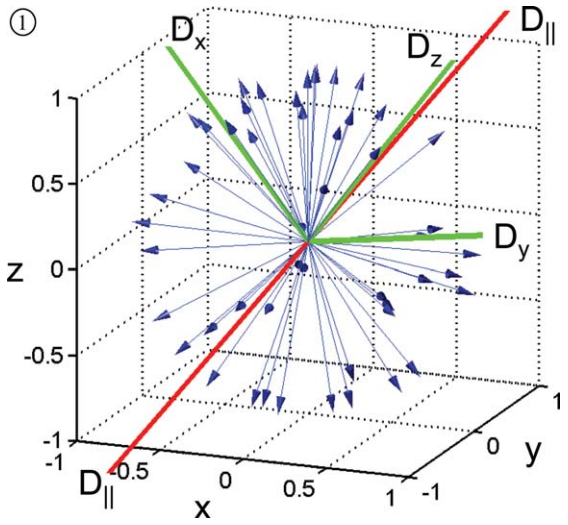
3.1. Computer simulations

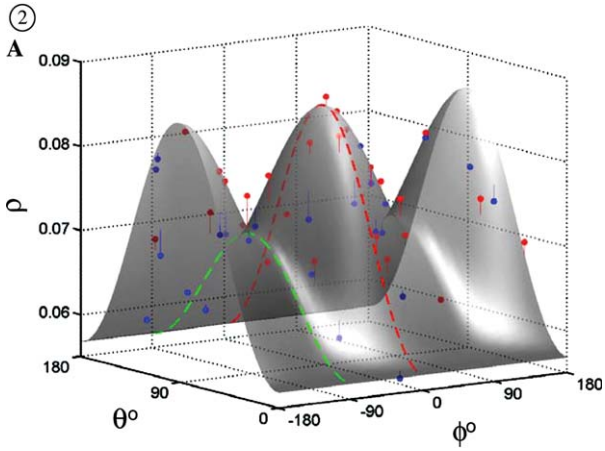
To demonstrate the utility of our program, we first apply it to synthetic ^{15}N relaxation data. This allows us to compare the results of the analysis with the true, input rotational diffusion model in order to verify the accuracy of the computational approach proposed

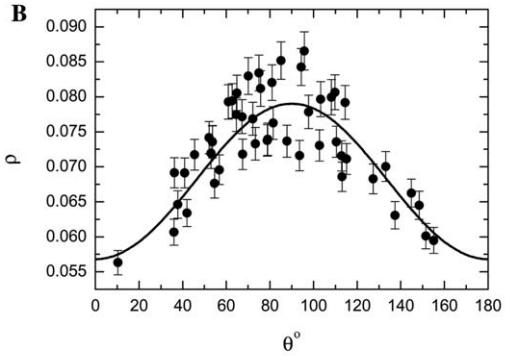
Fig. 1. Orientations of the 50 generated NH vectors (blue) and the principal axes of the derived diffusion tensor for the axially symmetric (red, labeled as $D_{\|}$) and fully anisotropic (green, labeled as D_x, D_y, D_z) diffusion models. The orientations of the NH vectors are distributed almost uniformly, and inferred from the low value (0.0415) of the generalized sampling parameter [29].

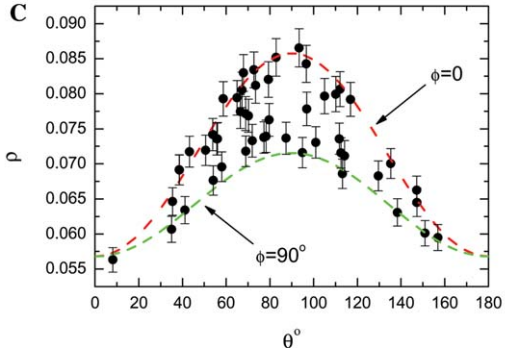
Fig. 2. Orientation dependence of ρ as a result of fitting the computer-generated "experimental" data to anisotropic rotational diffusion models. (A) The data were fit to a fully anisotropic diffusion model. The fitting surface is used, because in this model ρ depends on both θ and ϕ angles that define the orientation of a NH vector with respect to the diffusion tensor frame. The vertical lines represent deviations of the data points from their expected positions on the surface; data points above the surface are colored red, those underneath the surface are blue. (B) Data fit to an axially symmetric diffusion model. The large spread of the data around the fitting curve suggests that this model is not adequate. (C) A projection of the surface, shown in (A), onto the $\theta - \rho$ plane ($\phi = 0$) illustrates the fact that the spread of the data points seen in (B) is within the boundaries corresponding to $\rho(\theta)$ for $\phi = 0°$ and $\phi = 90°$, shown as dashed curves in (A). Relaxation data were generated as described in the text; the tensor parameters are listed in Table 1 for the anisotropy of 1.4 and rhombicity of 0.8.

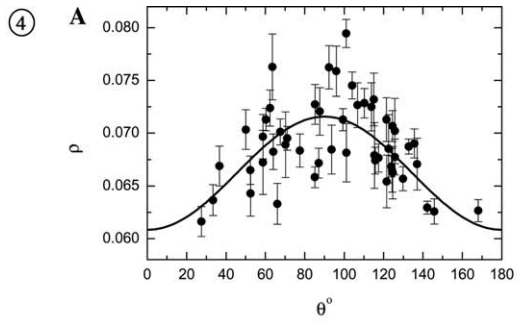
Fig. 4. Application of the method to the proximal Ub domain in K48-linked Ub_2. Orientational dependence of the parameter ρ obtained for (A) axially symmetric and (B,C) fully anisotropic models for the diffusion tensor, and the agreement between the experimental and calculated values of ρ for (D,F) axially symmetric and (E,G) fully anisotropic models. (C) A projection of the hilly surface shown in (B) onto the $\rho - \theta$ plane; the dashed lines in (B) and (C) represent the upper and lower boundaries corresponding to $\rho(\theta)$ for $\phi = 0°$ and $\phi = 90°$.

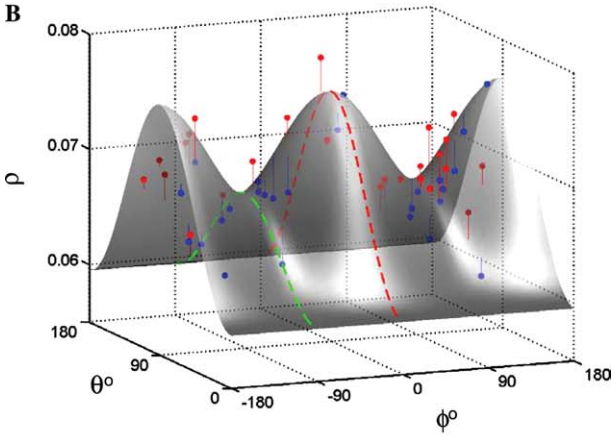

①
D_x
D_z
$D_{||}$
D_y
$D_{||}$
z
x
y

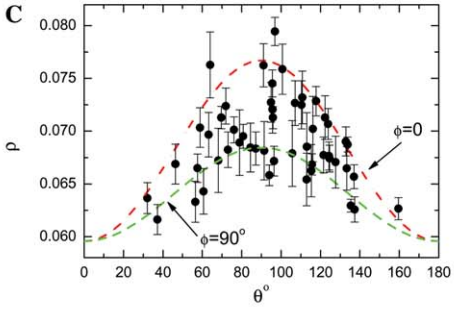

②
A
ρ
θ°
ϕ°

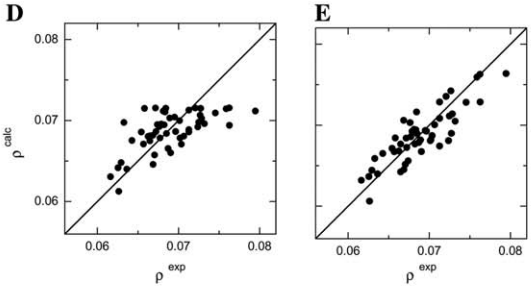

B
ρ
θ°

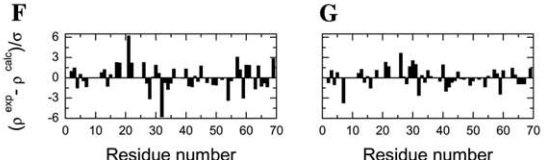

C
ρ
ϕ=0
ϕ=90°
θ°


④
A
ρ
θ°


B
ρ
θ°
ϕ°


C
ρ
ϕ=0
ϕ=90°
θ°


D
E
ρ^{calc}
ρ^{exp}
ρ^{exp}


F
G
$(\rho^{exp} - \rho^{calc})/\sigma$
Residue number
Residue number

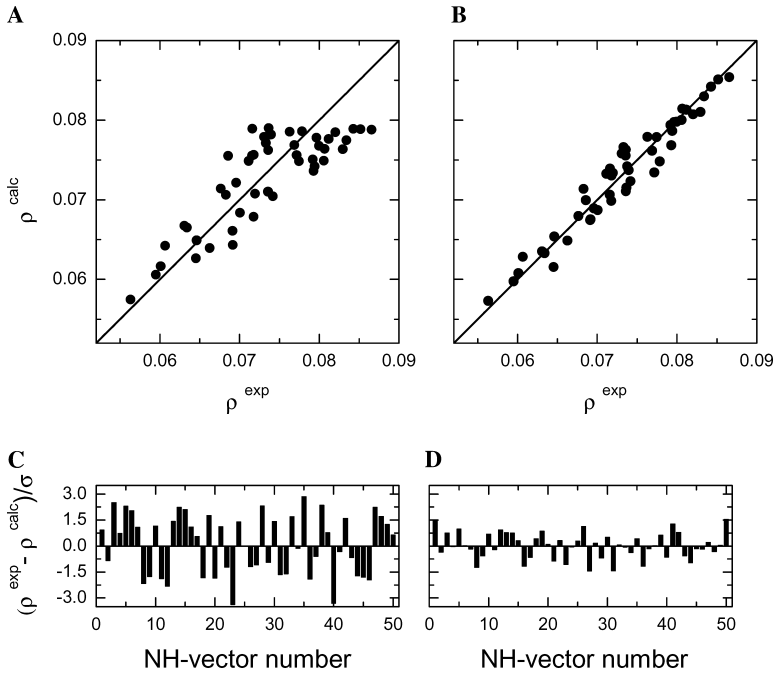


Fig. 3. The agreement between computer-generated "experimental" data for a fully anisotropic diffusion tensor and their back-calculated values using (A,C) axially symmetric and (B,D) fully anisotropic models of the overall tumbling. (C,D) The difference, $\rho^{exp} - \rho^{calc}$, divided by the experimental errors in ρ, for each NH vector. A "flat top" observed in (A) is a signature of the case when the data for a fully anisotropic diffusion tensor are analyzed using an oversimplified, axially symmetric model that is not capable of reproducing the dependence of the experimental data on the angle ϕ.

here. A set of 50 randomly uniformly distributed NH vectors was generated for this analysis (Fig. 1). Relaxation data (^{15}N R_1 and R_2, and $^{15}N\{^1H\}$ NOE) were generated for these vectors assuming an isotropic, axially symmetric, or fully anisotropic model for the overall diffusion tensor. The input characteristics of the diffusion tensor for each model are shown in Table 1, along with the results of the analyses. We assumed that the local backbone dynamics are characterized by the model-free parameters $S^2 = 0.9$ and $\tau_{loc} = 50$ ps, typical for the protein core. A 2% random "experimental" noise was added to all simulated relaxation data. The results (Table 1) show excellent agreement between the derived and the input diffusion tensors, when the proper overall rotational diffusion model is applied.

In each case, we have compared the performance of our method with that of the previous program DIFTENS [19] (with a 1° step grid search) in terms of accuracy and computation time. For both anisotropic models, the diffusion tensors obtained by the two different methods were virtually identical, while the Levenberg–Marquardt algorithm (ROTDIF) proved to be much faster (up to ~50-fold) than the optimized grid-search approach (DIFTENS): 21 s versus 17 min elapsed time, for the fully anisotropic model. Of the 100 initial guesses used for ROTDIF, 80 resulted in the global minimum characterized by the least value of the target function and 20 led to three other values of χ^2 representing local minima. Note that, in principle, this optimization strategy also provides an improvement in the accuracy of the tensor determination because none of the fitting parameters is limited to a grid.

3.2. Characteristic signatures of an inadequate diffusion model

Because the proper diffusion tensor model is not known a priori, it is important to be able to verify that the applied diffusion model is adequate. While the presence of rotational anisotropy is usually readily recognizable, and significant attention has already been paid in the literature to how to distinguish it from the isotropic tumbling, the presence of the rhombic component of the diffusion tensor could be more difficult to identify. Here, we use our computer-generated data to reveal the characteristic signatures of a fully anisotropic rotational diffusion.

We will discuss here in detail the application of an axially symmetric diffusion model to relaxation data generated for a fully anisotropic diffusion tensor with rhombicity of 0.8 (see Table 1). In this case the target function had two global minima, which were very close in their χ^2 values (147.7 and 158.7) and represented prolate and oblate approximations of the tensor, respectively. The presence of two minima has been shown [15] to be a general feature when approximating a truly anisotropic system by an axially symmetric tensor. The prolate solution was chosen based on a lower value of the target function. Naturally, the two-minima problem

Table 1
The results of the diffusion tensor determination from computer-generated spin-relaxation data for various input and applied diffusion tensor models

Model	D_{xx}[a]	D_{yy}[a]	D_{zz}[a]	α[b]	β[b]	γ[b]	τ_c[c]	ξ[d]	η[e]	χ^2	F[f]	P[g]	R[h]
Input: *iso*	2.08	2.08	2.08	—	—	—	8.0	0	0				
Full	2.05	2.07	2.12	158	50	78	8.00	1.03	0.371	39.9	0.07	0.93	0.68
	(0.07)	(0.07)	(0.08)	(53)	(41)	(95)	(0.14)	(0.04)	(0.3)				
Axial	2.06	2.06	2.12	155	48	—	8.00	1.03		40.1	1.24	0.3	0.68
	(0.03)	(0.03)	(0.07)	(32)	(53)		(0.12)	(0.03)					
Iso	2.08	2.08	2.08	—	—	—	8.01			43.3			0.7
	(0.01)	(0.01)	(0.01)				(0.02)						
Input: *axial*	1.84	1.84	2.57	70	60	—	8.0	1.4	0				
Full	1.83	1.87	2.55	71	58	33	8.0	1.38	0.094	40.1	0.8	0.45	0.29
	(0.10)	(0.09)	(0.2)	(5)	(3)	(52)	(0.3)	(0.12)	(0.02)				
Axial	1.85	1.85	2.55	72	59	—	8.01	1.38		41.6	125	3.7×10^{-22}	0.29
	(0.02)	(0.02)	(0.07)	(2)	(3)		(0.1)	(0.04)					
Iso	2.08	2.08	2.08	—	—	—	8.03			381			0.72
	(0.01)	(0.01)	(0.01)				(0.05)						
Input: *full*	1.79	1.89	2.57	70	60	170	8.0	1.4	0.2				
Full	1.81	1.88	2.57	71	60	150	8.0	1.39	0.15	35.6	2.3	0.12	0.22
	(0.07)	(0.07)	(0.14)	(4)	(3)	(32)	(0.2)	(0.08)	(0.02)				
Axial	1.84	1.84	2.57	71	61	—	8.0	1.39		39.3	137	6×10^{-23}	0.24
	(0.03)	(0.03)	(0.07)	(3)	(3)		(0.1)	(0.04)					
Iso	2.08	2.08	2.08	—	—	—	8.02			390			0.72
	(0.01)	(0.01)	(0.01)				(0.05)						
Input: *full*	1.64	2.03	2.57	70	60	170	8.0	1.4	0.8				
Full	1.62	2.03	2.56	72	59	166	8.04	1.40	0.83	28	94	1.3×10^{-16}	0.16
	(0.04)	(0.04)	(0.05)	(4)	(2)	(6)	(0.1)	(0.04)	(0.04)				
Axial	1.82	1.82	2.58	72	61	—	8.03	1.42		147.7	40	5.7×10^{-13}	0.40
	(0.05)	(0.05)	(0.14)	(5)	(7)		(0.25)	(0.07)					
Iso	2.07	2.07	2.07	—	—	—	8.05			538.1			0.72
	(0.02)	(0.02)	(0.02)				(0.06)						
Input: *full*	1.59	2.08	2.57	70	60	170	8.0	1.4	1.0				
Full	1.56	2.07	2.61	70	59	170	8.01	1.44	0.97	52.4	80	2.1×10^{-15}	0.19
	(0.04)	(0.05)	(0.06)	(4)	(3)	(6)	(0.1)	(0.04)	(0.06)				
Axial	1.81	1.81	2.63	69	62	—	8.0	1.46		243	29	9.2×10^{-11}	0.44
	(0.06)	(0.06)	(0.17)	(7)	(9)		(0.33)	(0.09)					
Iso	2.08	2.08	2.08	—	—	—	8.03			709			0.73
	(0.02)	(0.02)	(0.02)				(0.07)						

Numbers in the parentheses represent standard errors in the derived parameters. Here, "iso," "axial," and "full" stand for the isotropic, axially symmetric, and fully anisotropic models, respectively. The input lines list parameters for the input model, as indicated.

[a] The principal values of the diffusion tensor, in $10^7\ s^{-1}$.

[b] The Euler angles, in degrees, define the orientation of the principal axes frame of the diffusion tensor relative to the coordinate frame of the generated set of NH vectors.

[c] The overall rotational correlation time, $\tau_c = 1/(6D_{iso})$, in nanoseconds.

[d] The anisotropy of the tensor, see Section 1.

[e] The rhombicity of the tensor, see Section 1.

[f] The F value for the F test comparing the current model with a simpler one (in the row below): fully anisotropic versus axially symmetric or axially symmetric versus isotropic model. Large values of F justify the use of a more complex model.

[g] The probability that the improvement in the fit when applying a more complex model has occurred by chance.

[h] The quality factor, Eq. (3).

disappeared when the data were treated using a fully anisotropic model. Moreover, the target function for the fully anisotropic model was considerably lower ($\chi^2 = 28$), which is generally interpreted as a better fit. This is also confirmed by the F test ($F = 94$, Table 1) indicating that there is a 10^{-16} probability that the observed decrease in χ^2 was obtained by chance.

Prior to comparison of the results obtained for different diffusion models, it is instructional to consider the orientational dependence of the parameter ρ. The ^{15}N relaxation rates (hence the parameter ρ) depend on the orientation of the NH bond vector with respect to the principal axes frame of the tensor. This orientation is defined here by the polar angle θ and the azimuthal angle ϕ. The theoretical dependence of ρ as a function of θ and ϕ can be represented by a hilly surface shown in Fig. 2A for a fully anisotropic diffusion tensor (prolate). The overall shape of the surface reflects symmetry properties of ρ: $\rho(\theta,\phi) = \rho(180^\circ - \theta,\phi) = \rho(\theta,-\phi) = \rho(\theta,180^\circ - \phi) = \rho(\theta,\phi - 180^\circ)$, stemming from the fact

that the relaxation parameters are not sensitive to the directionality (sign) of the NH-vector coordinates. The actual heights of the "hills" and depths of the "valleys" (saddle points) depend on the principal values $\{D_{xx}, D_{yy}, D_{zz}\}$ of the diffusion tensor. The difference between the hills and the valleys depends on the rhombicity of the tensor and disappears when $D_{yy} = D_{xx}$, i.e., for an axially symmetric tensor, whereby ρ does not depend on the angle ϕ. The whole surface flattens and becomes a horizontal plane in the isotropic case ($D_{xx} = D_{yy} = D_{zz}$).

The quality of the data fit to anisotropic diffusion models is illustrated in Fig. 2 depicting the angular dependence of the relaxation data. The fully anisotropic model is in excellent agreement with the "experimental" data (Fig. 2A) that fit nicely into the hills and valleys of the $\rho(\theta, \phi)$-surface.

The axially symmetric diffusion model yields reasonably accurate values of the anisotropy of the tensor and of the overall correlation time and predicts the correct orientation of the z-axis of the diffusion tensor (Table 1, Fig. 1). However, it shows a significant vertical spread of the ρ values around the fitting curve (Fig. 2B), most pronounced around its maximum (at $\theta = 90°$). This is because applying an axially symmetric model to the fully anisotropic case amounts to fitting data to a surface with no hills or valleys in the ϕ-dimension. As a result, the data points located in various parts (hills and valleys) of the $\rho(\theta, \phi)$ surface (Fig. 2A) now appear spread in the vertical direction around the $\rho(\theta)$ fitting curve. The upper and lower boundaries of this area are given by $\phi = 0°$ and $\phi = 90°$ and correspond to those cases when the NH vector is orthogonal to the y- and x-axis, respectively, of the diffusion tensor (cf. Fig. 2A). A projection of the $\rho(\theta, \phi)$-surface onto the $\theta - \rho$ plane ($\phi = 0$), shown in Fig. 2C, illustrates this issue by showing that the data points spread in Fig. 2B indeed fall within these boundaries, indicated by the dashed curves in Fig. 2A. This observation suggests that a vertical spread of data points in the area around $\theta = 90°$ in the axially symmetric fit is a characteristic signature of the difference between D_{xx} and D_{yy} and can be used as an indicator that an axially symmetric model is inadequate.

Another characteristic signature of the case when an axially symmetric model is inadequate is illustrated in Fig. 3A depicting the correlation between the experimental and back-calculated data. The correlation plot has a characteristic "flat top": the upper limit of the calculated ρ values is always truncated compared to the maximal experimental data. This is because in the axially symmetric model, the fitting curve lies in the middle between the boundaries defined by $\phi = 0°$ and $\phi = 90°$. This curve sets the upper limit for the calculated values of $\rho(\theta)$ at $\theta = 90°$: these values cannot reach the hilltop as they do in the fully anisotropic model. The ρ^{calc} values predicted using fully anisotropic model are in excellent agreement with the "experimental" data (the correlation coefficient $r = 0.98$, the quality factor $R = 0.16$, Figs. 3B and D) whereas the agreement is noticeably worse ($r = 0.84$, $R = 0.4$) for the axially symmetric model (Figs. 3A and C).

In the opposite situation, when a more complex, fully anisotropic model was applied to relaxation data generated for the axially symmetric model (Table 1), the results show a large experimental uncertainty in the Euler angle γ that defines the orientation of the x- and y-axes of the tensor. The F test indicated that the fully anisotropic model was not statistically better than the axially symmetric model: there was a 45% probability that the observed marginal decrease in the χ^2 was obtained by chance. In addition, there was no improvement in the quality factor ($R = 0.29$) compared to the axially symmetric model. Similar conclusions hold when the data generated for isotropic model are treated assuming rotational anisotropy (Table 1).

3.3. Application to a real system: ubiquitin domain in a di-ubiqutin chain

To demonstrate the ROTDIF program on a real protein and to illustrate the ideas outlined above, here we apply it to the proximal ubiquitin (Ub) domain in the context of the K48-linked di-ubiquitin (Ub_2) chain [7]. Other applications of the program can be found in [4,8]. The NMR data, ^{15}N R_1, R_2, and $^{15}N\{^1H\}$ NOE, were measured at 14.1 T as described in [7]; the average experimental precision in the relaxation data was 2%. Our chemical shift perturbation data and residual dipolar coupling measurements indicate that the backbone structure of each Ub domain is not significantly perturbed in Ub_2, therefore protein coordinates for monomeric Ub [30] will be used in the analysis. Fifty-four out of the 73 observed NH groups were used for the rotational diffusion tensor determination. Residues that exhibited noticeable chemical shift perturbations, significant conformational exchange, or rapid motion on the fast timescale were excluded from the analysis. The results of the analysis are presented in Table 2 and Fig. 4.

First we applied an axially symmetric diffusion tensor model. The target function showed two global minima corresponding to a prolate ($\chi^2 = 190$) and an oblate ($\chi^2 = 194$) approximations of the tensor. The dependence of the parameter ρ on the NH vector orientation is shown in Fig. 6A for the prolate solution, the agreement between the experiment and the model is presented in Fig. 4D. Overall, the axially symmetric model could not reproduce well the experimental data. Consistent with the discussion above, the existence of the two minima, the observed vertical spread of the data (Fig. 4A), and the characteristic "flat top" in Fig. 4D, all indicate that the tensor has a significant rhombic component. Indeed, the agreement between the experiment and the theory improved significantly (Figs. 4B and E,

Table 2
Rotational diffusion tensor of Ub_2 derived from ^{15}N relaxation data measured for the proximal Ub domain

Model	D_{xx}	D_{yy}	D_{zz}	α	β	γ	τ_c	ξ	η	χ^2	F	P	R
Full	1.75	2.0	2.31	105	112	135	8.26	1.23	0.865	97	21	2.8×10^{-7}	0.36
	(0.04)	(0.05)	(0.06)	(6)	(6)	(10)	(0.1)	(0.04)	(0.1)				
Axial	1.89	1.89	2.25	96	107	—	8.28	1.19		190	15	4.2×10^{-7}	0.55
	(0.04)	(0.04)	(0.12)	(11)	(10)		(0.2)	(0.06)					
Iso	2.01	2.01	2.01	—	—	—	8.29			377			0.73
	(0.01)	(0.01)	(0.01)				(0.04)						

Numbers in the parentheses represent standard errors in the derived parameters. The meaning of all parameters and the units are the same as in Table 1. The atom coordinates were taken from the NMR structure of mono-Ub (PDB entry 1D3Z.pdb) [30].

Table 2) when a fully anisotropic model was used. Here, $\chi^2 = 97$, and the probability that this improvement in the fit is due to chance is $P = 2.8\times10^{-7}$. The orientational dependence of ρ that looks like a significant spread of data in Fig. 4A, now fits nicely the hilly surface (Fig. 4B) that takes into account an additional degree of freedom (angle ϕ) describing NH vector orientation (Figs. 4B and C).

3.4. Precision and accuracy of the diffusion tensor determination

Knowing the limits of the accuracy and precision in the NMR-derived diffusion tensors is critical for their use for protein dynamics analysis and for domain orientation in multidomain systems. Equipped with an efficient method of full diffusion tensor analysis, we are now in a position to explore these issues. Here, we focus only on the inaccuracy and imprecision in the tensor determination caused by the experimental noise; the effect of limited orientational sampling due to a non-uniform distribution of the NH vectors has been discussed elsewhere [29].

To assess the dependence of the derived rotational diffusion tensor on the quality of the experimental data, we used the same 50 randomly uniformly oriented NH vectors as above, and generated multiple synthetic sets of ^{15}N relaxation data assuming an axially symmetric or a fully anisotropic diffusion tensor. The τ_c and the orientation of the tensor were the same as in Table 1; the levels of experimental error ranged from 1 up to 10%. The anisotropy varied from 1.04 to 1.4 for an axially symmetric diffusion model. For a fully anisotropic model, where both the anisotropy and rhombicity could vary, we considered four different possibilities. First, the anisotropy was kept constant ($\xi = 1.3$) and the rhombicity varied from 0.1 to 1.0. Then we varied the anisotropy (from 1.05 to 1.4) for three fixed values of the rhombicity ($\eta = 1.0$, 0.4, and 0.2). Five hundred sets of relaxation data (R_1, R_2, and NOE) were generated for each combination of the experimental error, anisotropy, and rhombicity. Each generated data set was analyzed assuming a proper diffusion tensor model. The precision of the derived tensor was computed as the standard deviation over the 500 resulting tensors, while the accuracy was assessed as the deviation of the mean of these tensors from the true (input) value.

3.4.1. Precision

The derived uncertainties in the tensor parameters are depicted in Fig. 5 for the axially symmetric model. Here, the orientation of the tensor (angles α and β) turned out to be more sensitive than its magnitude (described by $D_\parallel$ and $D_\perp$) to the experimental errors and/or the degree of anisotropy. For a given experimental error, the uncertainties in the magnitude of the diffusion tensor increase almost linearly with the decrease in the anisotropy, while the errors in the angles increase faster, as a polynomial function. Similar behavior was observed for the fully anisotropic tensor model. For example, based on our analysis for a typical 2% level of experimental errors in R_1, R_2, and NOE, the expected uncertainties in the Euler

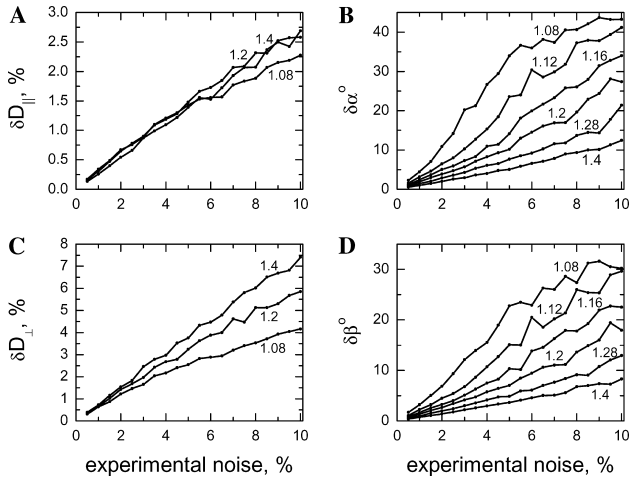


Fig. 5. Illustration of the expected imprecision of the derived diffusion tensor parameters as a function of experimental noise in ^{15}N relaxation parameters. The plots represent relative errors (in %) in (A) $D_\parallel$ and (C) $D_\perp$ and the absolute errors (in degrees) in (B) α and (D) β for various levels of rotational anisotropy, as indicated. The "experimental" data were generated and analyzed using an axially symmetric model. The input values for the orientation of the tensor ($\alpha = 70°$, $\beta = 60°$) and the τ_c ($= 8$ ns) were kept constant. The level of experimental noise in relaxation data was from 1 to 10%. These data indicate that the orientation of the tensor (angles α and β) is more sensitive than its magnitude ($D_\parallel$, $D_\perp$) to experimental errors and/or to the degree of anisotropy of the tensor. To reduce the crowdiness of the plots, only data for $\xi = 1.08$, 1.2, and 1.4 are shown in (A) and (C).

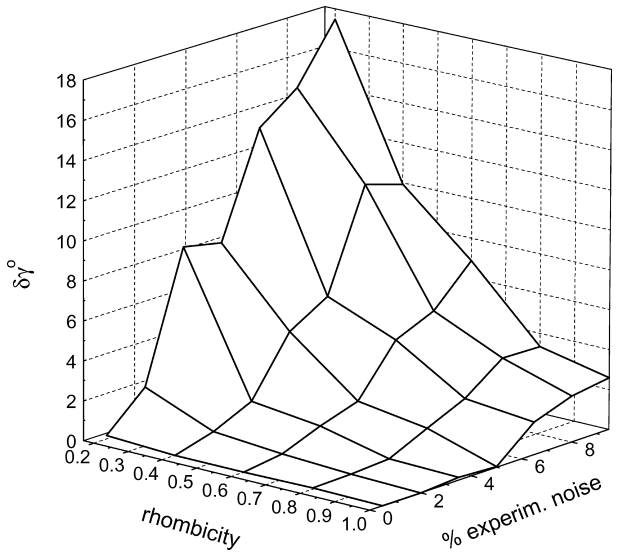


Fig. 6. The expected inaccuracy in the γ angle as a function of experimental noise and of the rhombicity of the diffusion tensor, for a constant anisotropy of 1.3. The inaccuracy was computed as $|\gamma_{exp} - \langle\gamma_{calc}\rangle|$, where $\langle\cdots\rangle$ denotes averaging over the set of 500 tensor values derived from fitting 500 sets of synthetic relaxation data for each combination of the rhombicity and the level of experimental noise.

angles α and β are 4° and 3°, respectively, assuming an axially symmetric model with small anisotropy of 1.2. The errors in $D_{\parallel}$ and $D_{\perp}$ were at 0.6 and 1.2%, respectively. These uncertainties increase for less precise experimental measurements. These data emphasize that the orientation of the diffusion tensor is significantly more sensitive to measurement errors than its principal values. The same tendency was observed in the case of a fully anisotropic tensor (data not shown). Here, the angle γ turned out to be most sensitive to experimental errors. Assuming the same levels of rotational anisotropy and experimental errors as above, the uncertainties in γ varied from 17.5° to 5.3° for the range of rhombicities from 0.2 up to the maximum value of 1.0, while the errors in α and β (4–6° and 2.5–2.2°, respectively) and in D_{xx}, D_{yy}, and D_{zz} (1.3–1.2, 1.2–1.15, and 1.2–1.14%, respectively) remained practically unaffected.

3.4.2. Accuracy

The experimental noise reduces the sensitivity of the analysis, thus making the tensor determination less accurate, even in the case when a correct diffusion model is applied. Our analysis indicates that, overall, the expected accuracy of the diffusion tensor determination is very good. The observed deviations from the input values were below 0.4% for the principal values of the tensor and below 1° and 0.5° for the angles α and β, respectively, for the range of tensor anisotropies from 1.1 to 1.4 and the rhombicities from 0 to 1, assuming a 2% level of experimental noise in the relaxation data. The inaccuracies in these parameters increased with the experimental errors and, for example, reached 1.4% (D_{xx}), 0.8% (D_{yy}, D_{zz}), and 2.8° (α, β) at a 5% experimental noise level and the lowest tensor anisotropy of 1.1. Here, again, for a fully anisotropic tensor the angle γ is the least accurately determined parameter (Fig. 4E). This angle deviates significantly from its input value when the rhombicity and/or anisotropy are small, hence the orientation of the x- and y-axes becomes poorly defined.

3.4.3. Sensitivity

Based on the error analysis we have also addressed the sensitivity of the diffusion tensor determination. In particular, we estimated the lower limits for the anisotropy and rhombicity of the tensor such that the tensor's orientation can be determined with a better than 10° uncertainty. Based on these criteria, the lowest identifiable anisotropy for an axially symmetric tensor was $\xi = 1.10$, assuming τ_c of 8 ns and a 2% level of experimental errors. For a fully anisotropic model, in order to have a less than 10° uncertainty in the γ angle, the anisotropy had to be greater than 1.12 (for $\eta = 1.0$) or greater than 1.2 ($\eta = 0.4$), for the same level of experimental errors and τ_c. The lowest identifiable rhombicity was 0.3, assuming similar precision criteria (including angle γ) and the tensor anisotropy of 1.3.

4. Conclusions

In the present paper, we have demonstrated a novel computational method for efficient and accurate determination of the rotational diffusion tensor of a molecule from spin-relaxation data. This is achieved via a six-dimensional optimization search using a constrained Levenberg–Marquardt algorithm combined with Monte Carlo generation of initial guesses. This algorithm is robust and provides significant time saving compared to our previous approaches. This method is implemented in our computer program ROTDIF. The method is applied here to real experimental data on ubiquitin domain in di-ubiquitin and to computer-generated data for isotropic, axially symmetric, and fully anisotropic diffusion tensors. The analysis of computer-generated relaxation data made it possible to identify characteristic signatures of data treatment using inadequate rotational diffusion models and to explore the effect of experimental errors on the accuracy and precision of the derived diffusion tensor.

Note added in proof

Further details of the analysis presented here and its application to interdomain orientation in di-ubiquitin can be found in D. Fushman, R. Varadan, M. Assfalg, O. Walker, Determining domain orientation in macromolecules by using spin-relaxation and residual dipolar

coupling measurements, Progress NMR Spectroscopy (2004) in press (available online as Articles in Press).

Acknowledgments

Supported by NIH Grant GM65334-01 to D.F. The ROTDIF software is freely available for download from the web site (www.vsnmr.org/FushmanLab) and from the authors upon request.

References

[1] J.M. Schurr, H.P. Babcock, B.S. Fujimoto, A test of the model-free formulas: effects of anisotropic rotational diffusion and dimerization, J. Magn. Reson. B 105 (1994) 211–224.
[2] P. Luginbuhl, K.V. Pervushin, H. Iwai, K. Wuthrich, Anisotropic molecular rotational diffusion in ^{15}N spin relaxation studies of protein mobility, Biochemistry 36 (1997) 7305–7312.
[3] D. Fushman, D. Cowburn, Studying protein dynamics with NMR relaxation, in: R. Sarma, M. Sarma (Eds.), Structure, Motion, Interaction and Expression of Biological Macromolecules, Adenine Press, Albany, NY, 1998, pp. 63–77.
[4] J.B. Hall, D. Fushman, Characterization of the overall and local dynamics of a protein with intermediate rotational anisotropy: differentiating between conformational exchange and anisotropic diffusion in the B3 domain of protein G, J. Biomol. NMR 27 (2003) 261–275.
[5] D. Fushman, R. Xu, 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 (1999) 10225–10230.
[6] P.M. Hwang, N.R. Skrynnikov, L.E. Kay, Domain orientation in b-cyclodextrin-loaded maltose binding protein: diffusion anisotropy measurements confirm the results of a dipolar coupling study, J. Biomol. NMR 20 (2001) 83–88.
[7] R. Varadan, O. Walker, C. Pickart, D. Fushman, Structural properties of polyubiquitin chains in solution, J. Mol. Biol. 324 (2002) 637–647.
[8] R. Varadan, M. Assfalg, A. Haririnia, S. Raasi, C. Pickart, D. Fushman, Solution conformation of Lys63-linked di-ubiqutin chain provides clues to functional diversity of polyubiquitin signaling, J. Biol. Chem. 279 (2004) 7055–7063.
[9] D. Fushman, D. Cowburn, Characterization of inter-domain orientations in solution using the NMR relaxation approach, in: N.R. Krishna (Ed.), Protein NMR for the Millenium, Biological Magnetic Resonance, vol. 20, Kluwer, Dordrecht, 2002, pp. 53–78.
[10] P. Bernado, T. Akerud, J. Garcia de la Torre, M. Akke, M. Pons, Combined use of NMR relaxation measurements and hydrodynamic calculations to study protein association. Evidence for tetramers of low molecular weight protein tyrosine phosphatase in solution, J. Am. Chem. Soc. 125 (2003) 916–923.
[11] D. Fushman, Determination of protein dynamics using ^{15}N relaxation measurements, in: O. Zerbe (Ed.), BioNMR in Drug Research, Wiley-VCH, New York, 2002, pp. 283–308.
[12] N. Tjandra, S.E. Feller, R.W. Pastor, A. Bax, Rotational diffusion anisotropy of human ubiquitin from ^{15}N NMR relaxation, J. Am. Chem. Soc. 117 (1995) 12562–12566.
[13] R. Bruschweiler, X. Liao, P.E. Wright, Long-range motional restrictions in a multidomain zinc-finger protein from anisotropic tumbling, Science 268 (1995) 886–889.
[14] L.K. Lee, M. Rance, W.J. Chazin, 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 (1997) 287–298.
[15] M. Blackledge, F. Cordier, P. Dosset, D. Marion, Precision and uncertainty in the characterization of anisotropic rotational diffusion by ^{15}N relaxation, J. Am. Chem. Soc. 120 (1998) 4538–4539.
[16] V. Copie, Y. Tomita, S.K. Akiyama, S. Aota, K.M. Yamada, R.M. Venable, R.W. Pastor, S. Krueger, 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 (1998) 663–682.
[17] G.M. Clore, A.M. Gronenborn, A. Szabo, 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 (1998) 4889–4890.
[18] P. Tsan, J.C. Hus, M. Caffrey, D. Marion, M. Blackledge, Rotational diffusion anisotropy and local backbone dynamics of carbon monoxide-bound *Rhodobacter capsulatus* cytochrome *c*, J. Am. Chem. Soc. 122 (2000) 5603–5612.
[19] R. Ghose, D. Fushman, D. Cowburn, 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, J. Magn. Reson. 149 (2001) 214–217.
[20] P. Dosset, J.C. Hus, M. Blackledge, D. Marion, Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data, J. Biomol. NMR 16 (2000) 23–28.
[21] J.G. de la Torre, S. Navarro, M.C.L. Martinez, F.G. Diaz, J.J.L. Cascales, HYDRO: a computer program for the prediction of hydrodynamic properties of macromolecules, Biophys. J. 67 (1994) 530–531.
[22] J. García de la Torre, M.L. Huertas, B. Carrasco, HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations, J. Magn. Reson. B 147 (2000) 138–146.
[23] M. Andrec, K.G. Inman, D.J. Weber, R.M. Levy, G.T. Montelione, A Bayesian statistical method for the detection and quantification of rotational diffusion anisotropy from NMR relaxation data, J. Magn. Reson. 146 (2000) 66–80.
[24] D. Fushman, D. Cowburn, Nuclear magnetic resonance relaxation in determination of residue-specific ^{15}N chemical shift tensors in proteins in solution: protein dynamics, structure, and applications of transverse relaxation optimized spectroscopy, in: T. James, U. Schmitz, V. Doetsch (Eds.), Methods in Enzymology, vol. 339, 2001, pp. 109–126.
[25] D. Fushman, T. Najmabadi-Haske, S. Cahill, J. Zheng, H.R. LeVine, D. Cowburn, The solution structure and dynamics of the pleckstrin homology domain of G protein-coupled receptor kinase 2 (beta-adrenergic receptor kinase 1). A binding partner of Gbetagamma subunits, J. Biol. Chem. 273 (1998) 2835–2843.
[26] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in C, Cambridge University Press, New York, 1992.
[27] G.M. Clore, D.S. Garrett, R-factor, free R, and complete cross-validation for dipolar coupling refinement of NMR structures, J. Am. Chem. Soc. 121 (1999) 9008–9012.
[28] P.R. Bevington, D.K. Robinson, Data reduction and error analysis for the physical sciences, second ed., McGraw-Hill, New York, 1992, pp. 205–209.
[29] D. Fushman, R. Ghose, D. Cowburn, The effect of finite sampling on the determination of orientational properties: a theoretical treatment with application to interatomic vectors in proteins, J. Am. Chem. Soc. 122 (2000) 10640–10649.
[30] G. Cornilescu, J.L. Marquardt, M. Ottiger, A. Bax, Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase, J. Am. Chem. Soc. 120 (1998) 6836–6837.