Journal of Biomolecular NMR **27:** 291–312, 2003.
KLUWER/ESCOM

Detection of nano-second internal motion and determination of overall tumbling times independent of the time scale of internal motion in proteins from NMR relaxation data

Göran Larsson, Gary Martinez, Jürgen Schleucher & Sybren S. Wijmenga*,**
Department of Medical Biochemistry and Biophysics Umeå University, SE-901 87 Umeå, Sweden

Received 3 February 2003; Accepted 11 June 2003

Key words: chemical shifts anisotropy, internal motion, ^{15}N relaxation, nano-second time scale, overall rotation correlation time, overall tumbling time, protein NMR, relaxation

Abstract

The usual analysis of ^{15}N relaxation data of proteins is straightforward as long as the assumption can be made that the backbone of most residues only undergoes fast (ps), small amplitude internal motions. If this assumption cannot be made, as for example for proteins which undergo domain motions or for unfolded or partially folded proteins, one needs a method to establish for each residue whether it undergoes fast (ps) or slow (ns) internal motion. Even then it is impossible to determine the correct overall tumbling time, τ_m^0, via the usual method from the ratio of the longitudinal and transverse relaxation times, if the majority of residues do not undergo fast, small amplitude internal motions. The latter problem is solved when τ_m^0 can be determined independent of the time scale, τ_i, or the amplitude, S^2, of the internal motion. We propose a new protocol, called PINATA, for analyzing ^{15}N relaxation data acquired at minimally two field strengths, where no a priori assumption about time scales or amplitude of internal motions needs to be made, and overall tumbling can either be isotropic or anisotropic. The protocol involves four steps. First, for each residue, it is detected whether it undergoes ps- or ns-internal motion, via the combination of the ratio of the longitudinal relaxation time at two fields and the hetero-nuclear NOE. Second, for each residue τ_m^0 and the exchange broadening, Rex, are iteratively determined. The accuracy of the determination of τ_m^0 is ca. $\pm$ 0.5 ns and of Rex ca $\pm$ 0.7 s^{-1}, when the relaxation data are of good quality and $\tau_m^0 > 5$ ns, $S^2 > 0.3$, and $\tau_i <\approx 3$ ns. Third, given τ_m^0 and Rex, step 1 is repeated to iteratively improve on the internal motion and obtain better estimates of the internal parameter values. Fourth, final time scales and amplitudes for internal motions are determined via grid search based fitting and χ^2-analysis. The protocol was successfully tested on synthetic and experimental data sets. The synthetic dataset mimics internal motions on either fast or slow time scales, or a combination of both, of either small- or large amplitude, superimposed onto anisotropic overall motion. The procedures are incorporated into MATLAB scripts, which are available on request.

Abbreviations: τ_m^0 – rotation correlation time; τ_m^{ap} – apparent rotation correlation time; T_1 – longitudinal relaxation time; T_2 – transverse relaxation time; R_1 – longitudinal relaxation rate; R_2 – transverse relaxation rate; NOE – Nuclear Overhauser Enhancement. T_1^X, T_2^X, and NOE^X are T_1, T_2, and NOE at a field of X MHz ^{1}H frequency, and similarly for any other relevant terms, e.g. R_1, R_2, Rex, τ_m^{ap}, etc. When for any of these terms the field is not specified the higher field is meant, e.g., NOE means NOE^{hf} and τ_m^{ap} means $\tau_m^{ap\_hf}$; where hf is the higher field; lf is the lower field. RT2 is defined as R_2^{hf} / R_2^{lf}. RT1 is R_1^{lf} / R_1^{hf}. $R_{\tau mapp}$ is $\tau_m^{ap\_hf}$ / $\tau_m^{ap\_lf}$. Rex, broadening due to conformational exchange. $R_{\tau mapp}n$ – $R_{\tau mapp}$ corrected for ps-im. RT1n – RT1 corrected for ps-im internal motion and normalized by its rigid limit value RT1^0. The term ns-im is broadly defined as internal motion with τ_i larger than 200 ps, while ps-im is defined as internal motion with τ_i is smaller than 200 ps; τ_i time scale for internal motion; S^2 – squared order parameter.

Introduction

The study of backbone dynamics via the relaxation of ^{15}N nuclear spins is now a standard component of the structural characterization of proteins by means of NMR (Ishima and Torchia, 2000; Kay, 1998; Korzhnev et al., 2001). Although many modifications both for the measurement as well as for the analysis of ^{15}N relaxation data have been described (Farrow et al., 1994; Lee and Wand, 1999; Mandel et al., 1995; Tjandra et al., 1995, 1996; Jin et al., 1998; Humphrey

*Present address: Laboratory of Biophysical Chemistry, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands.
**To whom correspondence should be addressed. E-mail: sybrenw@sci.kun.nl

et al., 1996) the main line, originally described by Kay et al. (1998) and later extended by Clore and co-workers (Clore et al., 1990a,b), essentially remains the same. After measurement of the experimental longitudinal (T_1) and transverse (T_2) relaxation times as well as the heteronuclear NOE at one, two or more fields, the relaxation data is interpreted in terms of the Lipari–Szabo (LS) model-free (Lipari and Szabo, 1982), or extended Lipari–Szabo (extended-LS) model-free (Clore et al., 1990b) approach in two steps (Farrow et al., 1994; Mandel et al., 1995; Tjandra et al., 1996).

In the first step the overall tumbling time (τ_m^0) is estimated, since a correct estimate of τ_m^0 is a primary requirement for the further analysis (Mandel et al., 1995; Tjandra et al., 1996; Korzhnev et al., 1997; Lee and Wand, 1999). Usually, τ_m^0 is estimated from the apparent overall tumbling time (τ_m^{ap}) which, for each residue, is calculated from the ratio of their T_1 and T_2 relaxation times (Farrow et al., 1994; Kay et al., 1989). For τ_m^{ap} to correctly estimate τ_m^0, the internal motion (τ_i) should be in the extreme narrowing limit, $(\omega\tau_i)^2 \ll 1$, and of a small amplitude. To select these residues two filters are usually employed (Tjandra et al., 1996; Lefevre et al., 1996). The first filter selects residues with a high NOE value (NOE $>$ 0.6), and second filter selects those residues that have T_1 and/or T_2 values close to the average T_1 and T_2. The remaining residues are then taken to have no exchange (Rex $= 0$) and to be affected only by fast ps-time scale motion of a limited degree. For these residues a τ_m^0 can then reliably be calculated from the ration of T_1 and T_2. The remaining residues are assumed to have a τ_m^0, which is equal to the average τ_m^0.

The second step uses the τ_m^0 (average and/or residue-specific) determined in the first step. Given τ_m^0, a LS description or extended-LS description of internal motion is derived. This step (Mandel et al., 1995) consists of mathematical optimization via the χ^2 minimization of the model-free parameters against the relaxation data, and selection of a motional model based on statistics. These two steps finally result in numbers for τ_m^0, the squared order parameters (S^2) and the time scales of internal motion (τ_i). A graphical procedure to extract the model-free parameters has also been suggested (Jin et al., 1998).

The analysis mentioned above is straightforward under the assumption that the backbone of most residues only undergoes fast (ps), small amplitude internal motions (Baber et al., 2001). However, if this assumption is incorrect or cannot be made, for example for proteins that undergo domain motions (Lefevre et al., 1996; Baber et al., 2001; Zdunek et al., 2003; Larsson et al., 2003) or for unfolded or partially folded proteins (Wright and Dyson, 1999; Dyson and Wright, 1998; Farrow et al., 1997), the analysis can fail. This is essentially for two reasons. First, the NOE filter cannot unequivocally establish t_i, because high NOE values (> 0.6) arise for residues with fast, ps-internal motion, as well as from residues with slow, ns-internal motion (Korzhnev et al., 2001). Since the second filter is based on deviations from the mean T_1 and T_2, this filter cannot remove residues affected by ns-internal motion, when a large number of residues are affected. Thus, residues may be selected with internal motion outside the extreme narrowing limit. Second, a correct estimate of τ_m^0 can only be made for residues with fast ps internal motion. Outside the extreme narrowing limit τ_m^{ap} underestimates τ_m^0. A wrong estimate of τ_m^0 (either on average or on a residue-specific basis) leads to wrong LS parameters and can even lead to a physically incorrect model for internal motion (Korzhnev et al., 1997).

Thus, to reliably analyze the ^{15}N relaxation data of proteins whose residues undergo ns-time scale internal motion two main problems need to be solved. First, a procedure that reliably detects the presence or absence of ns-time scale motion is needed. Second, τ_m^0 needs to be determined independent of the time scale of internal motion. In this paper we propose a practical alternative method for analyzing ^{15}N relaxation where no assumptions need to be made about time scales or amplitudes of internal motions. It requires that the relaxation parameters are measured at least at two or more magnetic fields. The protocol is based on the notion that whatever the time scale of internal motion, model selection, i.e., determination of its parameters can essentially be separated from the determination of τ_m^0, by focusing on the field dependence of the relaxation data. In addition, the motional model is easily assessed by inspection of simple two-dimensional graphs. The protocol consists of four iterative steps. It has been successfully tested on synthetic data that mimic ps or ns internal motion or a combination of both time scales, either with large or small motional amplitudes, superimposed on anisotropic overall motion. It is also demonstrated on published experimental relaxation data. The protocol, called PINATA, is implemented in MATLAB scripts, which are available on request. Although multiple-field NMR relaxation studies has been carried out before, they have mainly been focused on ^{15}N chemical shift anisotropy (Fush-

man et al., 1999; Canet et al., 2001), or spectral density mapping (Peng and Wagner, 1995; Papavoine et al., 1997; Vis et al., 1998), or assessing potential bias in the determination of τ_m of proteins (Lee and Wand, 1999). We have focused on the field dependence of the separate relaxation parameters, T_1, T_2, and the hetero nuclear NOE, which allows us to separate different timescales of motions for a ^{15}N nuclei in a protein backbone.

Theory

The equations for ^{15}N relaxation and its interpretation in terms of the Lipari–Szabo formalism (Lipari and Szabo, 1982) have been discussed extensively in the literature (Abragam, 1961; Lipari and Szabo, 1982; Farrow et al., 1994; Tjandra et al., 1996; Jin et al., 1998; Lefevre et al., 1996; Korzhnev et al., 1997). Here, we summarize those aspects most important for our protocol.

The longitudinal relaxation time (T_1), the transverse relaxation time (T_2), and the hetero nuclear NOE are given by (Abragam, 1961):

$$\begin{aligned} R_1 &= 1/T_1 = (3d + c)J(\omega_N) \\ &\quad + d(J(\omega_H - \omega_N) + 6J(\omega_H + \omega_N)), \end{aligned} \quad (1)$$

$$\begin{aligned} R_2 &= 1/T_2 \\ &= \tfrac{1}{6}(3d + c)(4J(0) + 3J(\omega_N)) \\ &\quad + \tfrac{1}{2}d(J(\omega_H - \omega_N) + 6J(\omega_H + \omega_N) \\ &\quad + 6J(\omega_H)) + \mathrm{Re}x, \end{aligned} \quad (2)$$

$$\begin{aligned} \mathrm{NOE} &= 1 + \frac{\gamma_H}{\gamma_N}d(-J(\omega_H - \omega_N) \\ &\quad + 6J(\omega_H + \omega_N))^* T_1. \end{aligned} \quad (3)$$

Here,

$$d = \left(\frac{\gamma_N \gamma_H h}{r_{HN}^3 2\pi}\right)^2 \left(\frac{\mu_0}{4\pi}\right)^2$$

and

$$c = \frac{2}{15}\omega_N^2 \Delta\sigma^2,$$

γ_N and γ_H are the ^{15}N and ^{1}H gyromagnetic ratio's, ω_N and ω_H are the corresponding angular resonance frequencies, h is Planck's constant, r_{HN} is the internuclear H-N distance assumed to be 1.02 Å, μ_0 is the permeability of free space, and $\Delta\sigma$ is the ^{15}N chemical shift anisotropy commonly assumed to be -170 ppm. The term Rex describes the additional broadening due to conformational exchange in the μs/ms-time range. The spectral density function J(ω) is defined as the Fourier transform of the total rotational auto-correlation function $C(t)$. We have observed that the simplified equations used in reduced spectral density mapping (Farrow et al., 1995; Lefevre et al., 1996; Ishima et al., 1995; Ishima and Nagayama, 1995a,b) are quite accurate and might as well have been used.

In general, any auto-correlation $C(t)$ can be approximated by a sum of exponentially decaying terms (see e.g., Lipari and Szabo, 1982; Viles et al., 2001). Although mathematically correct, such an approach may obscure a physically meaningful interpretation of the derived parameters if the number of exponentially decaying terms becomes too large. Lipari and Szabo (1982) have shown that the total rotational auto-correlation function $C(t)$ can be approximated as a simple sum of few exponential terms and that the parameters can be interpreted in a physically meaningful manner without invoking a specific motional model for the internal motion. They also established the conditions under which this 'model-free' approach is exact or approximate. The first assumption they make is that any $C(t)$ can be written as a product of a correlation function for overall tumbling $C_0(t)$ and one for internal motion $C_I(t)$ (Lipari and Szabo, 1982):

$$C(t) = C_0(t)C_I(t). \quad (4)$$

This factorization was shown to be rigorously correct if the internal motions and overall tumbling are not correlated, and the overall tumbling is isotropic. Equation 4 is not rigorously correct in case of anisotropic overall tumbling, even when internal and overall motions are uncoupled. However, it is a good approximation when the overall motion is axially symmetric and the internal motion is sufficiently fast (Lipari and Szabo, 1982; Baber et al., 2001; Schurr et al., 1994). In the original Lipari–Szabo formalism, $C_I(t)$ is approximated by the contribution on one time scale only. The corresponding spectral density function is then exact when all the internal motions are in the extreme narrowing limit. When internal motions occur on both fast (extreme narrowing limit) and slower time scales there is no general rigorous and exact description of $C_I(t)$. The simplest description is then the extended form of Lipari–Szabo formalism (extended LS) (Clore et al., 1990b), where $C_I(t)$ is approximated by the

sum of two exponential terms, describing the fast and slow time scales, respectively. This extension was originally proposed to describe backbone residues that undergo fast vibrational motions as well as slower motions due to dihedral angle transitions. The assumption that overall and internal motions are decoupled is then a good approximation (Lipari and Szabo, 1982). However, when the slow internal motion is the reorientation of an entire domain, the decoupling assumption is not rigorous, but difficult to avoid, as pointed out by Baber et al. (2001). Meirovitch and coworkers (Tugarinov et al., 2001) have developed a theory (SRLS) for isotropically tumbling molecules in where internal motion and overall tumbling are fully coupled. The SRLS theory still converges to the Lipari–Szabo formalism in the fast motional limit and of course in the rigid limit. Numerical simulations show that a Lipari–Szabo analysis (assuming decoupling), while coupling is present as described via SRLS, overestimates the squared-order parameter when the internal motion is on ns-time scale.

In PINATA the commonly used (extended) Lipari–Szabo approach to internal motion (Equation 4) has been implemented which is correct with the caveats discussed above. Internal motion on both fast and slow time scales and (axially symmetric) anisotropic overall motion can then be treated by describing $C(t)$ via Equation 4 and by taking for $C_I(t)$ the extended LS and for $C_0(t)$ the equations for (axially symmetric) anisotropic overall tumbling (Woessner, 1962). For anisotropic overall tumbling this yields an overall correlation function that is dependent on the orientation of the N-H bond vector with respect to the rotational diffusion tensor (Woessner, 1962). This treatment yields (for axially symmetric anisotropic tumbling) the following spectral density function for a ^{15}N amide nitrogen (Lipari and Szabo, 1982; Schurr et al., 1994; Tjandra et al., 1995; Baber et al., 2001; Korzhnev et al., 2001):

$$\begin{aligned} J(\omega) \;=\; & S_f^2 S_s^2 A_1 J_1(\omega, \tau_{m1}^0) \\ & + (1 - S_f^2) A_1 J_1(\omega, \tau_{1ef}) \\ & + S_f^2 (1 - S_s^2) A_1 J_1(\omega, \tau_{1es}) \\ & + S_f^2 S_s^2 A_2 J_2(\omega, \tau_{m2}^0) \\ & + (1 - S_f^2) A_2 J_2(\omega, \tau_{2ef}) \\ & + S_f^2 (1 - S_s^2) A_2 J_2(\omega, \tau_{2es}) \\ & + S_f^2 S_s^2 A_3 J_3(\omega, \tau_{m3}^0) \\ & + (1 - S_f^2) A_3 J_3(\omega, \tau_{3ef}) \\ & + S_f^2 (1 - S_s^2) A_3 J_3(\omega, \tau_{3es}). \end{aligned} \tag{5}$$

Here, A_x is defined as

$$\begin{aligned} A_1 \;&=\; 0.25(3\cos^2\Phi - 1)^2; \\ A_2 \;&=\; 3\cos^2\Phi \sin^2\Phi; \\ A_3 \;&=\; 0.75\sin^4\Phi \end{aligned} \tag{6}$$

and $J_x(w, t_x)$ as,

$$J_x(\omega, \tau_x) = \frac{2}{5}\frac{\tau_x}{1 + (\omega\tau_x)^2}. \tag{7}$$

The time constants for overall tumbling τ_{mx}^0 are defined as, $\tau_{m1}^0 = \tau_l^0$, $\tau_{m2}^0 = 6\tau_l^0\tau_s^0/(5\tau_s^0 + \tau_l^0)$, and $\tau_{m3}^0 = 3\tau_l^0\tau_s^0/(\tau_s^0 + 2\tau_l^0)$. Here, τ_l^0 and τ_s^0 are the tumbling times of the long and short axis of the diffusion tensor and the ratio $\tau_l^0/\tau_s^0 = D\perp/D\parallel$ defines the anisotropy of the overall tumbling. The orientation of the relaxation vector with respect to the long axis (z-axis) of the diffusion tensor is given by the angle Φ. We note in passing that non-axially symmetric anisotropic overall tumbling (diffusion tensor with three different components) adds two more terms to the description of the overall tumbling (five terms, A1–A5). The spectral density then not only depends on the angle Φ but also on the angle the N-H vector makes with the x-axis of the diffusion tensor. The time constants for the two internal motion contributions, τ_{xef} and τ_{xes}, are given by $\tau_{xef}^{-1} = (\tau_{mx}^0)^{-1} + \tau_{if}^{-1}$ and $\tau_{xes}^{-1} = (\tau_{mx}^0)^{-1} + \tau_{is}^{-1}$, respectively, with τ_{if} and τ_{is} the time constants for fast and slow internal motion. S_f^2 and S_s^2 are the order parameters of the two contributions with $S^2 = S_f^2 S_s^2$. Simpler motional models, such as for isotropic motion and/or for only one internal motion are simplifications of Equation 5. Higher-order approximations of the internal correlation function than the extended-LS description (three contributions or more) have not been invoked (see Section I of the Results and discussion). From the definition of τ_{xef} and τ_{xes} it follows that internal motions much slower than τ_{mx}^0 do not affect $J(\omega)$. This reflects the physical notion that if there is an independent overall motion common to all parts of the molecule, there cannot be a component in the spectral density that decays more slowly than the overall motion.

The apparent overall rotation correlation time is calculated from the ratio R_2 over R_1 and can be expressed as:

$$\tau_m^{ap} = \frac{1}{\omega_N}\sqrt{\frac{3}{2(1+a)}\left(\frac{R_2}{R_1} - \frac{7}{6}(1+a)\right)} \tag{8}$$

with $a = -0.02$. This equation is similar to the one employed by Farrow et al. (1995). In the absence of internal motion or when the time constants for internal motion are zero ($\tau_{is/f} = 0$), Equation 8 correctly estimates τ_m^0 within 1%. Thus, in the absence of internal motion and for small degrees of anisotropy, τ_m^{ap} is very close to the true correlation time, τ_m^0, and from Equations 5 and 8 it follows that $\tau_m^{ap} = \tau_m^0$ can be expressed as:

$$\tau_m^0 = \frac{\tau_l^0}{1 + \frac{\Delta}{2}\sin^2(\Phi)}, \tag{9}$$

where $\Delta = (\tau_l^0/\tau_s^0) - 1$. Hence, for small degrees of anisotropy, τ_m^{ap} contains structure information, namely on the angle Φ each relaxation vector (N-H vector) makes with the long axis of the diffusion tensor. Thus, in the case of anisotropic tumbling, τ_m^0 represents a residue-specific effective overall tumbling time that carries information about the relative orientations. The orientation of the N-H vectors give long-range structural information that can be used to orient individual domains of proteins (Zdunek et al., 2003). When the overall tumbling is isotropic, all residues have same overall tumbling time τ_m^0, and the orientation information from Equation (9) is lost.

Material and methods

The relaxation data acquired at two or more magnetic fields have been analyzed using PINATA described here, which consists of scripts written for Matlab Version 5.1. PINATA has been successfully tested on published relaxation data on M13 coat protein (gVIIIp) measured at 500, 600 and 750 MHz (Papavoine et al., 1997, 1998), as well as on synthetic relaxation data. The performance and the different steps of the Matlab protocol are described in detail in the Result and discussion section.

We have used the full equations (1)–(5) to generate synthetic T_1, T_2 and NOE data at 600 and 400 MHz ^{1}H frequency. The data was generated with Matlab. Different settings of τ_m^0, τ_i (τ_{if}, τ_{is}), S^2 (S_f^2, S_s^2) and Rex were used to produce the synthetic data (see Table 1).

Results and discussion

The results are presented and discussed in six sections. Section I provides a theoretical description of which motional parameters can be separately derived from relaxation data at multiple fields. Section II gives a flowchart of the actual analysis protocol. In Sections III, IV and V each step in the protocol is discussed in detail. The numerical tests are found in section IVc. Section VI describes the demonstration on published relaxation data.

I. Separation of motional parameters

Given a model for internal and overall motion, the relaxation data, R_1, R_2 and NOE can be calculated exactly from equations 1 to 5 for one or more fields. The reverse, the derivation of motional parameters (τ_m^0, S^2, τ_i, Rex, etc.) from the relaxation data (R_1, R_2 and NOE) is more complex, because it requires a mathematical fitting procedure. The complex dependence of the motional parameters on the relaxation data makes this fitting very difficult. We have (re) analyzed the interdependence of motional parameters and relaxation data in the context of the extended Lipari–Szabo model superimposed onto anisotropic overall motion (Equation (5)) and come to the following three conclusions with regard to the separation of motional parameters.

The qualitative information that the internal motion is purely ps-im ($\tau_i < 200$ ps) or that ns-im ($\tau_i > 200$ ps) is at least mixed in, can graphically be derived from $R_{\tau\text{mapp}}$ (or RT1) versus NOE plots, without the need to precisely specify τ_m^0 (Sections IIIa and IIIb). When pure ps-im is present, the condition $\omega\tau_e \ll 1$ generally also holds true, and the one-exponential approximation of $C_I(t)$ (or rather its Fourier transform, the spectral density function) is exact (Lipari and Szabo, 1982), i.e., whatever the complexity of $C_I(t)$ the corresponding spectral density functions cannot be distinguished. A one-contribution model for internal motion then always suffices, and $\tau_e (= \tau_i$ when $\tau_i \ll \tau_m^0)$ should then be interpreted as a S_k^2-weighted average time constant of the k internal motions. However, when ns-im is at least mixed in, $C_I(t)$ described by a one- or a two-exponential approximation (or more terms) leads to different spectral density functions. Consequently, a correct description of the internal motion may require more than one time scale. The discrimination between a one- and two-exponential contribution model and the determination of their parameters ($\{S_f^2, \tau_{if}\}$, $\{S_s^2, \tau_{is}\}$), can then be based on $R_{\tau\text{mapp}}$ (or RT1) and NOE, in combination with the R_1 value (Section IIIc).

Table 1. Parameters used to generate synthetic T_1, T_2 and ^{1}H-^{15}N hetero-nuclear NOE data

τ_m^0 (ns)	4.0	6.0	8.0	10.0	12.0	14.0	16.0
$\tau_{i(f,s)}$ (ps)	20	150	400	700	1000	1700	2000
$S^2_{(f,s)}$	0.4	0.6	0.8	1.0			
Rex (s^{-1})	0	2	12				
^{15}N CSA (ppm)	−150	−170	−190	−200			

A description of ns-im via a one-contribution model can be viewed as a first-order approximation; the derived τ_i and S^2 then represent average of the actual time scales and squared-order parameters of the different actual ns-im contributions. A description of ns-im via a two-contribution model is then a second-order approximation. To derive higher order approximations requires more and increasingly precise and accurate relaxation data. In practice it turns out that a two-contribution model with one ps-im contribution and one ns-im contribution, generally leads to a good fit to measured relaxation data (Mandel et al. 1995; Jin et al., 1998). Our analysis leads to same conclusion (e.g., see Section III). We therefore implemented in PINATA only a description of internal motion with a maximum of two-contributions.

The real overall rotation correlation time, τ_m^0, can be determined for each residue from the combination of τ_m^{ap}, $R_{\tau map}$ (or RT1), and NOE without the need to specify the number of terms in $C_I(t)$, and its parameters τ_i and S^2. The terms τ_m^{ap}, $R_{\tau map}$ (or RT1), and NOE form a more or less isolated parameter subspace that is used to determine the real residue-specific τ_m^0 independent of time scale and amplitude of internal motion (Sections IVb and IVc). The variation in the residue-specific τ_m^0 can be used to determine the shape of the diffusion tensor of the investigated molecule (Clore et al., 1998).

The ratio of R_2 rates (RT2) is essentially independent of τ_i and largely independent of S^2 and τ_m^0. Therefore, RT2 can be used to determine the exchange rate, Rex, if Rex is in the fast exchange regime (Section IVa). Given Rex and τ_m^0, S^2 can accurately be determined from the transverse relaxation rate, without specifying the time constant(s) for internal motion (Section IVa).

Thus, instead of directly using the measured quantities R_1, R_2, and NOE at two fields, it is better to focus on the field dependence of R_1 and R_2, i.e. RT2, RT1, and/or $R_{\tau\text{mapp}}$. In this way, certain motional parameters become largely separated. However, this separation is not strict. Therefore, any analysis protocol needs to be iterative (Section II).

II. Flowchart of the PINATA protocol

Based on the separation of parameters described above, the PINATA protocol consists of the following iterative steps (Figure 1). After reading in the experimental data, setting of some parameters (field strength, etc.), and plotting of the experimental data (step 0), the actual protocol starts.

Step 1. The experimental ($R_{\tau\text{mapp}}$, RT1 and NOE) data points are calculated and superimposed onto the theoretical $R_{\tau\text{mapp}}$ and/or RT1 vs. NOE graphs using a rough initial guess of t_m^0. From these plots it is directly visible whether a residue undergoes pure ps-im or that ns-im is at least mixed in. Thus, a qualitative determination of the internal motional model is obtained. In addition, rough parameters (S_f^2, τ_{if}, S_s^2, τ_{is}) can be derived.

Step 2. This is the core of the protocol. It entails an iterative determination of τ_m^0 and Rex for each residue by using the following experimental data: ([{τ_m^{ap}, $R_{\tau\text{mapp}}$ or RT1, NOE}, RT2]; the symbols '[' and '{' indicate the level of interdependence). Initially, τ_m^{ap} and $R\tau_{\text{mapp}}$ are calculated from the experimental data (from the ratio of R_2 over R_1) at the two fields. They are then corrected for ps-im, based on their respective NOE value, which gives τ_m^{ap-ps} and $R_{\tau\text{mapp}}n$. The $R_{\tau\text{mapp}}$n is subsequently used to correct τ_m^{ap-ps} for ns-im. This gives $\tau_m^{ap-ps-ns}$. Given $\tau_m^{ap-ps-ns}$ and a rough estimate for S^2 and τ_i (one-contribution model), a theoretical exchange free RT2 is calculated ($RT2^{\text{exfree}}$). Rex is determined from comparison of $RT2^{\text{exfree}}$ and the experimental RT2. This also gives exchange corrected R_2 values, R_2^{excor}. The R_2^{excor} values are then used to obtain new improved $\tau_m^{ap-ps-ns}$ values, which in turn are used to obtain improved R_2^{excor} (and Rex) in three iterations. The final $\tau_m^{ap-ps-ns}$ is a good estimate of τ_m^0, because the time scale of internal motion does not bias it.

Step 3. Given the estimate of the new τ_m^0 and R_2^{excor} step 1 is repeated, i.e., the motional model is established via $R_{\tau\text{mapp}}$ and RT1 vs. NOE plots.

Step 4. Via a grid search method the internal motional parameters (e.g. $\{S_f^2, \tau_{if}\}, \{S_s^2, \tau_{is}\}$) are refined using $\tau_m^{ap-ps-ns}$ (average or residue-specific). A classical χ^2-statistical analysis is then used to confirm whether a simple or more complex internal motion model is warranted. It is also possible to introduce the results of step 4 as input to step 2 and carry out a final iteration.

III. Determination of internal motion model (Step 1, Figure 1)

IIIa. Determination of internal motion model from $R_{\tau\text{mapp}}$ versus NOE graphs

Figure 2A shows an overlay of two $R_{\tau\text{mapp}}$ vs. NOE graphs. When the internal motion is fast ($\tau_i < 200$ ps), the contours with constant S^2 are strictly linearly dependent on the NOE and overlap in this regime, i.e., the slopes are independent of S^2. Moreover, these slopes are virtually independent of τ_m^0 (this is evident from the overlap of the linear part of the drawn and dotted S^2-contours calculated at 12 ns and 9 ns, respectively). In fact, the latter holds true as long as $\tau_m^0 > 6$ ns (data not shown). The linear dependence rather abruptly changes for τ_i outside the ps-im regime. In this ns-im regime ($\tau_i > 200$ ps), the S^2-contours reach their smallest value for $\tau_i \approx 1$ ns, and depend only weakly on τ_m^0 (compare drawn and dotted contours). We further note that the τ_i-contours in Figures 2A and 2C follow straight (dashed) lines extending from $R_{\tau\text{mapp}} = 1$ and NOE ≈ 0.82. In conclusion, as evident from the Figures 2A and C, the S^2-contours depend somewhat on τ_m^0, while the τ_i-lines essentially overlap for different τ_m^0 values.

It is convenient to correct the $R_{\tau\text{mapp}}$ values for their linear dependence on the NOE. In the $R_{\tau\text{mapp}}n$ vs. NOE graphs, the essentially constant linear slope in the ps-im regime is corrected (Figure 2C, Table 2). The S^2-contours now run horizontally in the ps-im regime. Consequently, when $R_{\tau\text{mapp}}n = 1$, τ_i is always faster than 200 ps, irrespective of the value of the NOE. On the other hand, when $R_{\tau\text{mapp}}n < 1$, τ_i is larger than 200 ps, or when a two-contribution internal motion model applies, it at least contains a contribution with $\tau_{is} > 200$ ps. Thus, the $R_{\tau\text{mapp}}n$ vs. NOE can be used to directly establish the motional model.

The spectral density functions are linear combinations of spectral density functions representing overall motion and internal motion with one (Lipari and Szabo, 1982) or two contributions (Lipari and Szabo, 1982; Clore et al., 1990b) or more. $R_{\tau\text{mapp}}$ and $R_{\tau\text{mapp}}n$ vs. NOE graphs can therefore also be regarded as linear combinations of different graphs representing different types of internal motion. For example, when in addition to the one-contribution internal motion, an additional ps-im of 20 ps with S^2 of 0.8 were present, the complete $R_{\tau\text{mapp}}n$ vs. NOE graph in Figure 2C would be moved horizontally to the left, i.e., to the point ($S^2 = 0.8$, $\tau_i = 20$ ps). Alternatively, when in addition ns-im were present, of say $S_s^2 = 0.8$ and $\tau_{is} \approx 2$ ns, the $R_{\tau\text{mapp}}n$ vs. NOE graph is moved down from $R_{\tau\text{mapp}}n = 1$ and NOE ≈ 0.82 along the τ_i-contour of 2 ns to the point $S^2 = 0.8$.

As shown in Figure 3B, $R_{\tau\text{mapp}}$ and thus $R_{\tau\text{mapp}}n$ are essentially independent of CSA, i.e., for -190 ppm $<$ CSA < -150 ppm, the $R_{\tau\text{mapp}}$ variation is smaller than 0.7%. Because $R_{\tau\text{mapp}}$ is essentially the product of RT2 and RT1, the opposing dependence of RT2 and RT1 on CSA (Figure 3A) is canceled out in $R_{\tau\text{mapp}}$.

Exchange broadening (Rex) may complicate the interpretation of the $R_{\tau\text{mapp}}n$ data because Rex increases $R_{\tau\text{mapp}}n$ whereas ns-im decreases $R_{\tau\text{mapp}}n$. The exchange broadening could therefore cancel out the effects of ns-im in the $R_{\tau\text{mapp}}n$ graph. Rex can, however, be determined reliably from RT2 (see section IVa below) and $R_{\tau\text{mapp}}n$ can be corrected for exchange ($R_{\tau\text{mapp}}n^{\text{excor}}$). Thus, the $R_{\tau\text{mapp}}n^{\text{excor}}$ graphs can reliably be used to determine the presence of ns-im.

In conclusion, $R_{\tau\text{mapp}}$ or $R_{\tau\text{mapp}}n$ vs. NOE graphs form master curves, as they are essentially independent of τ_m^0 and CSA. They can be used to detect the presence of ns-im independent of the fact whether the overall tumbling is isotropic or anisotropic, because $R_{\tau\text{mapp}}$ or $R_{\tau\text{mapp}}n$ vs. NOE is independent of τ_m^0. When $R_{\tau\text{mapp}}n < 1$ (depending on error margin, see below), it is safe to conclude that ns-im is at least mixed in. The relative error in the experimental $R_{\tau\text{mapp}}$ values is half the sum of the relative errors in the T_1 and T_2 values at the two fields ($R_{\tau\text{mapp}} \approx$ const (RT1 $\times$ RT2)$^{1/2}$). The errors in T_1 and T_2 can be estimated to be ca. 1% and 3%, respectively, for good-quality relaxation data. Thus, the error on $R_{\tau\text{mapp}}$ lies around 4%. In this way the error on the data points is directly evident. It is therefore clear which conclusions can be drawn: When $R_{\tau\text{mapp}}n < 0.96$, it can safely be concluded that there is ns-im. Note that this conclusion can be drawn independent of the value of the NOE.

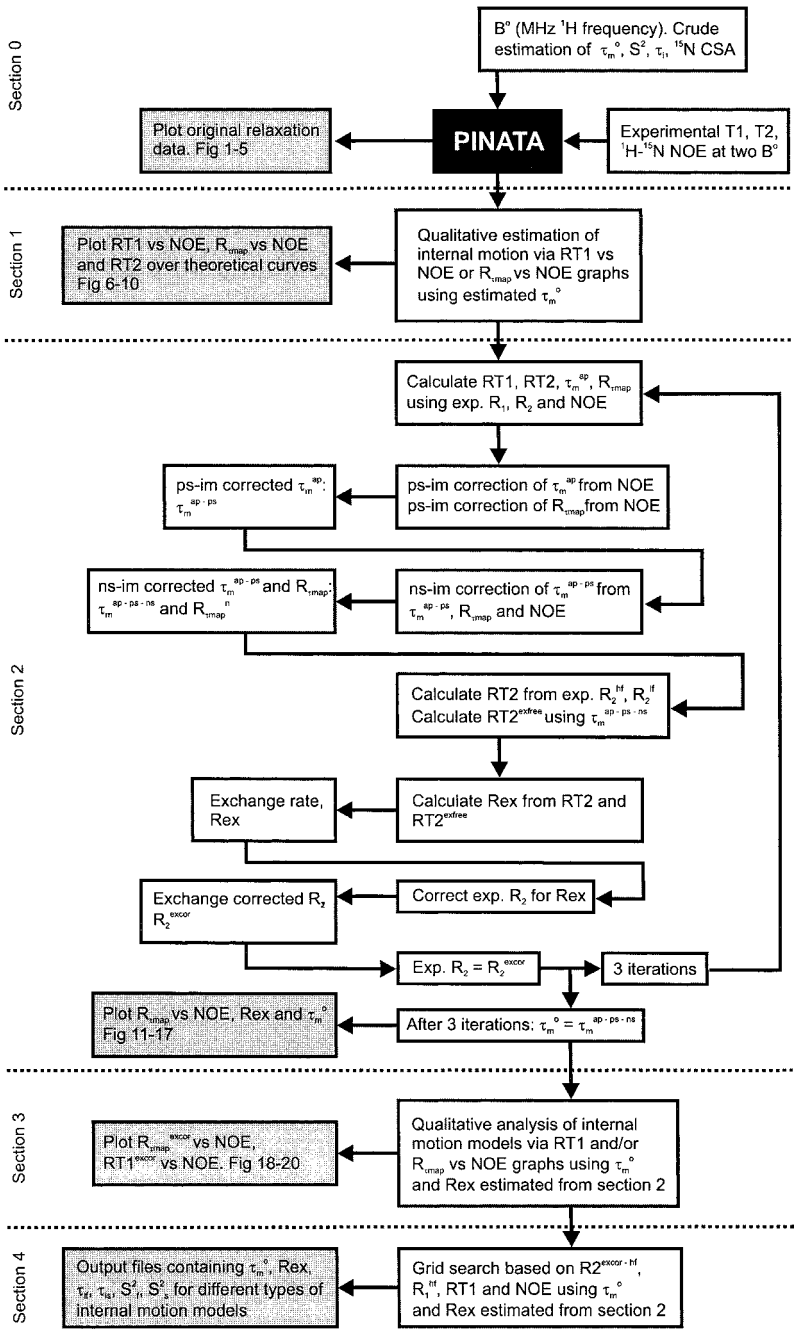

Figure 1. Flowchart of the analysis protocol PINATA. Section 0: Input of relaxation data, settings of some parameters, and plotting of original relaxation data. Section 1: Qualitative analysis of internal motion models via RT1 and/or $R_{\tau\mathrm{map}}$ vs. NOE graphs using a rough estimate of τ_m^0. Section 2: Residue specific iterative estimation of τ_m^0 and Rex. First τ_m^{ap} is calculated using experimental R_1 and R_2 values at the two fields (Equation (8)), using a one-contribution model (LS) using the S^2 and τ_i approximations set in section 0. Then $R_{\tau map}$ and τ_m^{ap} is corrected for ps internal motion (τ_m^{ap-ps}) and ns internal motion ($\tau_m^{ap-ps-ns}$). After ps and/or ns correction, the Rex contribution is calculated from RT2 and the experimental R_2 is corrected for Rex (R_2^{excor}), which is used to calculate new τ_m^{ap} and the whole section 2 is iteratively repeated 3 times. This finally gives $\tau_m^{ap-ps-ns}$ that can be estimated to be τ_m^0. Section 3: Qualitative analysis of internal motion models via RT1 and/or $R_{\tau\mathrm{map}}$ vs. NOE graphs using τ_m^0 and Rex estimated in Section 2. Section 4: Grid search based on $R_2^{\mathrm{excor-hf}}$, R_1^{hf}, RT1 and NOE to obtain all internal motion parameters (τ_{if}, τ_{is}, S_f^2 and S_s^2) using estimated τ_m^0 (residue specific or average) from Section 2.

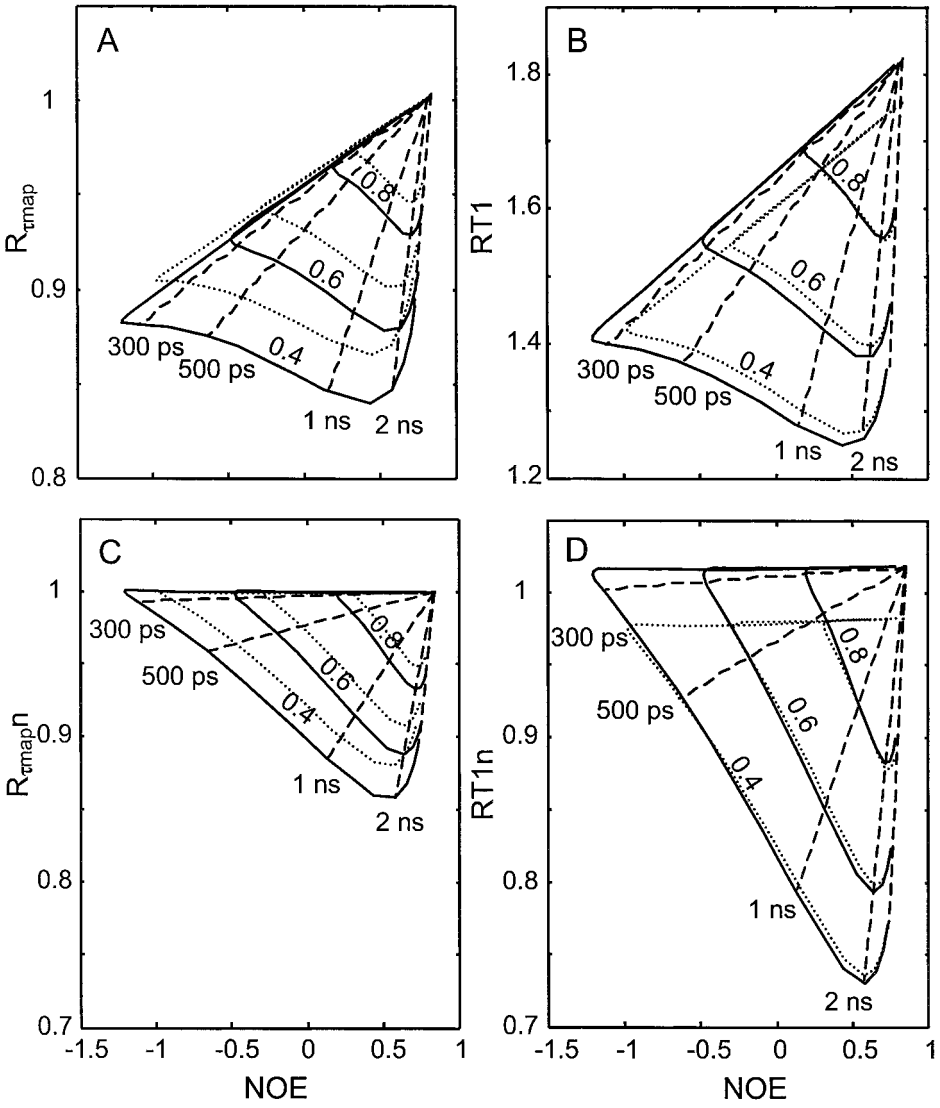

Figure 2. $R_{\tau mapp}$ ($= \tau_m^{ap\_600}/\tau_m^{ap\_400}$) (A) and RT1 ($= R_1^{400}/R_1^{600}$) (B) vs. the NOE^{600}. The normalized $R_{\tau mapp}n$ and RT1n are shown in panel C and D, respectively. The full equations (Equations 1–3 and 5) was used in the calculations, assuming a one internal motion model with an isotropic overall tumbling, no conformational exchange (Rex = 0) is assumed, and the ^{15}N CSA was set to −170 ppm. The data points were calculated for τ_i ranging from 20 ps to 6 ns and with S^2 ranging from 1 to 0.40. The solid contour lines connect points with constant S^2 (S^2-contours) for $\tau_m^0 = 12$ ns, and the dotted contour lines are S^2-contours for τ_m^0 of 9 ns. The values of 12 ns and 9 ns correspond to the effective tumbling time τ_m^0 of residues with the NH vector parallel or perpendicular to the long axis of an anisotropically tumbling molecule with an axial symmetry and an anisotropy of 1.5 ($=\tau_l/\tau_s$). The τ_i-contours (broken lines) are only shown for $\tau_m^0 = 12$ ns for τ_i equal to 300 ps, 500 ps, 1 ns and 2 ns (labels). The RT1n values are normalized by the rigid limit value RT1, RT1^0 (= RT1 when $S^2 = 1.0$ and $\tau_m^0 = 10$ ns).

$R_{\tau mapp}n < 0.96$ corresponds to a contribution of ca. 8% when $\tau_i = 2$ ns (Figure 2C).

IIIb. Determination of internal motion model from RT1 versus NOE graphs

The RT1 vs. NOE graphs (Figure 2B) have a shape very similar to that of the corresponding $R_{\tau mapp}$ vs. NOE graphs (Figure 2A). Analogous to the $R_{\tau mapp}$ the S^2-contours in RT1 depend linearly on the NOE as long as the internal motion is in the ps-im regime. Hence, RT1 can, like $R_{\tau mapp}$, be corrected for ps-im (Table 2). The normalized RT1n vs. NOE graph (Figure 2D) is again very similar to the $R_{\tau mapp}n$ vs. NOE graph (Figure 2C). As for $R_{\tau mapp}$, the RT1n

Table 2. Coefficients in the equations used for ps- and ns-im correction[a]

	1	2	3	4	5	6	7	8	9	10
c[#]	1.17	10	0.14	0.11	[c]					
d	0.5	0.0023	14	3.1×10^{-5}	0.0057	1.72×10^{-4}	0.001	1.3	0.9	
e[b]	18	6	0.164	[c]	10	0.003	88	15	0.25	0.15

[a]The (recursive) equations used in the corrections are:

$$\tau_m^{ap-ps} = \tau_m^{ap} + \left\{c_1 + (\tau_m^{ap-ps} - c_2) * c_3 + c_4 * (c_5 - \mathrm{NOE}^{hf})\right\} * (c_5 - \mathrm{NOE}^{hf})$$

$$R_{\tau\mathrm{map}}n = R_{\tau\mathrm{map}} + \left[1 + (1 - S_s^2) * d_1\right] * (\left[d_2 - (\tau_m^{ap-ps} * 10^9 - d_3)^2 * d_4\right] + \left\{\left[d_5 - (\tau_m^{ap-ps} * 10^9 - d_3)^2 * d_6\right] + d_7 * \left[d_8 - (\tau_m^{ap-ps} * 10^9 - d_3)^2 * d_9\right] * (c_5 - \mathrm{NOE}^{hf})\right\} * (c_5 - \mathrm{NOE}^{hf}))$$

$$\tau_m^{ap-ps-ns} = \tau_m^{ap-ps-ns} + \left\{\left[e_1 + (\tau_m^{ap-ps-ns} - e_2)^2 * e_3\right] + (e_4 - \mathrm{NOE}^{hf})^2 * \left[e_5 + (\tau_m^{ap-ps-ns} - e_2)^3 * e_6\right]\right\} * (1 - R_{\tau\mathrm{map}}n) + \left[e_7 + (e_4 - \mathrm{NOE}^{hf})^2 * e_8\right] * (1 - R_{\tau\mathrm{map}}n)^2 + e_9 * (e_4 - \mathrm{NOE}^{hf}) - e_{10}.$$

$RT1n = \{RT1 + (0.823 - \mathrm{NOE}^{hf}) * 0.2\}/RT1^0$, where $\mathrm{RT1}^0 = \mathrm{RT1}$ ($S^2 = 1$, $\tau_i = 0$) with estimated τ_m^0 as described in Section IIIb.

[b]The coefficients c and e are multiplied with 10^9.

[c]The coefficients c_5 and e_4 are 0.823 and 1.0, respectively.

values have been normalized to the rigid limit value ($S^2 = 1.0$) with $\mathrm{RT1}^0$ calculated for $\tau_m^0 = 10$ ns. The similarity of the $R_{\tau\mathrm{map}}n$ vs. NOE and RT1n vs. NOE graphs stems from the fact that $R_{\tau\mathrm{map}} \approx$ const $(\mathrm{RT1} \times \mathrm{RT2})^{1/2}$ and that RT2 is effectively independent of the time scale of internal motion. Thus, $R_{\tau\mathrm{map}} \approx$ const. $(\mathrm{RT1})^{1/2}$ vs. NOE is nearly coincident with $R_{\tau\mathrm{map}}n$ vs. NOE. Note however that the RT1n vs. NOE graph has a lower minimum than $R_{\tau\mathrm{map}}n$ vs. NOE (the minimum $R_{\tau\mathrm{map}}n$ is 0.86, Figure 2C; whereas that of RT1n is 0.74, Figure 2D). This immediately implies that that RT1n has a higher sensitivity to ns-im than $R_{\tau\mathrm{map}}n$. Moreover, the error on the experimental RT1n data points is smaller than for $R_{\tau\mathrm{map}}n$, because it derives from the relative error of T_1 at the two fields. Given, a 1% error in T_1, the estimated error in RT1n is 2%.

To investigate the dependence on CSA we calculated $\mathrm{RT1}^0$ as a function of CSA varying between -150 to -190 ppm (Figure 3B). At the extreme points of CSA (-150 and -190 ppm) $\mathrm{RT1}^0$ still remains within $\pm$ 3% of its value at CSA -170 ppm.

We have also investigated the variation in $\mathrm{RT1}^0$ with τ_m^0 (Figure 2D and 3A). At $\tau_m^0 = 6$ ns, $\mathrm{RT1}^0$ is ca. 9.5% below its value at 10 ns, while at 14 ns it is 3.8% above. This error can, however, be effectively reduced by improving the estimate of τ_m^0, which can be as low as $\pm$ 0.5 ns (Section IVc). A rough estimate of τ_m^0, within $\pm$ 1 ns, leads to a variation in $\mathrm{RT1}^0$ of only $\pm$ 3.8% at 6 ns and only 0.4% at 14 ns (Figure 3A). In view these error estimates on RT1n, it is safe to take RT1n < 0.96 as the detection limit for ns-im.

In conclusion, RT1n can be used to detect the presence of ns-im, independent of the fact whether the overall tumbling is isotropic or anisotropic (because of the weak dependence of RT1n on τ_m^0). The absence of an Rex effect on RT1n, the lower experimental error on RT1n, together with the higher sensitivity to the presence of ns-im, therefore makes RT1n a better parameter for detecting ns-im than $R_{\tau\mathrm{map}}n$.

IIIc. Determination of internal motion model and qualitative assessment of its parameters, summary

As follows from Sections IIIa and IIIb the internal motion model and its parameters can be determined from experimental RT1n (or $R_{\tau\mathrm{map}}n^{\mathrm{excor}}$) and NOE values independent of the fact whether the overall motion is isotropic or anisotropic; the following practical rules apply (see Figures 2 and 3).

- If RT1n > 0.96 (or $R_{\tau\mathrm{map}}n^{\mathrm{excor}} > 0.96$), only ps-im is present.
- If RT1n < 0.96 (or $R_{\tau\mathrm{map}}n^{\mathrm{excor}} < 0.96$), ns-im is at least mixed in.
- If NOE > 0.6 and RT1n < 0.96 (or $R_{\tau\mathrm{map}}n^{\mathrm{excor}} < 0.96$), $\tau_i > 1.0$ ns.

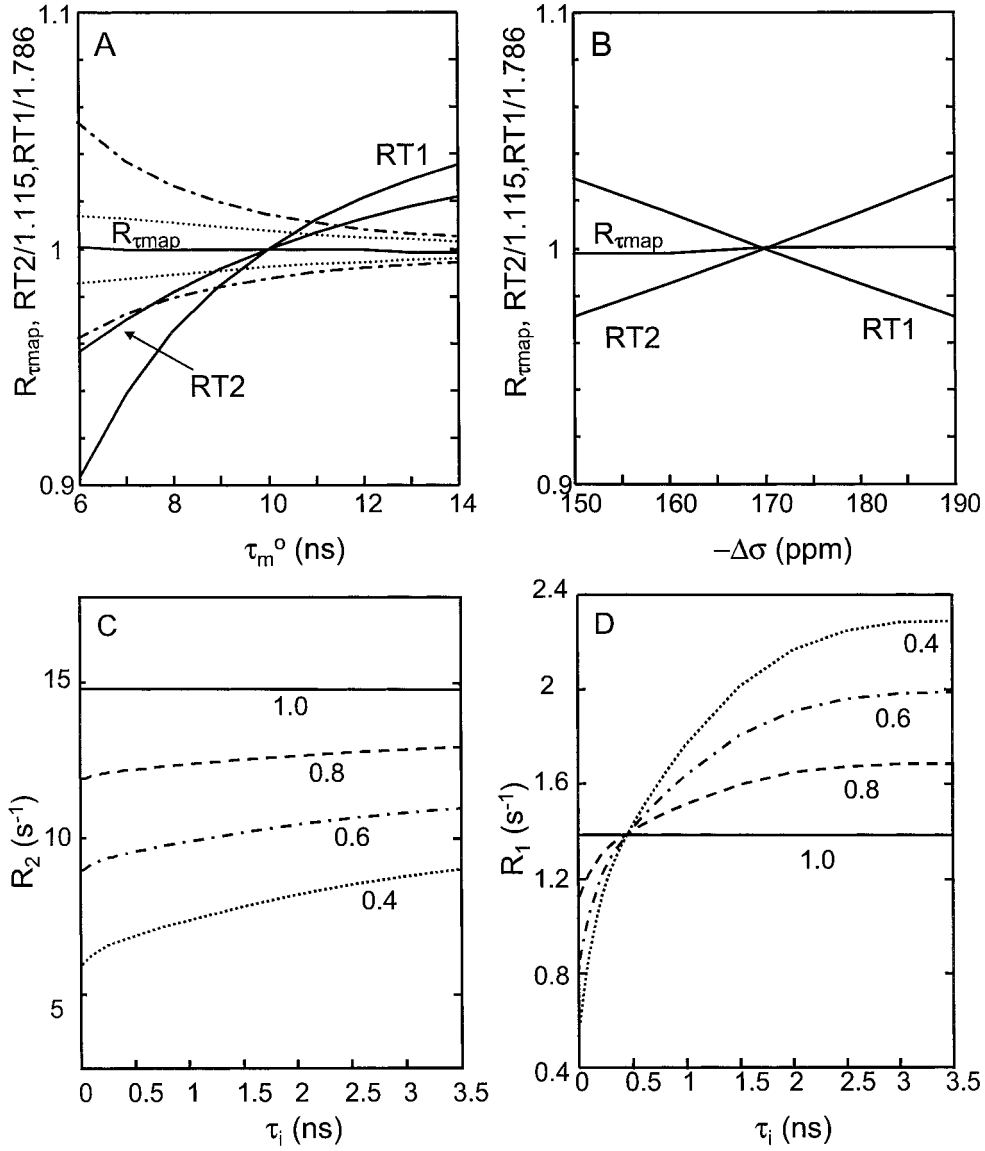

Figure 3. Dependence of $R_{\tau mapp}$, RT1, and RT2 at $S^2 = 1.0$ on t_m^0 indicated as labeled solid lines (A) and ^{15}N CSA (B). RT1 and RT2, with $S^2 = 1.0$, are normalized to their values at $t_m^0 = 10$ ns and CSA $= -170$ ppm, respectively. In panel A, the relative deviations of $R_{\tau mapp}$ (solid) and the normalized RT1 (dashed-dotted) and RT2 (dotted), when t_m^0 is ± 1 ns away from its actual value are also given. Field strengths of 400 and 600 MHz were used in the calculations. Panels C and D show the dependence of R_2 and R_1 on t_i for different S^2 values at 600 MHz; $t_m^0 = 10$ ns and CSA $= -170$ ppm. In the calculation the full equations was used (Equations (1)–(5)). Internal motion is represented by a one-contribution model with Rex $= 0$.

- If NOE < 0.0, $\tau_i < 1$ ns independent of the value of RT1n (or $R_{\tau mapp} n^{excor}$).
- If RT1n > 0.98 (or $R_{\tau mapp} n^{excor} > 0.98$) and NOE < 0, τ_i lies around 0.2 to 0.3 ns.

Note that the conclusion concerning the absence or presence of ns-im can be drawn from the value of RT1n (or $R_{\tau mapp} n^{excor}$) independent of the value of the NOE. The value of the NOE does only affect the combination of the exact time scale and amplitude of the internal motion. The motional parameters and the type of internal motion model can be estimated more precisely than in the above list based on the combination of RT1n (or $R_{\tau mapp} n$), NOE and the experimental R_1 value as illustrated via the following examples.

Suppose that RT1n is ca. 0.94 and the NOE is ca. 0.67. As can be seen in Figure 2D, this would correspond to a one contribution model with $S^2 = 0.75$ and $\tau_i = 1.3$ ns. The uncertainty in the values of S^2 and τ_i due to error in RT1n and NOE can directly be estimated from Figure 2D by simply establishing S^2 and τ_i values from Figure 2D using RT1n $\pm\,\sigma$. and NOE $\pm\,\sigma$. For a two contribution model, different combinations of solutions are possible given RT1n is ca. 0.94 and the NOE is ca. 0.67, e.g., one with $S_s^2 = 0.8$, $\tau_{is} = 1.7$ ns and $S_f^2 = 0.8$, $\tau_{if} = 0.02$ ns ($S^2 = S_f^2 \times S_s^2 = 0.64$). To decide whether the one- or two-contribution model applies, the information on the R_1 value can be used. R_1 is maximal for a one-contribution model, because an additional contribution of ps-im always reduces R_1 (Figure 3D). Here the R_1 at 600 MHz for the one-contribution model with $S^2 = 0.75$ and $\tau_i = 1.3$ ns is approximately 1.5 s^{-1}. Thus, if the experimental R_1 value equals 1.5 s^{-1} within experimental error, the one-contribution model applies. If R_1 is smaller than 1.5 s^{-1}, the internal motion consists of at least two contributions.

Finally, we note that contributions from additional time scales in the ns-im regime lead in principle to different spectral density functions (Section I). However, the above analysis shows that detection of such additional contributions beyond the two-contribution model is going to be difficult, since it constitutes a higher order approximation. To detect these additional time scales would require more relaxation data, e.g., at more fields, and of very high accuracy.

IV. Estimation of Rex and t_m^0 (Step 2, Figure 1)

IVa. Estimation of Rex from RT2

The R2 relaxation rate is roughly independent of τ_i when $\tau_m^0 > 6$ ns (Figure 3C, see also Jin et al. (1998)) and is well approximated by

$$R_2 = 1/T2 \approx \tfrac{4}{15}(3d + c)\left(S^2\tau_m^0 + (1 - S^2)\tau_e\right) + R_{\rm ex}. \quad (10)$$

Thus, R_2 depends mainly on τ_m^0, $S^2 (= S_s^2 \times S_f^2)$, and Rex. When Rex $= 0$, R_2 is proportional to the product of S^2 and τ_m^0. RT2 then becomes:

$$RT2 \approx (3d + c_{600})/(3d + c_{400}) \quad (11)$$

and is independent of S^2 and τ_m^0. RT2 is then solely determined by the ratio of the chemical shift anisotropy term at the two fields (c_{600} and c_{400}). The $\Delta\sigma$ (= CSA) may vary within -170 ± 20 ppm (Fushman et al., 1998, 1999). This potential variation in the CSA, ca. $\pm\,12\%$, leads to a variation of only $\pm\,3.0\%$ in RT2. In conclusion, RT2 is roughly independent of τ_m^0, CSA, S^2 and τ_i, and only its dependence on Rex remains. Equation 11 is not exact and thus variations in RT2 due to τ_m^0, CSA, S^2 and τ_i may be larger than expected. We have therefore calculated RT2 using the full equations, for different τ_m^0 values and internal motion parameters (Figures 4A–D). If Rex $= 0$, the RT2 values indeed fall within a relatively narrow range (when $\tau_m^0 \geqslant 6$ ns, $1.07 <$ RT2 < 1.14). The effect of a potential variation of the CSA on RT2 is weak (Figure 4B; CSA: $\pm\,12\%$ leads to RT2: $\pm\,3\%$). Thus, the conclusion remains the same; RT2 only weakly depends on τ_m^0, internal motion parameters, and CSA.

Consequently, to predict the exchange free RT2 ($RT2^{\rm exfree}$) within a few % only rough estimates (within ca. 20%) of S^2 and τ_m^0 are needed and a distinction between ps-im and ns-im. The experimental error in RT2 is ca. 6%, which overwhelms this systematic error. Thus, Rex can accurately be determined from the difference between experimental RT2 and $RT2^{\rm exfree}$. This does not require prior knowledge of the complexity and extent of the internal motion and exact knowledge of τ_m^0. The level of accuracy of determining exchange broadening is better than 2 s^{-1} (Figures 4(A–D).

Finally, it is to be noted that we assume a B_0-squared dependence for Rex, i.e., the broadening is due to exchange in the fast limit. This is an approach that is most commonly used (see, e.g., review of Korhznev et al., 2001). In principle, one may distinguish between fast and slow exchange by the number of resonance lines present in the NMR spectra per exchanged spin. However, observation of single resonances in NMR spectra does not necessarily mean that the exchange is fast. This problem has been considered and recipes proposed on how to estimate the timescale of an exchange process based on CPMG data (see for example Korhznev et al., 2001, and references therein).

IVb. Estimation of τ_m^0 independent of the time scale of internal motion

Figure 5A shows a plot of τ_m^{ap} vs. NOE. The S^2-contours follow curves quite similar to those of $R_{\tau{\rm mapp}}n$ or RT1n. In the ps-im region τ_m^{ap} is strictly linearly dependent on NOE and the slope is only weakly dependent on τ_m^{ap} as long as $\tau_m^0 > 5$ ns. Hence, a correction of the linear dependence on the NOE can be made and a ps corrected τ_m^{ap} (τ_m^{ap-ps}) can be ob-

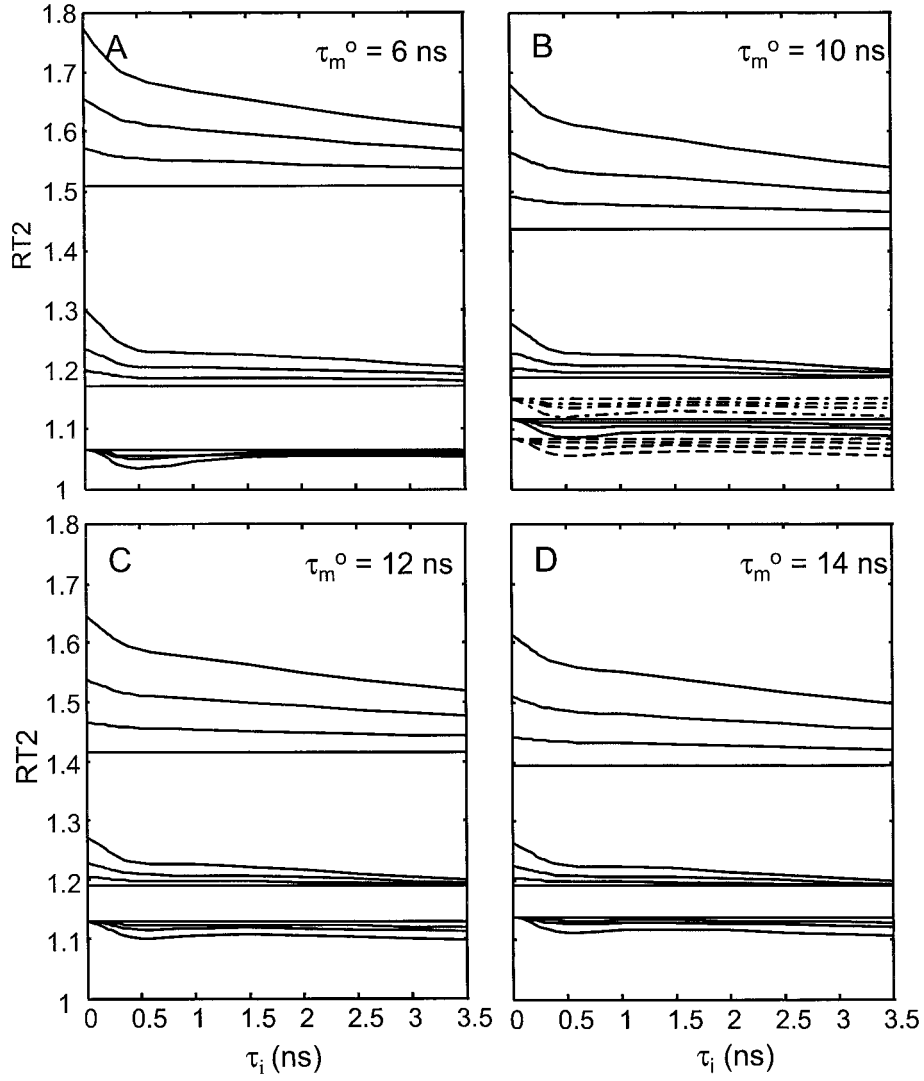

Figure 4. RT2 ($= R_2^{600}/R_2^{400}$) vs. τ_i at $\tau_m^0 = 6$ ns (A), 10 ns (B), 12 ns (C) and 14 ns (D). The full equations are used, i.e., Equation 2 for R_2 and Equation 5 for the spectral density function. One contribution model is assumed for internal motion. In each panel three sets of four RT2 values are shown as drawn lines, corresponding to Rex $= 0$ s^{-1} (lowest), 2 s^{-1} (middle), and 12 s^{-1} (high). Within each set, four RT2 values corresponding to S^2 values of 1, 0.8, 0.6 and 0.4 are shown. CSA is -170 ppm in all cases, except in panel B, where also are shown the RT2 values for CSA $= -150$ ppm (broken lines) and -190 ppm (broken dotted lines) with Rex $= 0$.

tained (Figure 5B, Table 2). In the ns-im region τ_m^{ap-ps} can lie substantially below the true rotation correlation time, τ_m^0. Note also that requiring NOE > 0.6, as done in the usual analysis to remove residues with slow internal motion, does not guarantee a correct estimation of τ_m^0.

Figure 5C shows a three-dimensional plot of τ_m^{ap-ps} vs. $R_{\tau\text{mapp}}n$ and NOE. To first order, the S^2-contours lie in a plane parallel to the NOE axis. Thus, τ_m^{ap-ps} is independent of the NOE and linearly dependent on $R_{\tau\text{mapp}}n$, i.e., $\tau_m^{ap-ps} \approx \tau_m^0 - \alpha(R_{\tau mapp}n - 1)$, where α is the slope of the plane. Consequently, this equation can be used to correct τ_m^{ap-ps} for internal motions slower than 200 ps (ns-im). The correction only depends on $R_{\tau\text{mapp}}n$. This yields $\tau_m^{ap-ps-ns}$, which in fact corresponds to τ_m^0. In other words, τ_m^0 can be estimated independent of the time scale or complexity of internal motion.

In practice this correction turns out to be rather rough (± 1 ns) and dependencies are not exactly linear. We have determined optimal recursive correction equations (Table 2). The iterative procedure developed

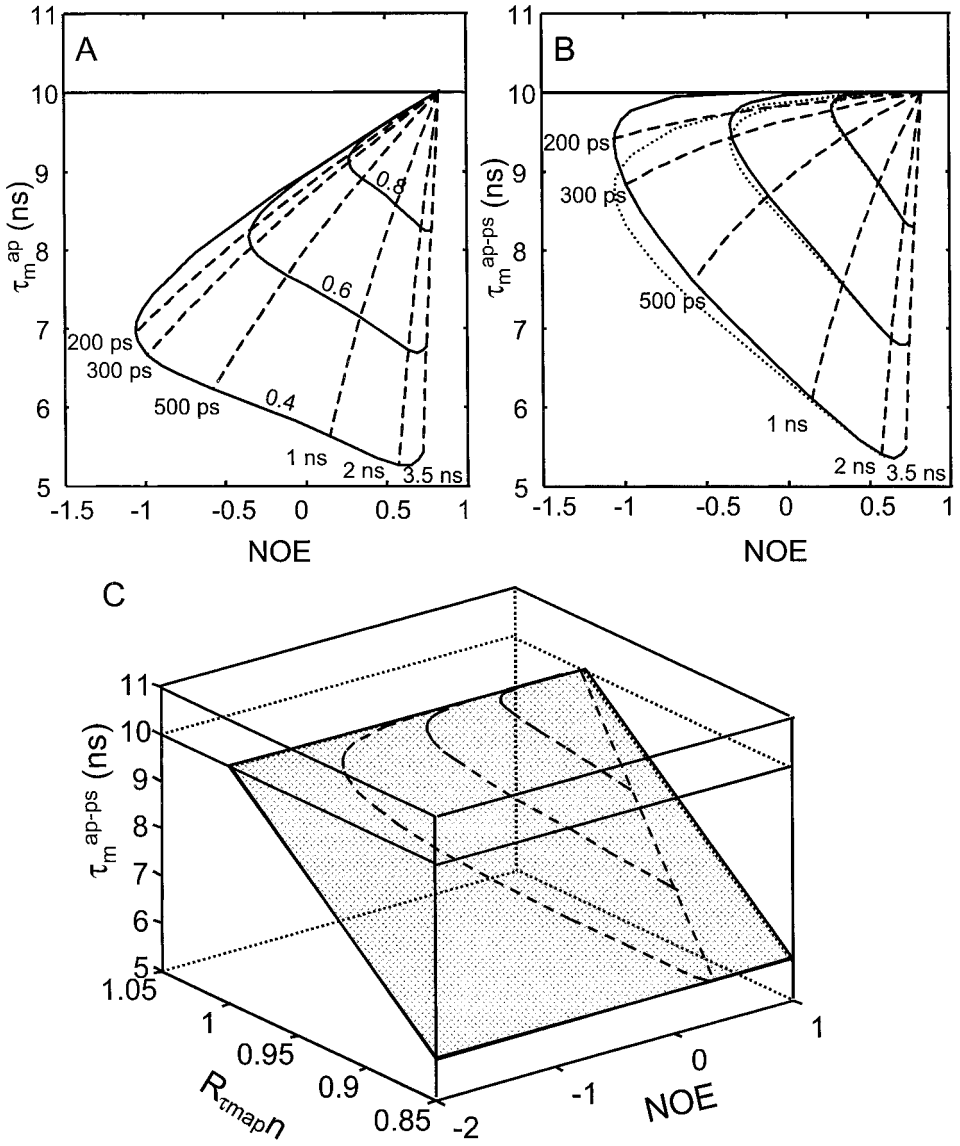

Figure 5. The τ_m^{ap} (A) and τ_m^{ap-ps} (B) as a function of the NOE, and τ_m^{ap-ps} as a function of both $R_{\tau mapp}n$ and NOE (C). The parameter settings are as in Figure 2, $\tau_m^0 = 10$ ns, and Rex $= 0$.

to derive τ_m^0 is written in MATLAB and proceeds as described in the flowchart (Figure 1).

IVc. Numerical tests and error considerations

To test how well the iterative determination of τ_m^0 and Rex works under different situations, we have carried out extensive tests on a variety of simulated data. The most important results are illustrated in Figure 6. The correction on τ_m^{ap} is excellent, i.e., $\tau_m^{ap-ps-ns} \approx \tau_m^0 \pm 0.5$ ns over a wide range of $S^2 (1 < S^2 <\approx 0.4)$ and τ_i values ($0 < \tau_i <\approx 2.0$ ns) (Figures 6A–D). For larger τ_i values up to 3 ns, the estimation remains correct but requires S^2 values progressively closer to 1. The range of τ_m^0 values for which the approach works well is from ca. 6 ns up to 16 ns (or higher). Whether the internal motion consists of one- or two contributions does not affect the correct determination of τ_m^0 (compare Figures 6A and 6B). A wrong value of CSA by up to ca. $\pm 30\%$ hardly affects the correct estimation of τ_m^0 (Figure 6C), although for smaller S^2 values, $\tau_m^{ap-ps-ns}$ become somewhat overcorrected. As long as $S^2 \geqslant 0.6$, the over-correction is insub-

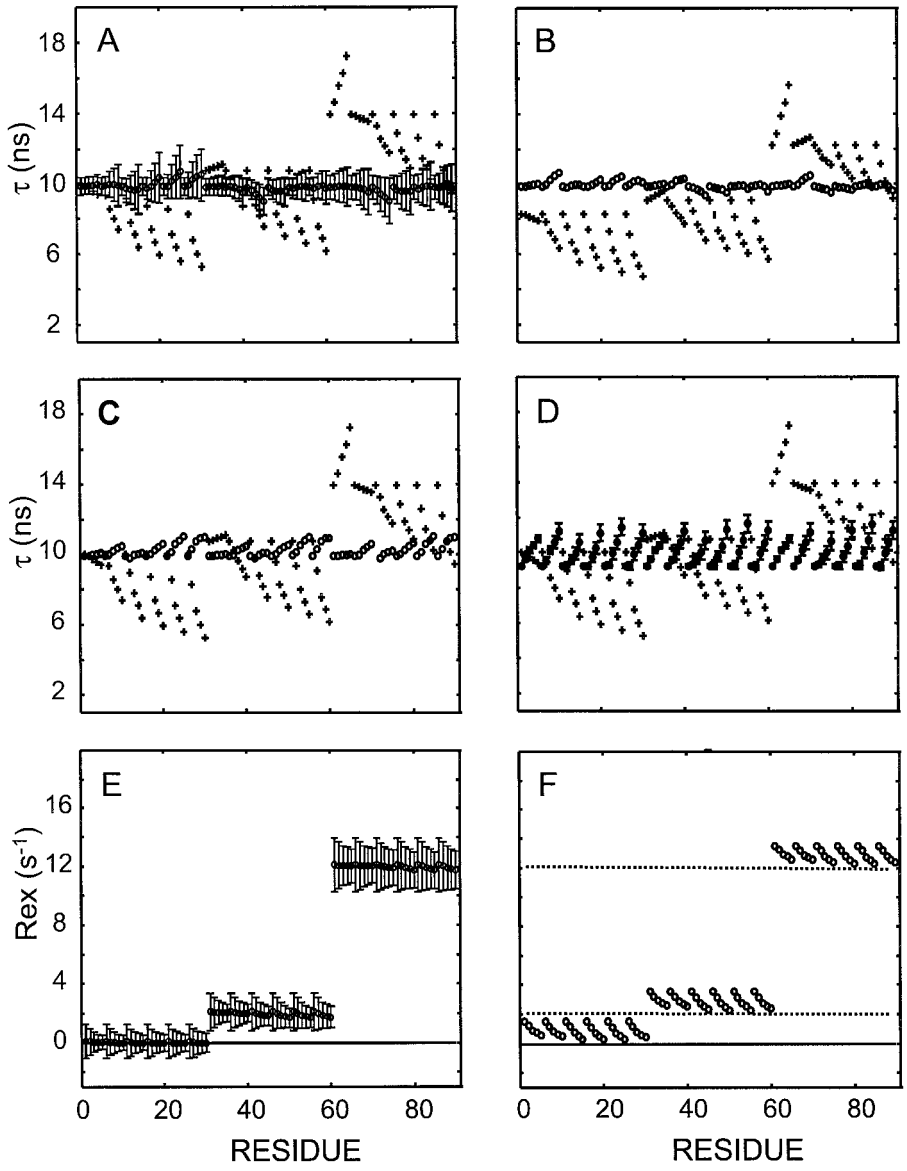

Figure 6. Test of correction of τ_m^{ap} for ps- and ns-im and determination of Rex from RT2. Panels A to C show $\tau_m^{ap-ps-ns}$ (o) and τ_m^{ap} (+), and panels E and F show Rex, as a function of residue number, where the residue number represents different conditions. S^2 runs from 1.0 in steps of 0.2 to 0.4 for residues 1 to 5; the same set of S^2 values applies for residues 6 to 10 and so on until the last group of residues, 86 to 90. The τ_i values are 0.02 ns (residues 1 to 5), 0.15 ns (residues 6 to 10), 0.4 ns (residues 11 to 15), 0.7 ns (residues 16 to 20), 1.0 ns (residues 21 to 25), 1.7 ns (residues 26 to 30). Rex = 0 s^{-1} for residues 1 to 30, Rex = 2 s^{-1} for residues 31 to 60, and Rex = 12 s^{-1} for residue 61 to 90. To calculate the synthetic relaxation data (R_1, R_2, and NOE) at 600 MHz and 400 MHz, a one-contribution internal motion model was assumed, except in B (see below). The full equations were used (Equations 1–3 and 5). In all cases the CSA is assumed to be −170 ppm, except in C and F (see below). The τ_m^0 is 10 ns. (A) $\tau_m^{ap-ps-ns}$ with error bars derived from the experimental errors σ on R_1 of 1% and R_2 of 2%; the error bars on $\tau_m^{ap-ps-ns}$ (and Rex see E below) are obtained by rerunning protocol with different combinations of $R_1 \pm \sigma$ and $R_2 \pm \sigma$. (B) $\tau_m^{ap-ps-ns}$ derived when internal is described by a two contribution model; the additional 'slow' internal motion has $\tau_{is} = 2.5$ ns and $S^2 = 0.8$. The error bars have been omitted for clarity. (C) $\tau_m^{ap-ps-ns}$ is calculated assuming CSA = −170 ppm, while the test data set was generated with CSA = −200 ppm. (D) $\tau_m^{ap-ps-ns}$ obtained using an average R_2. The reduced experimental error, which is now only based on R_1, can be seen on the smaller error bars. Note the closeness to the actual value of 10 ns. (E) Rex estimated with error bars resulting from experimental errors of 1% on R_1 and 2% on R_2, and calculated as described under A. (F) Rex estimated using the wrong CSA value; conditions identical to those described under C.

stantial, $\tau_m^{ap-ps-ns} \leqslant \tau_m^0 + 0.5$ ns. An important observation with regard to the data in Figures 6A–C, is also that the presence or absence of a Rex term has no effect on the value of $\tau_m^{ap-ps-ns}$.

As shown in Figure 6E, Rex itself is determined very accurately, i.e., within ca. ± 0.2 s^{-1} of its real value when the correct CSA is used. The effect of a potential variation of the CSA on the Rex estimate is illustrated in Figure 6F, where Rex is estimated using a CSA value of -180 ppm, while the actual value is -200 ppm. This leads to an overestimation of Rex by ca. 0.9 s^{-1} on average (using a value of -170 ppm for the CSA instead of -180 ppm leads to an overestimation of Rex by ca. 1.2 s^{-1} on average). The variation in the CSA is thus absorbed into Rex.

The errors discussed so far are systematic and due to the approximations during the derivation of the analytical correction factors for $\tau_m^{ap-ps-ns}$ and in the estimation of Rex. Another matter is how the experimental error on the relaxation data affects the estimate of τ_m^0 and Rex (Figures 6A and 6D). To estimate the error on $\tau_m^{ap-ps-ns}$ and Rex the following approach was implemented. The $\tau_m^{ap-ps-ns}$ and Rex calculation protocol (Figure 1, Section 2) is executed for three combinations of $R_1 \pm \sigma$ and $R_2 \pm \sigma$ (and optionally the NOE) at the two fields, chosen so that the maximum, middle and minimum values of $\tau_m^{ap-ps-ns}$ and Rex, consistent with the error R_1 and R_2, are obtained. The final $\tau_m^{ap-ps-ns}$ and Rex are the average of these values and their rmsd is taken as the error. This approach can simply be turned into a true Monte Carlo estimation by execution of the protocol for a large number R_1 and R_2 values within their error range.

The experimental error on $\tau_m^{ap-ps-ns}$ turns out to be about $\pm$ 0.8 ns for good quality data (1% in T_1 and 2% in T_2). The error on Rex is directly related to the error in T_2 as it is derived from RT2 and is about ± 0.7 s^{-1}, ± 0.9 s^{-1} and ± 1.4 s^{-1} for Rex = 0, 2 and 12 s^{-1}, respectively. In conclusion, the systematic errors are clearly smaller than these experimental errors.

A few final comments should be made. The experimental error on R_2 is in practice larger than in R_1, often by as much as a factor of two (see e.g., Farrow et al., 1994; Fischer et al., 1998; Loria et al., 1999; Korzhnev et al., 2001). It may therefore be more advantageous to use the average R_2, $\langle R_2 \rangle$, instead of R_2 itself, for calculating $R_{\tau\text{mapp}} n^{\text{excor}}$, which is needed to calculate $\tau_m^{ap-ps-ns}$. Using an average R_2 effectively removes the experimental error on R2 as a source for experimental error in $R_{\tau\text{mapp}} n$ and thus in $\tau_m^{ap-ps-ns}$. Due to the fact that R_2 does not depend much on the time scale for internal motion, such an approach may not be detrimental to the accuracy. Both aspects are borne out by the numerical tests (Figure 6D). For the whole range of the test data ($0.4 \leqslant S^2 \leqslant 1$, 0.02 ns $\leqslant \tau_i \leqslant$ 1.7 ns), $\tau_m^{ap-ps-ns}$ ranges between 9.2 ns and 11.8 ns (rmsd of 0.8 ns). When $S^2 \geqslant 0.6$ and τ_i up to 1.7 ns, $\tau_m^{ap-ps-ns}$ remains within ca. ± 0.5 ns from the actual value of 10 ns (range between 9.5 ns and 10.7 ns, rmsd 0.27 ns). The experimental error on $\tau_m^{ap-ps-ns}$, which is based on the 1% error in R_1, is indeed considerably reduced and even smaller or equal to the systematic error (± 0.3 ns, error bars in Figure 6D) when $\langle R_2 \rangle$ is used.

It is of interest to consider how PINATA performs with respect to the determination of τ_m^0 in non-ideal situations. The $\tau_m^{ap-ps-ns}$ estimate progressively deteriorates further away from the ideal situation, $\tau_m^0 <$ 5 ns and $S^2 < 0.4$. However, even when $\tau_m^0 < 5$ ns, PINATA still performs essentially correct as long as τ_i is well within the ps-im regime and $S^2 > 0.4$ (i.e. NOE >0.2), because ps-im and ns-im corrections are insignificant. For larger τ_i values the parameters for ps-im and ns-im correction need to be adjusted, which can be done via test calculations. When $\tau_m^0 >$ 5 ns and $\tau_i > 3$ ns the ns-im correction of τ_m^{ap} becomes incorrect for smaller S^2 (e.g., < 0.7) and PINATA takes τ_m^{ap} as the best estimate of τ_m^0.

The effect on the $\tau_m^{ap-ps-ns}$ and Rex values of an error in the NOE can in the protocol simply be taken into account by including the potential variation in the NOE in the re-execution of the protocol (see above and Figure 1, step 2). However, the error in the NOE is negligible as follows from the following considerations. Let us assume a potential error in the NOE of $\pm$ 0.1. This leads to an error in $\tau_m^{ap-ps-ns}$ of ca 1% ($= 0.117/\tau_m^{ap-ps-ns}$; see Table 2) due to the ps-im correction; the ns-im correction depends only to second order on the NOE (Table 2) and does not contribute via the NOE to the error.

The $\tau_m^{ap-ps-ns}$ values represent τ_m^0. In the case of anisotropic tumbling, the residue-specific τ_m^0 contains global structural information (Equation 9). In the usual methodology to analyze ^{15}N relaxation data, residues with NOE > 0.6 and residues with T_1 and T_2 values close to their average are selected to extract this information (Tjandra et al., 1996). Subsequently, the diffusion tensor is estimated from the τ_m^{ap} of the selected residues (Tjandra et al., 1996). The reduced

number of residues may prevent a correct estimate of the diffusion tensor (Renner and Holak, 2000). With the protocol presented here no residue selection needs to be done. Consequently, a more reliable estimate of the diffusion tensor is obtained. In addition, more orientations N-H relaxation vectors are available for more 'global' structural information. Error considerations show that even for modest degrees of anisotropy the helix orientations can be well determined from good quality relaxation data.

V. Final order parameters and time scales via grid search fitting (Step 4, Figure 1)

After Rex and τ_m^0 have been determined independent of the model for internal motion, and the internal motional model has been qualitatively assessed via analysis of $R_{\tau\mathrm{map}}n$ and RT1n vs. NOE graphs, the final order parameters and time constants for internal motion are determined via grid search fitting. The following target function has been used:

$$\chi_\upsilon^2 = \frac{1}{\nu}\left\{\frac{\left(R_2^{hf\_\exp} - R_2^{hf\_\mathrm{calc}}\right)^2}{\sigma^2_{R_2^{\mathrm{hf}}}} + \frac{\left(R_1^{hf\_\exp} - R_1^{hf\_\mathrm{calc}}\right)^2}{\sigma^2_{R_1^{\mathrm{hf}}}} + \frac{\left(RT1^{\exp} - RT1^{\mathrm{calc}}\right)^2}{\sigma^2_{\mathrm{RT1}}} + \frac{\left(\mathrm{NOE}^{hf\_\exp} - \mathrm{NOE}^{hf\_\mathrm{calc}}\right)^2}{\sigma^2_{\mathrm{NOE}^{\mathrm{hf}}}}\right\}. \quad (12)$$

Here, $\nu = N - p$ is the number of degrees of freedom left after fitting N data points with the fitting function that has p adjustable parameters (Bevington, 1969). The σ's are the experimental error estimates. The subscript '*hf*' stands for high-field and identifies that R_2^{hf}, R_1^{hf}, etc., are measured or calculated at the highest field. We use the above target function based on the idea that Rex is already determined from RT2 ratios, so that only R_2^{hf} is still a free parameter (exchange corrected R_2 at highest field), while R_1^{lf} ($lf =$ low field) is incorporated in to RT1. The fitting routines, which are written in MATLAB, employ grid search to ensure that no minima are missed. No restrictions are put on the relaxation equations and in principle every parameter can be optimized. First a rough grid is used to scan the target error function using parameter estimates from the earlier steps as starting values. Then a finer grid is applied. Internal motion models with either one (M1; τ_i and S^2 fitting parameters) or two contributions (M2; τ_{if}, S_f^2, τ_{is} and S_s^2 fitting parameters) are tested. The S_s^2 can either be kept fixed at a uniform value or optimized together with τ_{if} and S_f^2. To keep the number of adjustable parameters as small as possible, τ_{is} is kept fixed at a uniform value during a fitting run. However, it can be optimized via different fitting runs. The optimization can be carried out with either a residue-specific τ_m^0 or an average τ_m^0. The quality of the fit of a given model is statistically assessed via the standard χ_υ^2 statistics (Bevington, 1969).

Alternatively, one can use the Modelfree (Mandel et al., 1995) or DASHA (Orekhov et al., 1996) programs to optimize the parameters using the model selection based on $R_{\tau\mathrm{map}}n^{\mathrm{excor}}$ or RT1n.

VI. Demonstration of protocol on experimental data

The protocol has also been applied to the published experimental ^{15}N-relaxation data of the M13 coat protein (gVIIIp) complexed with SDS micelles, which has been measured at 500, 600 and 750 MHz (Papavoine et al., 1997, 1998). The 45 residue long gVIIIp contains two helices. One helix is inserted into the SDS micelle, whereas the other is bound to the micelle surface. The two helices show distinctly different relaxation behavior and thus are expected to have different degrees of ps-im and ns-im for the residues in the two helices. Therefore, this set of relaxation data promises to be suitable for demonstrating our protocol. First, for determining the presence or absence of ps-im and/or ns-im via our protocol, which functions even if all residues in the protein would be involved in ns-im. Secondly, for determining the real residue-specific τ_m^0 independent of the time scale of internal motion. The latter provide global structural information in case of anisotropic tumbling. The NMR structure of gVIIIp was derived from short-range classical NMR data (NOEs and J-couplings). Generally, it is difficult to derive global structural features from classical NMR data. However, in this case the classical NMR data sufficed to reasonably well define the relative orientation of the two helices, i.e., to define global structural features. In other words, the global structural information we derive via our protocol from the relaxation data becomes testable. The analysis is based on the R_1, R_2 and NOE data collected at 500 and 750 MHz.

VIa. Analysis of internal motion via RT1n and $R_{\tau mapp}n$ versus NOE *plots*

The exchange-corrected ($R_{\tau mapp}n^{excor}$) and the RT1n values vs. NOE are displayed in Figures 7A and 7B, respectively. The $R_{\tau mapp}n$ values are the ratio of the residue-specific apparent overall tumbling times derived from R_1 over R_2 ratio at the two magnetic fields, while RT1n are the R_1 ratio's at the two magnetic fields (exact definition in Section III). $R_{\tau mapp}n$, RT1n, and NOE follow directly from the measured experimental data and do not depend on a motional model. Both plots in Figure 7 show essentially the same features except that the RT1n vary over a larger range (between ca. 1 to 0.8). In addition, RT1n values are not affected by possible conformational exchange and have smaller errors than the $R_{\tau mapp}n^{excor}$ values (ca. 1 to 0.89). As can be seen some residues have RT1n values around 1, while others have RT1n values considerably smaller than 1. The residues with RT1n (or $R_{\tau mapp}n^{excor}$) around 1 are only affected by ps-im, while those with RT1n (or $R_{\tau mapp}n^{excor}$) < 0.97 experience ns-im or a mixture of ns- and ps-im (see Section III). Note that this conclusion holds true, irrespective of the fact whether the overall tumbling is isotropic or anisotropic, and independent of the value of the NOE, independent of the value of the exact value CSA and the exact value of the overall tumbling time(s) (see Section III).

To investigate whether the $R_{\tau mapp}n$ and RT1n values correlate with expected rigidity, we have in Figures 7C and 7D separated the residues according to their structural environment. The residues in the helix which is inserted into the micelle, generally have RT1n $= 1 \pm 0.05$ (Figure 7D). Thus, they experience only ps-im and a description with one contribution of internal motion is exact. This is expected for a well-formed helix in a rigid core. In contrast, the residues in the helix on the SDS surface cluster around $\langle$RT1n$\rangle \approx$ 0.87, showing the presence of ns-im (Section IIIc). A description with either one or two contributions of internal motion is then required. If a one-contribution model holds, the time scale and order parameter of the internal motion can be directly read off from the RT1n vs NOE graph (Figure 7D). A $\langle$RT1n$\rangle$ of ca. 0.87 and a $\langle$NOE$\rangle$ of ca 0.58 correspond to τ_i of ca 1 ns with S^2 of ca. 0.67. The error margins on τ_i and S^2 follow directly from Figure 7D and the experimental error in RT1n (ca. 2%) and NOE (ca 0.05). In the case of a two-contribution model the data can be interpreted as different combinations of ps- and ns-im (see section IIIc). For example, the internal motion can consist of a ns-im with $\tau_{is} = 1.5$ ns, so that its contribution (S_s^2) is ca. 0.75, together with a ps-im with varying but small values for its S_f^2 (ca. 0.8) (Figure 7D, dotted lines). Whether a one- or a two-contribution model describes the data best, can be established from a simple comparison of the experimental and predicted R_1-values (see also Section IIIc). The surface helix has experimental R_1-values of 1.25 ± 0.02 s^{-1} (at 750 MHz, residue 10 to 16). For a one-contribution model with $\tau_{is} = 1.0$ ns and $S^2 = 0.7$ and with $\tau_m^0 = 9.5$ ns (see below), the predicted R_1 is ca. 1.41 s^{-1}. This value is too high. Therefore a two-contribution model needs to be considered. For $\tau_{is} = 1.5$ ns and $S_s^2 = 0.75$ and $\tau_{if} = 0.02$ ns and $S_f^2 = 0.85$ and $\tau_m^0 = 9.5$ ns, which fits the RT1n and NOE data in Figure 7D, the predicted R_1 is 1.26 s^{-1}. This R1 value is close to the experimental R_1 of 1.25 s^{-1}. Thus, a two-contribution model indeed is needed to explain the relaxation data for the surface helix. The surface helix residues undergo both ps-im as well as ns-im. The ps-im is of an amplitude and time scale such as usually found in well-defined helices. The ns-im can probably be ascribed to domain motion of the complete helix over the micelle surface.

In conclusion, the internal motion of the two helices as derived from the relaxation data indeed correlates nicely with the structural environment.

VIb. The real residue-specific overall tumbling times τ_m^0 and global structural information

Figure 8A shows the residue-specific $\tau_m^{ap-ps-ns}$ values for the M13-SDS complex obtained via our analysis protocol (Section IV). Note that the $\tau_m^{ap-ps-ns}$ values are corrected for internal motion on either ps- or ns-time scale. They represent the real residue-specific overall tumbling times τ_m^0. If anisotropic tumbling is present, these values depend on the N-H vector orientation and thus can provide global structure information (Sections IVb and IVc).

The $\tau_m^{ap-ps-ns}$ values seen in Figure 8A are not equal but show variation along the amino acid sequence, e.g., for the surface helix they are different from those for the helix inserted into the micelle. This shows that the complex tumbles anisotropically. From the distribution of the $\tau_m^{ap-ps-ns}$ values we estimate an anisotropy ($= \tau_l^0/\tau_s^0$) of 1.5 to 2.0. The former value is estimated from $\left\langle \tau_m^{ap-ps-ns} \right\rangle$ of the inserted helix (12.3 ns) and of the surface helix (9.8 ns) via $(\tau_l^0/\tau_s^0) \approx 1 + (2 * ((12.3/9.8) - 1))$ (Equation 9) and the latter from the maximum and minimum values of $\tau_m^{ap-ps-ns}$ of 12.8 ns and 8.2 ns, respectively.

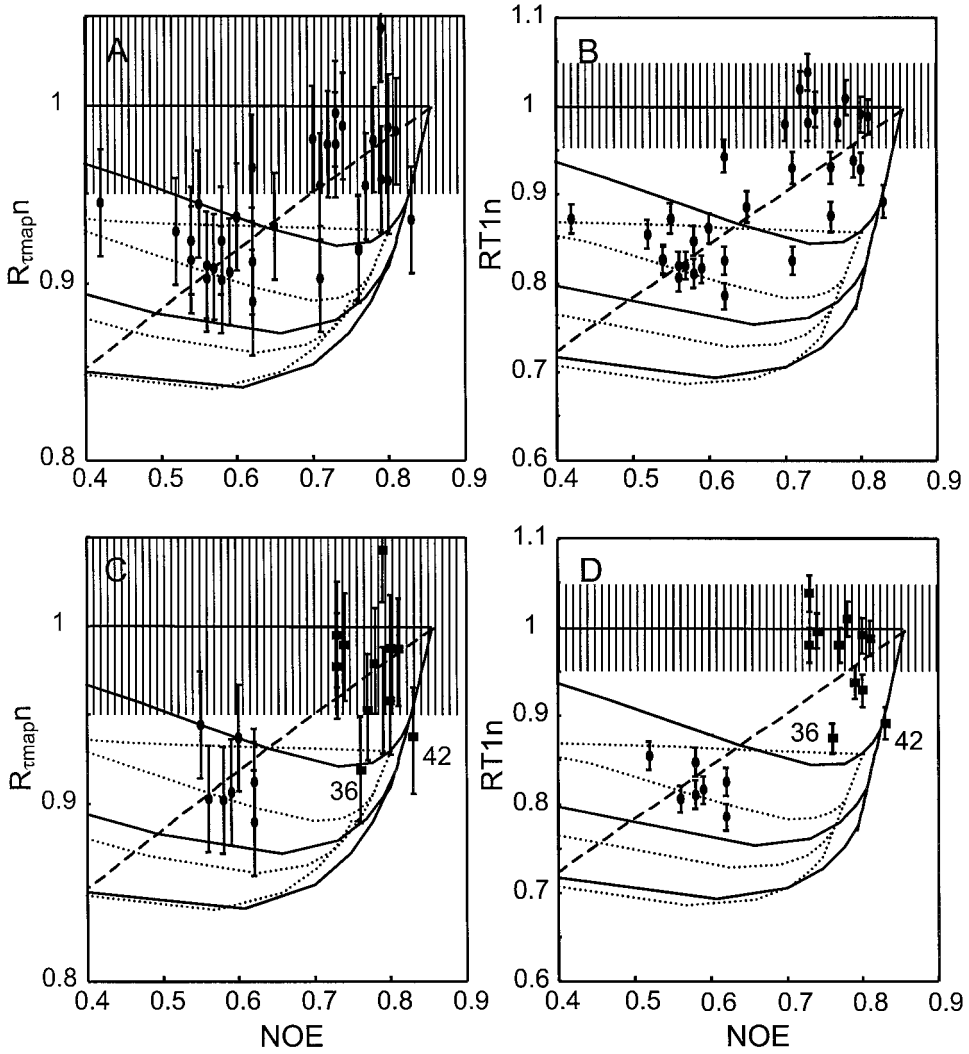

Figure 7. $R_{\tau mapn}$ (A, C) and RT1n (B, D) vs. NOE plots of ^{15}N-relaxation data from the M13 gVIII coat protein (gVIIIp) in complex with SDS micelles at 750 and 500 MHz (Papavoine et al., 1997, 1998). In panels A and B all residues are shown. In panels C and D only the residues are shown, which are part of the helix on the surface of the SDS micelle (10–16; circles) or part of the helix that is inserted in to the micelle (30–43; squares). Two residues are labeled with its corresponding residue number. The normalization constant $RT1^0$ used in the experimental RT1n is calculated assuming CSA $= -170$ ppm and t_m^0 equal to $\left\langle \tau_m^{ap-ps-ns} \right\rangle$ (see text). The error on the data points is based on 1% error in R_1 and 2% error in R_2. The areas with vertical drawn lines indicate the regime where only ps-im is present. The theoretical S^2-contours at S^2 1, 0.8, 0.6 and 0.4, are calculated using the full equations (Equations 1–3 and 5) with $t_m^0 = \left\langle \tau_m^{ap-ps-ns} \right\rangle$ and CSA $= -170$ ppm in two ways. (1) A one-contribution model for internal motion is assumed (solid contour lines) with t_i running from 20 ps to 6 ns. The dashed line is the t_i-contour at 1 ns. (2) An additional internal motion is assumed to be present with $S_s^2 = 0.8$ and $t_{is} = 2$ ns (dotted contour lines).

Within each helix, the $\tau_m^{ap-ps-ns}$ values in Figure 8A are effectively the same. This is expected, because within helices, the N-H vectors are oriented almost parallel to the helix axis and thus have the same orientation throughout the whole helix. Comparing the two helices, the $\left\langle \tau_m^{ap-ps-ns} \right\rangle$ of the helix inserted into the micelle is ca 12.3 ns, while for the surface helix ca 9.8 ns. This difference shows that within the complex the two helices have different orientations. The $\left\langle \tau_m^{ap-ps-ns} \right\rangle$ of the helix inserted into the micelle is close to the maximum. Therefore, this helix must be oriented nearly parallel to the long axis of the diffusion tensor and thus to the long axis of the complex. In contrast, the $\left\langle \tau_m^{ap-ps-ns} \right\rangle$ of the residues in the surface

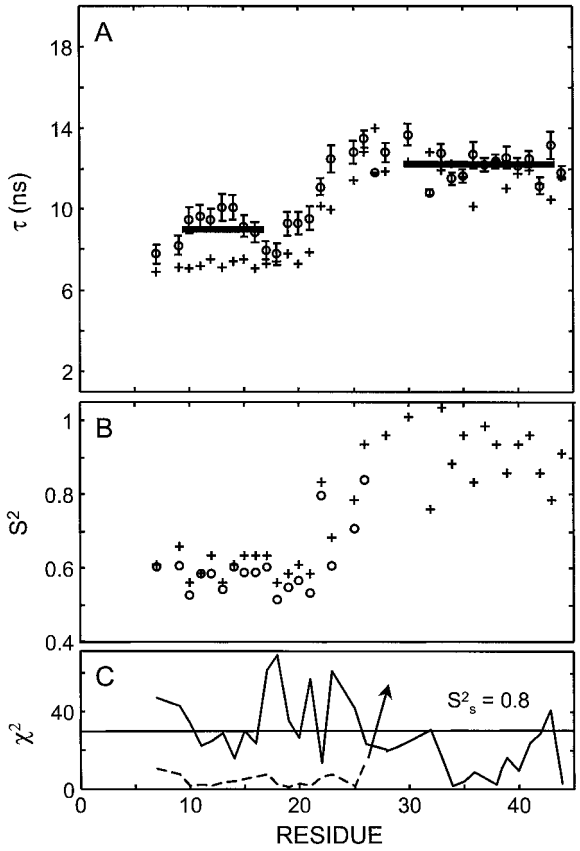

Figure 8. Overview of fitting results for M13 coat protein gVIIIp in complex with the SDS micelles. (A) τ_m^{ap} (+) and $\tau_m^{ap-ps-ns}$ (o with error bars) versus residue number. The errors on $\tau_m^{ap-ps-ns}$ are obtained from the errors in R_1 (1%) and R_2 (2%) at the two fields by recalculation of $\tau_m^{ap-ps-ns}$ for combinations of $R_1^{hf/lf}$, $R_1^{hf/lf} \pm \sigma$, $R_2^{hf/lf}$, and $R_2^{hf/lf} \pm \sigma$, so that its upper and lower bound values are found (hf and lf indicate high and low field, respectively, and σ the experimental error). (B) Final S^2 of the one-contribution (+) and two-contribution models (o). The S_s^2 of the two-contribution model is 0.8 and t_{is} is 1.5 ns. (C) The χ^2 for the one- (dashed line) and two-contribution models (drawn line) with $S_s^2 = 0.8$ and $\tau_{is} = 1.5$ ns.

helix is close to the smallest value. This helix must therefore be close to but not exactly perpendicular to the long axis of the complex (ca. 60°, Equation 9).

In conclusion, real residue-specific overall tumbling times were determined, i.e. they are not affected by ps- and/or ns-im, and global structure information has been derived from their residue-specific variation. Note that Papavoine et al. employed the usual analysis to derive overall tumbling times from R_1 over R_2 ratio's, i.e., no correction for ns-im could be made. They therefore had to assume that the residues in the surface helix experience the same tumbling time as those in the inserted helix (ca. 11.3 ns). Consequently, they could not derive global structural information from the relaxation data. It is interesting to compare our global structural information with the NMR structure (Papavoine et al., 1998). This structure was determined from classical short-range NMR constraints, such as NOEs and *J*-couplings. Generally, it is difficult to derive global structural features from such data. However, in this case the data sufficed to reasonably well define the relative orientation of the two helices, i.e., to define global structural features. The set of NMR structures shows that the two helices are approximately perpendicular. Thus, our analysis of the relaxation data confirms this 'global' structural feature of the gVIIIp in the complex.

Recently we have derived the global structure of the apoCII-SDS complex (Zdunek et al., 2003). ApoCII consists of three helices attached to the surface of the SDS micelle. The classical NMR data did not define the relative orientation of these helices. Our analysis of relaxation data measured at two fields showed that all residues were affected by ns-im. The real residue-specific overall tumbling times provided 'global' structural information, which together with the other restraints was sufficient to define the 'global' structure of the complete apoCII-SDS complex.

VIc. Final internal motion parameters by fitting via grid search

Given the qualitative model assessment, including detection of the presence of ns-im (Section VIa) and the internal motion corrected residue-specific overall tumbling times (Section VIb), the final internal motion parameters are obtained by fitting via the grid search method (see Section V). The residues of the helix inserted into the SDS micelle are indeed fitted best with a one-contribution model with ps-im and S^2 of around 0.9 (Figures 8B and 8C). In contrast, for the surface helix the χ^2-residuals are only acceptable when a two-contribution model with a mixture of ps- and ns-im is invoked (Figure 8C). The $S^2 (= S_s^2 S_f^2)$ values of the residues in the surface helix are ca. 0.6 (Figure 8B).

It is interesting to compare these final internal motion parameters with those of Papavoine et al. (1997). They derived the overall tumbling times via the usual method, i.e., from R_1 over R_2 ratio's but without correction for ps- and/or ns-im, and had to assume that the residues in the surface helix experienced the same tumbling time as those in the inserted helix (ca. 11.3 ns). As described in Section VIb, we find that the real overall tumbling time of residues in the surface

helix is ca. 9.8 ns, after correction for ns-im, while those in the inserted helix have on average 12.3 ns (Figure 8A). Smaller tumbling times tend to increase the S^2. Thus, our average value of S^2 for the surface helix of 0.6 is indeed up from the value of ca. 0.5 found by Papavoine et al. (1997). The reverse but to lesser extent holds (11.3 ns versus 12.3 ns) for the residues of the helix inserted into the micelle.

Their results are similar to ours because they measured the ^{15}N relaxation at multiple fields, and most importantly ca half of the residues in the gVIIIp undergo only pure ps-im of small amplitude. Consequently, for these residues the overall tumbling time is correctly estimated via the usual method. The underestimation of the overall tumbling time due to the residues in the surface helix, which experience ns-im, is thus compensated by the correct value of the residues in the helix inserted into the micelle. Is there a worse case scenario? They occur for proteins with domain motions or for partially folded proteins. An example is the apoCII-SDS complex (Zdunek et al., 2003), where all apoCII residues experience a significant degree of ns-im. Applying the usual method for estimating the overall tumbling time (no ps-im and ns-im correction) would here underestimate the overall tumbling time considerably (ca. 12 ns versus 9 ns for the first two helices). This underestimation leads to an overestimation of the real S^2 (which is 0.66) by a factor of 1.33 (estimated from the equation for R_2). Furthermore, Korhznev et al. (1997) have shown that analyzing relaxation data at only one field with an underestimated tumbling time obscures the presence of ns-im. That is, a perfectly correct fit is obtained for a one-contribution ps-im model even when both ps- and ns-im are present. Fitting the relaxation at two fields simultaneously, but with an underestimated overall tumbling time, still underestimates the contribution of ns-im and thus obscures the presence of ns-im at least partly. In our protocol, we focus on the field dependence of the relaxation data (via the RT1n versus NOE plots and in the fitting target function, Equation 12) and employ the correct overall tumbling time, which circumvent these potential pitfalls.

Concluding remarks

We successfully tested and demonstrated a new protocol (PINATA) for analyzing within the context of the Lipari–Szabo formalism ^{15}N-relaxation data measured at two fields at least. The two main new features of the protocol are the following. With this method the presence or absence of ns-im can unambiguously be established irrespective whether the overall motion is isotropic or anisotropic, and independent of the value of the NOE, even when all residues are affected by ns-im. In addition, the real residue-specific overall tumbling time, τ_m^0, is determined independent of the time scale and amplitude of internal motion. These results are obtained by focusing on the field dependence of the relaxation data. Thus, in contrast to the usual analysis, no priori assumption about time scales or amplitude of internal motions needs to be made. The PINATA protocol takes ca 1 to 1.5 h to analyze the relaxation data of a 150-residue protein and has a graphical interface to quickly identify the internal motion model. The decoupling assumption in the Lipari–Szabo description of internal motion which is employed in PINATA is dropped in the theory developed by Meirovitch and coworkers (Tugarinov et al., 2001) for isotropically tumbling molecules. This SRLS theory still converges to the Lipari–Szabo formalism in the fast motional limit and of course in the rigid limit. Because PINATA focuses on detection of deviations from ps-im and rigid limit behavior detecting the absence or presence of ns-im via PINATA does not depend on the assumption of absence or presence of coupling between internal and global motion. The PINATA protocol therefore opens the way for the quantitative analysis of the dynamics of proteins, which undergo domain motions or are unfolded or partially folded. The PINATA protocol is implemented in MATLAB scripts, which are available on request.

Acknowledgements

This work was supported by grants from the Swedish Natural Science Research Council (SW), the Bioteknik Medel Umeå University (SW), the Royal Swedish Academy of Science (GL), and the Kempe Foundation (GL). Janny Hof is thanked for critical reading of the manuscript.

References

Abragam, A. (1961) *The Principles of Nuclear Magnetism*, Clarendon Press, Oxford.

Baber, J.L., Szabo, A. and Tjandra, N. (2001) *J. Am. Chem. Soc.*, **123**, 3953–3959.

Bevington, P.R. (1969) *Data Reduction and Error Analysis*, McGraw-Hill Book Co., New York, NY.

Canet, D., Barthe, P., Mutzenhardt, P. and Roumestrand, C. (2001) *J. Am. Chem. Soc.*, **123**, 4567–4576.
Clore, G.M., Driscoll, P.C., Wingfield, P.T. and Gronenborn, A.M. (1990a) *Biochemistry*, **29**, 7387–7401.
Clore, G.M., Szabo, A., Bax, A., Kay, L.E., Driscoll, P.C. and Gronenborn, A.M. (1990b) *J. Am. Chem. Soc.*, **112**, 4989–4991.
Clore, G.M., Gronenborn, A.M., Szabo, A. and Tjandra, N. (1998) *J. Am. Chem. Soc.*, **120**, 4889–4890.
Dyson, H.J. and Wright, P.E. (1998) *Nat. Struct. Biol.*, **5**, 499–503.
Farrow, N.A., Muhandiram, D.R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D. and Kay, L.E. (1994) *Biochemistry*, **33**, 5984–6003.
Farrow, N.A., Zhang, O.W., FormanKay, J.D. and Kay, L.E. (1997) *Biochemistry*, **36**, 2390–2402.
Farrow, N.A., Zhang, O.W., Szabo, A., Torchia, D.A. and Kay, L.E. (1995) *J. Biomol. NMR*, **6**, 153–162.
Fischer, M.W.F., Majumdar, A. and Zuiderweg, E.R.P. (1998) *Prog. Nucl. Magn. Reson. Spectrosc.*, **33**, 207–272.
Fushman, D., Tjandra, N. and Cowburn, D. (1998) *J Am. Chem. Soc.*, **120**, 10947–10952.
Fushman, D., Tjandra, N. and Cowburn, D. (1999) *J. Am. Chem. Soc.*, **121**, 8577–8582.
Humphrey, W., Dalke, A. and Schulten, K. (1996) *J. Mol. Graphics*, **14**, 33–38.
Ishima, R. and Nagayama, K. (1995a) *Biochemistry*, **34**, 3162–3171.
Ishima, R. and Nagayama, K. (1995b) *J. Magn. Reson.*, **B108**, 73–76.
Ishima, R. and Torchia, D.A. (2000) *Nat. Struct. Biol.*, **7**, 740–743.
Ishima, R., Yamasaki, K. and Nagayama, K. (1995) *J. Biomol. NMR*, **6**, 423–426.
Jin, D.Q., Andrec, M., Montelione, G.T. and Levy, R.M. (1998) *J. Biomol. NMR*, **12**, 471–492.
Kay, L.E. (1998) *Nat. Struct. Biol.*, **5**, 513–517.
Kay, L.E., Torchia, D.A. and Bax, A. (1989) *Biochemistry*, **28**, 8972–8979.
Korzhnev, D.M., Billeter, M., Arseniev, A.S. and Orekhov, V.Y. (2001) *Prog. Nucl. Magn. Reson. Spectrosc.*, **38**, 197–266.
Korzhnev, D.M., Orekhov, V.Y. and Arseniev, A.S. (1997) *J. Magn. Reson.*, **127**, 184–191.
Larsson, G., Schleucher, J., Onions, J., Hermann, S., Grundström, T. and Wijmenga, S.S. (2003) *J. Mol. Biol.*, in press.
Lee, A.L. and Wand, A.J. (1999) *J. Biomol. NMR*, **13**, 101–112.
Lefevre, J.F., Dayie, K.T., Peng, J.W. and Wagner, G. (1996) *Biochemistry*, **35**, 2674–2686.
Lipari, G. and Szabo, A. (1982) *J. Am. Chem. Soc.*, **104**, 4546–4559.
Loria, J.P., Rance, M. and Palmer, A.G. (1999) *J. Am. Chem. Soc.*, **121**, 2331–2332.
Mandel, A.M., Akke, M. and Palmer, A.G. (1995) *J. Mol. Biol.*, **246**, 144–163.
Orekhov, V.Y., Nolde, D.E., Golovanov, A.P. and Arseniev, A.S. (1996) *Appl. Magn. Reson.*, **9**, 581–588.
Papavoine, C.H.M., Christiaans, B.E.C., Folmer, R.H.A., Konings, R.N.H. and Hilbers, C.W. (1998) *J. Mol. Biol.*, **282**, 401–419.
Papavoine, C.H.M., Remerowski, M.L., Horstink, L.M., Konings, R.N.H., Hilbers, C.W. and vandeVen, F.J.M. (1997) *Biochemistry*, **36**, 4015–4026.
Peng, J.W. and Wagner, G. (1995) *Biochemistry*, **34**, 16733–16752.
Renner, C. and Holak, T.A. (2000) *J. Magn. Reson.*, **145**, 192–200.
Schurr, J.M., Babcock, H.P. and Fujimoto, B.S. (1994) *J. Magn. Reson.*, **B105**, 211–224.
Tjandra, N., Feller, S.E., Pastor, R.W. and Bax, A. (1995) *J. Am. Chem. Soc.*, **117**, 12562–12566.
Tjandra, N., Wingfield, P.T., Stahl, S. and Bax, A. (1996) *J. Biomol. NMR*, **8**, 273–284.
Tugarinov, V., Liang, Z., Shapiro, Y.E., Freed, J.H. and Meirovitch, E. (2001) *J. Am. Chem. Soc.*, **123**, 3055–3063.
Viles, J.H., Duggan, B.M., Zaborowski, E., Schwarzinger, S., Huntly, J.J.A., Kroon, G.J.A., Dyson, H.J. and Wright, P.E. (2001) *J. Biomol. NMR*, **21**, 1–9.
Vis, H., Vorgias, C.E., Wilson, K.S., Kaptein, R. and Boelens, R. (1998) *J. Biomol. NMR*, **11**, 265–277.
Woessner, D.E. (1962) *J. Chem. Phys.*, **37**, 647–654.
Wright, P.E. and Dyson, H.J. (1999) *J. Mol. Biol.*, **293**, 321–331.
Zdunek, J., Martinez, G., Schleucher, J., Ippel, H.J., Lycksell, P.-O., Yin, Y., Nilsson, S., Shen, Y., Olivecrona, G. and Wijmenga, S.S. (2003) *Biochemistry*, **42**, 1872–1889.