Journal of Biomolecular NMR, 6 (1995) 423–426
ESCOM

J-Bio NMR 323

Application of the quasi-spectral density function of ^{15}N nuclei to the selection of a motional model for model-free analysis

Rieko Ishima[a,*], Kazuhiko Yamasaki[b] and Kuniaki Nagayama[a,c]

[a]*Nagayama Protein Array Project, ERATO, JRDC, Tsukuba Research Consortium, Pilot Laboratory, 5-9-1 Tokodai, Tsukuba 300-26, Japan*
[b]*Protein Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565, Japan*
[c]*Department of Pure and Applied Sciences, College of Arts and Sciences, University of Tokyo, Komaba, Meguro, Tokyo 153, Japan*

Received 31 July 1995
Accepted 11 November 1995

Keywords: Protein; Relaxation; RNase HI

Summary

Parameters used in model-free analysis were related to simulated spectral density functions in a frequency region experimentally obtained by quasi-spectral density function analysis of ^{15}N nuclei. Five kinds of motional models used in recent model-free analyses were characterized by a simple classification of the experimental spectral density function. We demonstrate advantages and limitations of each of the motional models. To verify the character of the models, model selection using experimental spectral density functions was examined.

Recently, magnetic relaxation of ^{15}N nuclei has been used to probe the backbone dynamics of proteins in solution (Kay et al., 1989; Clore et al., 1990a; Peng and Wagner, 1992a; Stone et al., 1992; Farrow et al., 1994; Tycko, 1994). Model-free analysis has been employed extensively in the interpretation of such relaxation data (Lipari and Szabo, 1982a,b; Clore et al., 1990b). In the original model-free formalism used by Lipari and Szabo (1982a,b), the precision of their motional model was evaluated using simulation of spectral density functions. Pioneering work by Peng and Wagner (1992b) enabled one to obtain experimental spectral density functions and to compare them to the model-free parameters. However, a set of five motional models used in recent model-free analyses has not been evaluated by spectral density functions (Clore et al., 1990b; Stone et al., 1992; Cheng et al., 1994; Mandel et al., 1995; Yamasaki et al., 1995).

The number of motional models in use today has grown to five, despite the original name of 'model-free'. The physical motional models may still be of the model-free type, but the number of five may be too large to call it 'model-free'. The models are applied to distinguish differences in motional properties. The model yielding the smallest values of χ^2 for each backbone amide is principally selected. The parameter χ^2 is defined as (Palmer et al., 1991):

$$\chi^2 = (R_{1,exp} - R_{1,cal})^2 / \sigma_1^2 + (R_{2,exp} - R_{2,cal})^2 / \sigma_2^2 + (R_{NOE,exp} - R_{NOE,cal})^2 / \sigma_{NOE}^2 \quad (1)$$

where R_1, R_2 and R_{NOE} are the longitudinal, transverse and NOE relaxation rates, respectively. Parameter σ is the uncertainty stemming from experimental errors (Palmer et al., 1991). In this method, simplified models are selected because of the limited number of experimental data (three points in most cases). The χ^2 values from models with different numbers of unknown parameters are compared using an *F*-distribution. While this statistical method is quantitative, it may not describe all types of motions in proteins. A limited number of experimental data may cause improper selection of the motional models. It is therefore necessary to characterize the five models systematically using experimentally derived spectral density functions.

In this report, we characterize the five motional models employed in recent model-free analyses using the quasi-spectral density function (QSDF) of ^{15}N nuclei (Farrow et al., 1995; Ishima and Nagayama, 1995a,b). The QSDF

*To whom correspondence should be addressed at: Division of Molecular and Structural Biology, Ontario Cancer Institute, Princess Margaret Hospital, 610 University Avenue, Toronto, ON, Canada M5G 2M9.

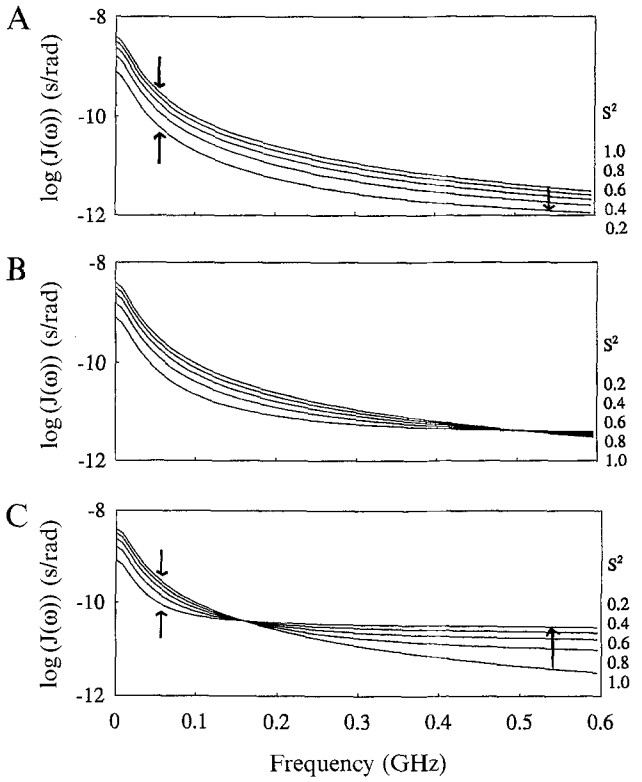

Fig. 1. Spectral density function for model 2 (Table 1) at various correlation times for internal motion (τ_e) and various generalized order parameters (S^2). The rotational correlation time of a molecule, τ_R, was assumed to be 10 ns as a typical value in protein NMR. The τ_e value was assumed to be (A) 1 ps, (B) 10 ps and (C) 100 ps. The S^2 value was changed as indicated in the graphs. The arrows indicate positions at 0 and 540 MHz.

is a simplification of the original proposal by Peng and Wagner (1992b) and is advantageous in that three values of a spectral density function, J(0), $J(\omega_N)$ and $J(\omega_H+\omega_N)$, are determined using the same experimental data as used in the model-free analysis. First, we examine the dependencies of the model-free parameters on simulated values of J(0), $J(\omega_N)$ and $J(\omega_H+\omega_N)$ at the magnetic field strengths employed in the experiment. Next, motional models are related to the experimental spectral density function and verified by a model selection using J(0) and $J(\omega_H+\omega_N)$. All analyses presented here have been performed using ribonuclease HI from *E. coli* (RNase HI; Yamasaki et al., 1995). Relaxation data as reported earlier (Yamasaki et al., 1995) were used, in which the angular frequencies at ω_N and $\omega_H+\omega_N$ correspond to 60 and 540 MHz, respectively.

Figure 1 shows the calculated profiles of a spectral density function using the Lipari–Szabo formalism (model 2 in Table 1). We used this model to demonstrate the contribution of internal motions to the spectral density function. At a correlation time for the internal motion $\tau_e = 1$ ps and at 540 MHz (Fig. 1A), the value of the spectral density function increases as a function of the generalized order parameter (S^2). In contrast, at $\tau_e = 100$ ps (Fig. 1C), the corresponding value decreases as a function of S^2. Consequently, at any S^2 value < 1.0, the value of the spectral density function at 540 MHz is always larger at $\tau_e = 100$ ps (Fig. 1C) than at $\tau_e = 1$ ps (Fig. 1A). J(0) decreases at any τ_e value when S^2 decreases. J(60 MHz) shows the same tendency as seen for J(0) in the range of τ_e shown in Fig. 1. However, the variety in J(60 MHz) values reduces when τ_e increases.

Therefore, dependencies of the model-free parameters (for model 2) on the values of the calculated spectral density function can be summarized as follows: (i) an increase in τ_e is related to an increase in J(540 MHz); (ii) a decrease in S^2 is related to a decrease in J(0); and (iii) J(0) increases when an exchange term (R_{ex}) contributes significantly to the relaxation. The R_{ex} term has been explained in a previous report (Ishima and Nagayama, 1995a), where R_{ex} is involved in J(0) in our definition. J(60 MHz) is one of three important data points, although a relationship between J(0) and S^2 (case ii) is representatively substituted for that of J(60 MHz). These results indicate that the parameters are linearly related to the values of a spectral density function at 0 and 540 MHz.

Based on Fig. 1 and similar data for the other models (not shown, as model 2 is essential), J(0) and J(540 MHz) can be categorized into six regimes, using the definitions of 'Large', 'Medium' or 'Small' for J(0) and 'Large' or 'Small' for J(540 MHz). The values for these criteria are given by the combination of averages and standard deviations of J(0) and J(540 MHz). The categories are summarized in Table 1: a significant contribution of the R_{ex} term categorizes J(0) as 'Large'. A large decrease in S^2 is represented by the order parameter for slow internal motion (S_s^2), which categorizes J(0) as 'Small'. No contribution of either R_{ex} or S_s^2 categorizes J(0) as 'Medium'. The J(540 MHz) values are categorized as 'Large' or 'Small', depending on τ_e, as described in Fig. 1. Consequently, as the parameters S_s^2, R_{ex} and τ_e have been given in an individual model, motional models can be characterized by J(0) and J(540 MHz). A decrease in J(0) is expressed by model 3, which contains the S_s^2 term. An increase in J(0) is expressed by models 4 and 5, which contain R_{ex}. Models 2 and 3 can express an increase in J(540 MHz), using the τ_e term. These regimes were related to the model-free models as described above (Table 1). The possibility that both J(0) and J(540 MHz) values become 'Small' was not used, since a globular protein is assumed. In such a protein, internal motion is restricted while a rotation of the molecule is dominant in $J(\omega)$.

To verify the character of the models, a model selection using J(0) and J(540 MHz) obtained by QSDF analysis was performed. In this communication, we use the following criteria to separate 'Large', 'Medium' and 'Small': the $J(540\text{ MHz})_{ave} + J(540\text{ MHz})_{dev}$ value was used as a criterion to separate 'Large' and 'Small' in J(540 MHz).

TABLE 1
MODEL-FREE MODELS, PARAMETERS AND CATEGORIES OF A SPECTRAL DENSITY FUNCTION AT 0 AND 540 MHz

Model number[a]	Model function J(ω)	Parameters[b]	Simple categories[c]	
			J(540 MHz)	J(0)
5	$0.4\{S^2\tau_R/(1+\omega^2\tau_R^2)+(1-S^2)\tau_i/(1+\omega^2\tau_i^2)\}$, πR_{ex}	S^2, τ_e, R_{ex}	Large	Large
2	$0.4\{S^2\tau_R/(1+\omega^2\tau_R^2)+(1-S^2)\tau_i/(1+\omega^2\tau_i^2)\}$	S^2, τ_e	Large	Medium
3	$0.4\,S_f^2\{S_s^2\tau_R/(1+\omega^2\tau_R^2)+(1-S_s^2)\tau_i/(1+\omega^2\tau_i^2)\}$	S^2, τ_e, S_s^2	Large	Small
4	$0.4\,S^2\tau_R/(1+\omega^2\tau_R^2)$, πR_{ex}	S^2, R_{ex}	Small	Large
1	$0.4\,S^2\tau_R/(1+\omega^2\tau_R^2)$	S^2	Small	Medium
No			Small	Small

[a] The number is the same as described previously (Yamasaki et al., 1995).
[b] S^2 is the generalized order parameter; S_s^2 is the order parameter for slow internal motion; R_{ex} is the fluctuation of the chemical shifts evaluated as an exchange term; τ_R is the rotational correlation time of the molecule; τ_e is the correlation time of internal motion; $\tau_i^{-1}=\tau_R^{-1}+\tau_e^{-1}$. S_f^2 is given by $S^2=S_f^2S_s^2$. R_{ex} is given as an additional term in the transverse relaxation time.
[c] The categorization is described in the text.

Here, the 'ave' and 'dev' are abbreviations for an average and a standard deviation for all residues in the calculation. The $J(0)_{ave}+J(0)_{dev}$ and $J(0)_{ave}-J(0)_{dev}$ values were used to separate 'Large', 'Medium' and 'Small' in J(0). Experimental errors in the values for J(0) and J(540 MHz) were not taken into account here. The result is shown in Fig. 2B. The selected models were indicated on the values for J(0) and J(540 MHz). As a reference, Fig. 2A shows the values for J(0) and J(540 MHz), which were calculated using the model-free parameters determined by statistical model selection (Yamasaki et al., 1995). In Fig. 2A the J(0) values determined using model 3 were lower than the average, while J(540 MHz) values determined using models 2 and 3 were larger than the average. The smallest values for J(540 MHz), bounding the plot as a flat bottom, are given by models 1 and 4 (Fig. 2A), whose spectral density functions are given by one Lorentzian term (Table 1).

In Fig. 2B, the models of 80 residues out of 124 investigated residues agree with those given by the statistical method (Yamasaki et al., 1995). We also tested two other criteria for the model selection, with $J(0)_{ave}\pm0.5J(0)_{dev}$ and $J(0)_{ave}\pm0.25J(0)_{dev}$. Selected models of 12 residues did not agree with those given by the statistical analysis employ-

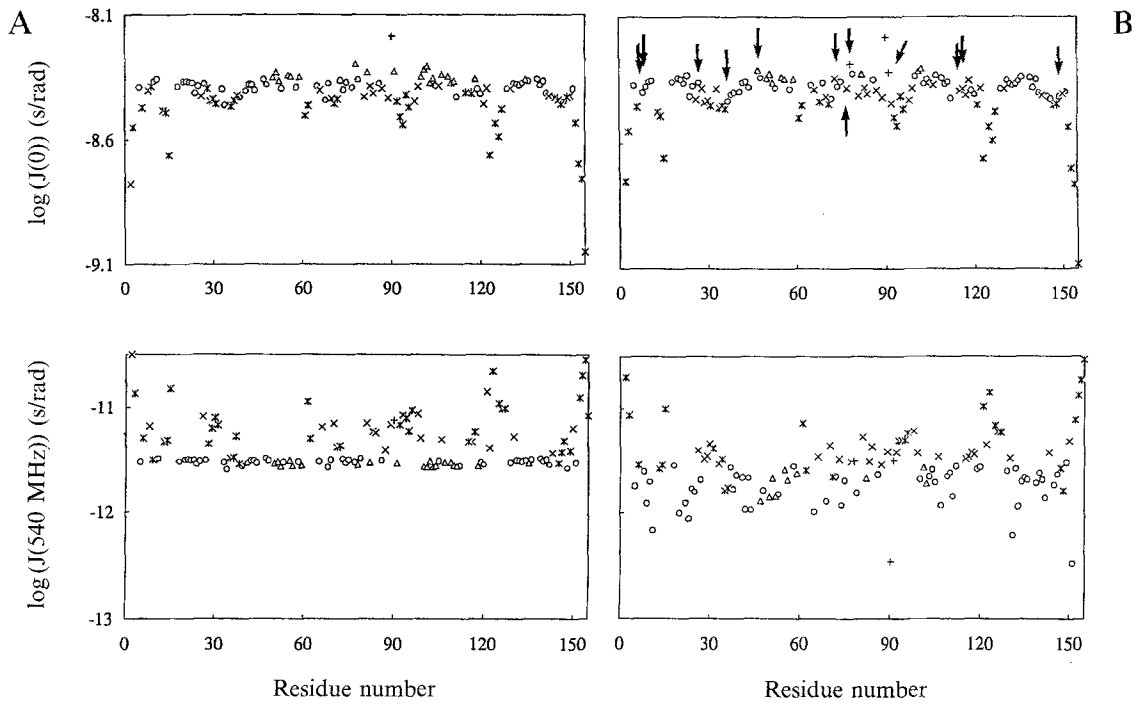

Fig. 2. The J(0) and J(540 MHz) values obtained by (A) the results of model-free analysis and (B) QSDF analysis of RNase HI. For each value, the motional models, 1, 2, 3, 4 and 5 (Table 1) are plotted as ○, x, *, Δ and +, respectively. In (A), the values were calculated using parameters determined in the model-free analysis, and the statistical model selection was performed to select motional models. In (B), the values were obtained using QSDF analysis of the relaxation data, and the model selection was performed using the values of the spectral densities. The ^{15}N relaxation measurements and their model-free analysis were performed by Yamasaki et al. (1995). Arrows in (B) indicate residues whose model numbers are different from those determined by the statistical method in all three trials (see text).

ing any three criteria. The models of five residues among the 12 were changed to more complicated ones with a decrease in χ^2 values: from model 1 to model 2, from model 1 to model 4, or from model 4 to model 5. For these five residues, the complicated models were not statistically distinguished in the previous analysis by Yamasaki et al. (1995). The models of seven residues among the 12 were changed to simpler ones than those previously reported: from model 3 to model 1, from model 2 to model 1, or from model 3 to model 2. The χ^2 values for these residues were larger than those in the literature. In these residues, however, the χ^2 values are still less than the 10% critical value (Yamasaki et al., 1995), or the *F*-distributions were still more than the 10% critical value. Therefore, disagreements in the 12 residues were not significant. Furthermore, five residues among the 12 were the same as those chosen by Mandel et al. (1995), who also employed the model-free analysis for RNase HI using ^{15}N relaxation at 500 MHz. Our result of the model selection showed arbitrarity of the criteria, although it basically agreed with the results of the statistical method. This ensures characterization of the motional models by spectral density functions, as shown in Table 1.

In this report, first, parameters used in the model-free analysis were characterized by the spectral density function. It was shown that J(60 MHz) becomes insensitive to S^2 when τ_e increases (Fig. 1). This explains well the results of our previous report (Ishima and Nagayama, 1995a), where it was found that experimental $J(\omega_N)$ values did not significantly change in proteins. Therefore, it is interpreted that a significant contribution of internal motion on the nanosecond time scale is certainly present (Ishima et al., 1995). The importance of the frequency range observed was shown. It is a critical factor that affects the motional models and has been described in evaluating the rotational correlation times (τ_R). These values are slightly different between various kinds of relaxation data applied (Tjandra et al., 1995). One major reason will be differences in frequencies provided by individual relaxation data.

Second, motional models used in the model-free analysis were characterized by a categorization of J(0) and J(540 MHz). A simple categorization is qualitatively satisfactory. The reason will be a strong correlation between parameters and individual values of the spectral density, which is caused by the limited number of experimental data. Our results show that a major variety of the experimental spectral density functions could be described by the five motional models, although there are minor function forms that cannot be expressed in the five model-free models. One minor function form is that determined by both R_{ex} and S_s^2. In this form, J(0) is judged as 'Medium' and neither R_{ex} nor S_s^2 is used in the motional model. We do not know a physical motional model for the minor form, although it exists in proteins (Powers et al., 1992). Another minor form is observed in the case where significant uncertainty exists in $J(\omega_H+\omega_N)$. Model 1 or 4 will be selected in this case. However, these models cannot express large values of $J(\omega_H+\omega_N)$ if these exist, and are therefore different from the major part in model 1 or 4. A proper model in this case is a function defined only between the frequencies 0 and ω_N. In conclusion, model-free models can basically express variety of observed spectral density functions. As described in this report, it is important to interpret relaxation data in terms of spectral density functions.

Acknowledgements

We greatly thank Drs. S. Bagby and P. Morin at the Ontario Cancer Institute (Toronto, Canada), and K. Fukushima in the Nagayama Protein Array Project (Tsukuba, Japan), for reading of the manuscript.

References

Cheng, J.W., Lepre, C.A. and Moore, J.M. (1994) *Biochemistry*, **33**, 4093–4100.

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.

Farrow, N.A., Muhandiram, 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., Forman-Kay, J.D. and Kay, L.E. (1995) *Biochemistry*, **34**, 868–878.

Ishima, R. and Nagayama, K. (1995a) *Biochemistry*, **34**, 3162–3171.

Ishima, R. and Nagayama, K. (1995b) *J. Magn. Reson. Ser. B*, **108**, 73–76.

Ishima, R., Yamasaki, K., Saito, M. and Nagayama, K. (1995) *J. Biomol. NMR*, **6**, 217–220.

Kay, L.E., Torchia, D.A. and Bax, A. (1989) *Biochemistry*, **28**, 8972–8979.

Lipari, G. and Szabo, A. (1982a) *J. Am. Chem. Soc.*, **104**, 4546–4559.

Lipari, G. and Szabo, A. (1982b) *J. Am. Chem. Soc.*, **104**, 4559–4570.

Mandel, A.M., Akke, M. and Palmer III, A.G. (1995) *J. Mol. Biol.*, **246**, 144–163.

Palmer III, A.G., Rance, M. and Wright, P. (1991) *J. Am. Chem. Soc.*, **113**, 4371–4380.

Peng, J.W. and Wagner, G. (1992a) *J. Magn. Reson.*, **98**, 308–332.

Peng, J.W. and Wagner, G. (1992b) *Biochemistry*, **31**, 8571–8586.

Powers, R., Clore, G.M., Stahl, S.J., Wingfield, P.T. and Gronenborn, A. (1992) *Biochemistry*, **31**, 9150–9157.

Stone, M.J., Fairbrother, W.J., Palmer III, A.G., Reizer, J., Saier, M.H. and Wright, P.E. (1992) *Biochemistry*, **31**, 4394–4406.

Tjandra, N., Kuboniwa, H., Ren, H. and Bax, A. (1995) *Eur. J. Biochem.*, **230**, 1014–1024.

Tycko, R., *Nuclear Magnetic Resonance Probes of Molecular Dynamics*, Kluwer, Dordrecht, 1994.

Yamasaki, K., Saito, M., Oobatake, M. and Kanaya, S. (1995) *Biochemistry*, **34**, 6587–6601.