Journal of Biomolecular NMR (2006) 36:215–224
DOI 10.1007/s10858-006-9083-0

Article

A suite of *Mathematica* notebooks for the analysis of protein main chain ^{15}N NMR relaxation data

Leo Spyracopoulos*
Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7

Received 21 June 2006; Accepted 30 August 2006

Key words: model-free analysis, ^{15}N NMR relaxation, protein dynamics

Abstract

A suite of *Mathematica* notebooks has been designed to ease the analysis of protein main chain ^{15}N NMR relaxation data collected at a single magnetic field strength. Individual notebooks were developed to perform the following tasks: nonlinear fitting of ^{15}N-T_1 and -T_2 relaxation decays to a two parameter exponential decay, calculation of the principal components of the inertia tensor from protein structural coordinates, nonlinear optimization of the principal components and orientation of the axially symmetric rotational diffusion tensor, model-free analysis of ^{15}N-T_1, -T_2, and $\{^1H\}$–^{15}N NOE data, and reduced spectral density analysis of the relaxation data. The principle features of the notebooks include use of a minimal number of input files, integrated notebook data management, ease of use, cross-platform compatibility, automatic visualization of results and generation of high-quality graphics, and output of analyses in text format.

Abbreviations: AIC – Akaike's information criteria.

Introduction

The combination of multinuclear, multidimensional protein NMR spectroscopy and the Lipari-Szabo model-free approach has allowed for detailed characterization of pico to nanosecond timescale dynamic processes in proteins in a residue specific fashion (Lipari and Szabo, 1982a, b, Kay et al., 1989). Long-term development of the original Lipari-Szabo approach (Clore et al., 1990, Palmer et al., 1991, Mandel et al., 1995) and advances in pulse sequence methodology (Farrow et al., 1994) have allowed for new insights into biological processes carried out by proteins through investigation of the role of fast time scale motions into the affinities of ligands for proteins (Akke et al., 1993), protein stability (Yang and Kay, 1996), and enzyme function (Nicholson et al., 1995). Methods, applications, and functional interpretation of fast time scale dynamics for proteins have recently been comprehensively reviewed (Jarymowycz and Stone, 2006).

Main chain amide ^{15}N-T_1, and -T_2 relaxation times are determined from nonlinear fits of changes in peak intensities in 2D ^{1}H–^{15}N HSQC NMR spectra to a two-parameter exponential decay. The $\{^1H\}$–^{15}N NOE is given by the ratio of the ^{1}H-saturated/unsaturated peak heights. ^{15}N relaxation data are typically acquired at a single magnetic field strength due to spectrometer time constraints. Subsequent to determination of per residue T_1, T_2, and $\{^1H\}$–^{15}N NOE values, it is customary to perform a Lipari-Szabo, or model-free analysis (Lipari and Szabo, 1982a, b). For modern model-free analyses, the spectral density is written in terms of an order parameter (S^2) that is indicative of the degree of spatial restriction for a given internuclear bond vector, with one or two kinds of internal

*To whom correspondence should be addressed.
E-mail: leo.spyracopoulos@ualberta.ca

correlation times, and the assumption of a single, global molecular rotational correlation time (Mandel et al., 1995). However, prior to conducting a model-free analysis, the rotational tumbling of the protein is usually characterized. It has long been appreciated that rotational diffusion anisotropy for non-spherical molecules is reflected in the relaxation rates T_1 and T_2 (Woessner, 1962). For proteins of known structure, the components of the rotational diffusion tensor can be calculated from T_1/T_2 (Tjandra et al., 1995), or calculated theoretically (de la Torre et al., 2000). For proteins of unknown structure, the components of the diffusion tensor may be calculated from the distribution of T_1/T_2 values for the protein (Clore et al., 1998).

For unfolded proteins, or proteins that weakly self-associate, it is difficult to extract meaningful values of S^2. For these cases, main chain dynamics are often assessed using one of two approaches: full or reduced spectral density mapping (Peng and Wagner, 1992, Farrow et al., 1995). Reduced spectral density mapping is usually the methodology of choice given that fewer experimental measurements are required.

There are a number of programs available for the analysis of main chain amide ^{15}N NMR relaxation data from proteins (Orekhov et al., 1995, Mandel et al., 1995, Cordier et al., 1998, Blackledge et al., 1998, Ghose et al., 2001, Idiyatullin et al., 2003). Herein, we describe a suite of *Mathematica* notebooks for analysis of ^{15}N NMR relaxation data from proteins that have been designed primarily for ease of use, integrated data management and visualization, and cross-platform compatibility. There are individual notebooks capable of calculating ^{15}N-T_1 and -T_2 relaxation times from peak intensity decay data, determining the components of the inertia tensor, determining the components of the axially symmetric, rotational diffusion tensor, performing standard model-free analysis, and performing spectral density analysis. The notebooks require only four input files: three lists of peak intensities determined using the 'rh' utility from the program Sparky for ^{15}N-T_1, -T_2, and $\{^1$H$\}$–^{15}N NOE, and a PDB coordinate file if analysis of anisotropic rotational diffusion is necessary. The notebooks were designed to have an integrated dependency on input and output files, that is, output from one notebook is read as input in subsequent notebooks. Additional features of the notebooks include: self-contained, concise instructions for ease of use, compatibility of *Mathematica* across many computer platforms, graphical display of input data and results that can be easily exported as vector graphics files in the PostScript language, and output of analyses in text format.

Materials and methods

Protein production and NMR spectroscopy

Production of [$U-^{15}$N; $U-^{13}$C; $U-^{2}$H($\sim$50%)]-Mms2, acquisition and analysis of 2D ^{1}H–^{15}N HSQC NMR spectra for measurement of ^{15}N-T_1, T_2, and $\{^1$H$\}$–^{15}N NOE are described in (Spyracopoulos et al., 2005).

Notebook for analysis of $^{15}N-T_1$ and T_2

The notebook requires as input a file directory and file name for the relaxation decay data (*Mathematica* SetDirectory and ReadList commands). The file must have the format: "$x\ y_1\ y_2\ z_1 \ldots\ z_n$", where x, y, and z are real numbers separated by spaces with x = residue number, y_1 = decay constant, y_2 = decay constant error, $z_1 \ldots z_n$ are peak heights 1 to n, where n is the total number of relaxation delays used. It should be noted that y_1 and y_2 are not required by the notebook, and are included as a matter of convenience as they are included in the output from the utility for the program Sparky for relaxation decay peak picking ('rh' command). If the Sparky program is not employed, any real number can be substituted for y_1 and y_2. Additional input includes the value of the noise from the NMR spectra, the relaxation delays in ms, the number of Monte Carlo simulations to be conducted (typically 500), a flag that allows the notebook to analyze data for only the first residue, or all residues, a flag to write the parameters from the fit to a text file, an upper constraint for the T_2 value (dependent on protein), the confidence level for the critical χ^2 (typically 95%), and a noise scale factor (typically between 1 and $\sqrt{2}$). This notebook and subsequently described notebooks have self-contained, succinct instructions.

^{15}N-T_1 and -T_2 relaxation decays are subjected to nonlinear fits to a two parameter exponential decay using the Nonlinear Regress function from the standard *Mathematica* package Statistics 'Nonlinear Fit' to minimize the value of χ^2. A χ^2 goodness-of-fit test is administered and the exact critical χ^2 is compared to that determined from Monte Carlo simulations for a single residue, as previously described (Palmer et al., 1991). The χ^2 distribution from the Monte Carlo simulations (Figure 1a), 500 random picks from the exact χ^2 distribution, and the exact χ^2 distribution (Figure 1b), are graphically displayed. The experimental and critical χ^2 values are plotted as a function of residue number. In the next stage of the notebook, the mean and standard deviation for the experimental χ^2 values are calculated, and a trimmed data set is determined after excluding experimental χ^2 values that exceed the mean $\chi^2 \pm \sigma$. Assuming that the data set is large enough, and properly samples the ideal χ^2 distribution, the mean for the trimmed data set should be equal to the number of degrees of freedom, which is equal to the number of relaxation delays minus the number of fitted parameters. The noise scale factor can be adjusted between 1 and $\sqrt{2}$ until there is better agreement between the trimmed mean and the number of degrees of freedom. Next, histograms of experimental and trimmed χ^2 values are displayed for comparison to the exact and Monte Carlo χ^2 distributions (Figure 1c). It should be noted that if the noise scale factor must exceed $\sqrt{2}$ in order for $\chi^2 <$ critical χ^2 for the majority of residues, then it is possible that there is a systematic error that is not accounted for in the root mean square noise in the spectra, or the data are not appropriately fit by a two parameter exponential decay (Palmer et al., 1991).

The relaxation decay constants (T_1 or T_2) and their errors from the covariance matrix are plotted as a function of residue number. The notebook also tabulates residue number, T_1 or T_2, T_1 or T_2 error, I_0, I_0 error, χ^2, and critical χ^2, where I_0 is a fitted parameter for the peak intensity at zero time, for all residues, and an additional table that includes only residues whose experimental $\chi^2 <$ critical χ^2. The

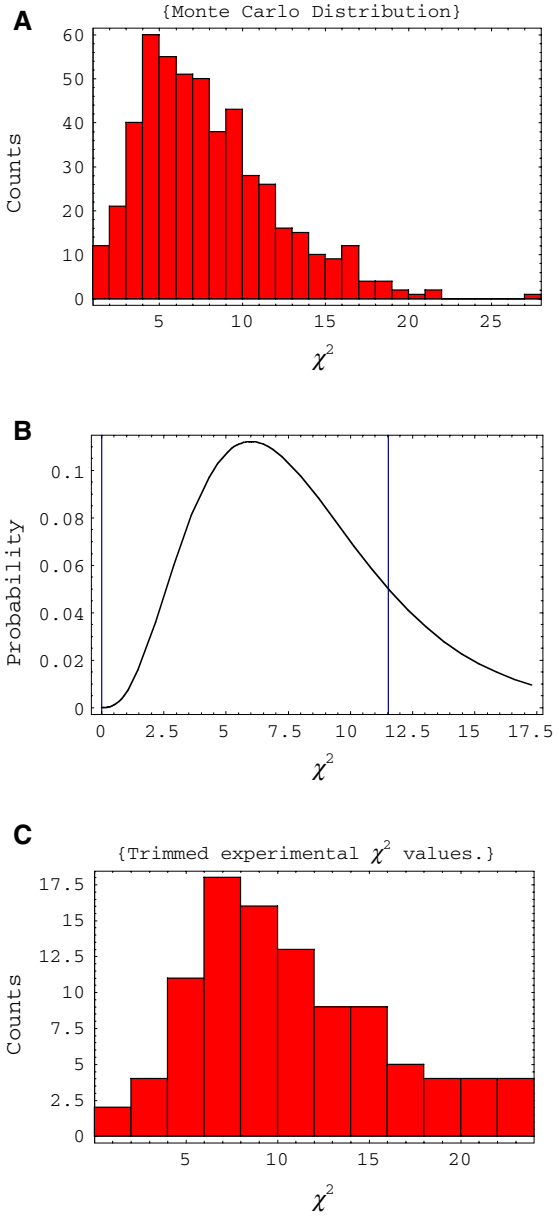

Figure 1. Graphical output for statistical analyses conducted within the notebook for analysis of ^{15}N-T_1 and - T_2 relaxation decays. (A) Histogram of χ^2 values for 500 Monte Carlo simulations of fits to a two-parameter exponential decay for ^{15}N-T_1 for residue 8 from the protein Mms2. The experimental decay was fit to the function $I_0 \exp(-t/T_1)$ with $I_0 = 316548$, $T_1 = 763.2$, the noise in the spectrum was 8784, and there are 8 degrees of freedom (10 relaxation decays). (B) Exact χ^2 distribution for 8 degrees of freedom. The right grid line ($\chi^2 = 11.5$) is plotted at the 95% confidence level. (C) Trimmed χ^2 distributions. This figure and all subsequent figures (except Figure 8) were produced directly by the various notebooks described in the text. Figures were exported as PostScript vector graphics.

notebook provides per residue plots of peak intensity against relaxation delay, and the associated curve fit. Each plot includes residue number, decay parameters, χ^2 and critical χ^2, and a warning if χ^2 exceeds the critical χ^2 (Figure 2).

Finally, one of the development criteria was that the notebooks should operate in an integrated fashion. Thus, textual output from the ^{15}N-T_1 and -T_2 notebook is employed by notebooks that determine the isotropic rotational correlation time, components of the rotational diffusion tensor, and perform model-free and reduced spectral density analyses.

Notebook for determination of the components of the inertia tensor

The matrix describing the moment of inertia for a given protein is calculated from the atomic coordinates according to the equation given in (Foote and Raman, 2000). The eigenvalues and eigenvectors of the 3×3 inertia tensor are calculated with the *Mathematica* matrix operation commands Eigenvalues and Eigenvectors. The eigenvalues, or principal axes of the inertia tensor, are normalized to the largest component. The eigenvectors are then translated to the center of mass coordinates, and the eigenvector matrix and its transpose are used as rotation matrices to rotate the Cα atoms of the protein into the frame of the principal axes. After multiplication by an arbitrary scale factor, the eigenvectors are plotted in 3D along with a trace through the Cα atoms of the protein of interest using the commands Graphics3D and ScatterPlot3D from the Graphics3D standard package (Figure 3). The package RealTime3D allows for real time rotation of the 3D plot.

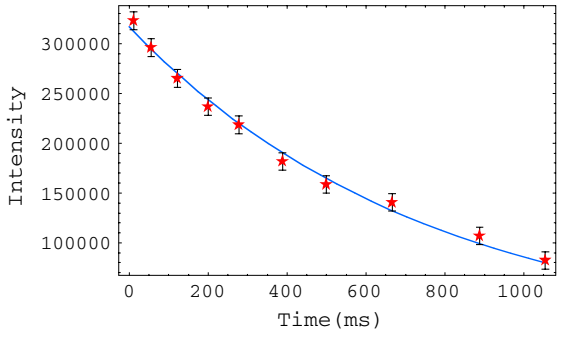

Figure 2. Graphical output for nonlinear curve fits conducted within the notebook for analysis of ^{15}N-T_1 and T_2 relaxation decays. Experimental ^{15}N-T_1 decay curve and associated curve fit for residue 8 from the protein Mms2. Additional details are given in the legend to Figure 1A.

Notebook for determination of isotropic correlation time, R_{ex} terms, and components of the rotational diffusion tensor

The principal components of the rotational diffusion tensor for axially symmetric, prolate and oblate anisotropic rotational diffusion are determined from T_1/T_2 in a similar fashion as previously described by Bax and co-workers (Tjandra et al., 1995). The notebook requires as input text files that are generated by the ^{15}N-T_1 and -T_2 notebook. In addition, a file for the $\{^1$H$\}$–^{15}N NOE is required, with format: "x y_1 y_2 z_1 z_2", where x, y, and z are real numbers separated by spaces with x = residue number, y_1 = unsaturated peak intensity, y_2 = saturated peak intensity, z_1 = NOE, and z_2 = NOE error. However, the notebook does not utilize y_1 and y_2,

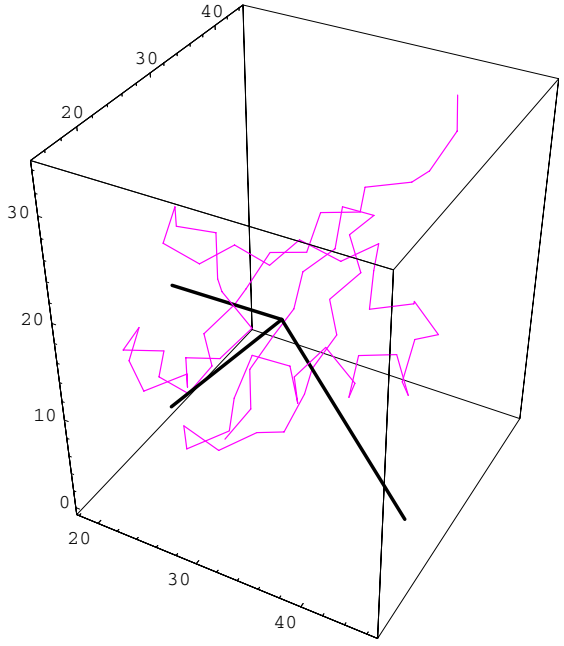

Figure 3. Graphical output generated within the notebook for determination of the principal axes of the inertia tensor. The coordinates from PDB ID 1UBQ were used to generate a trace through the Cα atoms of ubiquitin, and the center of mass coordinates and eigenvector matrix were used to plot vectors corresponding to the principal axes of inertia. The figure can be rotated in real time.

therefore any real number can be substituted for these values. Additional input parameters include a flag that allows the notebook to analyze data for a user-defined number of residues, or all residues, a flag to perform an isotropic correlation time analysis only, or an axially symmetric, anisotropic rotational correlation time analysis, a variable to set the spectrometer frequency, and a cut-off value for the $\{^1H\}$–^{15}N NOE, below which residues are excluded from the analysis due to flexibility. The NOE cut-off value is typically 0.65, but it should be adjusted according to the overall correlation time of the protein and the static magnetic field strength. The user is prompted to locate a protein PDB file using a file browser, if the option is chosen to perform an analysis of anisotropic rotational diffusion anisotropy. Finally, the notebook uses Stokes' law and Perrin's equations to determine constraints for the values of the principal components of the rotational diffusion tensor (Cantor and Schimmel, 1980). Therefore, the notebook also requires a parameter for the value of $D_{\|}/D_{\perp}$ that can be estimated from the coordinates of the protein using the program HYDRONMR (de la Torre et al., 2000). In addition, a parameter for temperature, a parameter to make small adjustments to solution viscosity and a parameter for the number of residues in the protein are also required to generate constraints for D_x, D_y, and D_z.

The notebook generates per residue plots of R_1, R_2, and NOE, and their associated errors, creates a table of R_1, error, R_2, error, NOE, error, and N–H^N direction cosines (calculated from the PDB file), for residues with NOE values greater than the specified cut-off value. A per residue plot of the product R_1R_2 and its error is generated, and gridlines for the average $R_1R_2 \pm \sigma$ are displayed (Figure 4). These data are used to provide per residue estimates for R_{ex} terms according to the method of Bracken and co-workers (Kneller et al., 2002).

Subsequent to removing residues on the basis of NOE, per residue T_1 and T_2 data are used to test for the presence of conformational exchange (Tjandra et al., 1995). These residues are excluded from further analyses, and a table of R_1, error, R_2, error, NOE, error, and N–H^N direction cosines is displayed for residues not found to be undergoing chemical exchange. A per residue plot of the T_1/T_2 ratio and error, and a histogram of the T_1/T_2 ratios (Clore et al., 1998), are displayed.

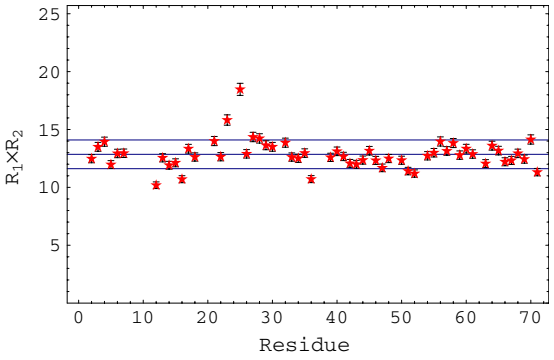

Figure 4. Graphical output generated within the notebook for determination of the principal axes of the rotational diffusion tensor. R_1R_2 is plotted as a function of residue number with gridlines for $\langle R_1R_2 \rangle \pm \sigma$ displayed. The data are used to provide estimates for R_{ex}.

The final section of the notebook performs a numerical optimization of the per residue sum of the squared differences between experimental T_1/T_2 ratios and theoretical T_1/T_2 ratios using the NMinimize command with respect to the parameters describing the overall correlation time for isotropic rotational diffusion (Kay et al., 1989), or D_x, ($D_y = D_x$), D_z, and the angles θ and ϕ that describe the orientation of D_z in the PDB frame for anisotropic, axially symmetric rotational diffusion (Tjandra et al., 1995). The values of the optimized parameters, χ^2, and $\chi^2 + 2k$ (AIC values for model selection) are displayed. The notebook generates a text file containing residue numbers and direction cosines for the N–H^N vectors calculated from the PDB file for subsequent use in the notebook that performs model-free analysis.

Notebook for model-free analysis

A model free analysis (Lipari and Szabo, 1982a, b) is performed using five forms for the spectral density function as outlined in (Mandel et al., 1995), with the option of accounting for axially symmetric rotational diffusion anisotropy (Barbato et al., 1992). Text files that are generated by the ^{15}N-T_1 and -T_2 notebook are used as input for this notebook, other input includes the text file containing N–H^N direction cosines generated by the diffusion tensor notebook. Further variables and parameters include a flag that gives the option to conduct the analysis using either isotropic or

anisotropic rotational tumbling, a flag to activate Monte Carlo analysis and a parameter to set the number of Monte Carlo trials to conduct, a flag to analyze only the first or all residues, a parameter to set the overall correlation time for isotropic rotational tumbling, and a parameter to set the spectrometer frequency. The chemical shift anisotropy for the backbone amide nitrogen, and the length of the N–H^N bond are set to −172 ppm, and 1.02 Å, respectively.

For a given residue, numerical optimization of the parameters for the five spectral density models is conducted using the NMinimize command with the following constraints: $1 \geq (S_f^2 \text{or } S_s^2) \geq 0$, $\tau_c > \tau_f > 0$, $0 \leq R_{ex} < 50\ s^{-1}$, and for the two-time scale model (Clore et al., 1990) $\tau_f = 0$, and $\tau_c \geq \tau_s \geq 0$. Following optimization, models are selected according to AIC (d'Auvergne and Gooley, 2003). If the option is chosen, the selected model for a given residue is then subjected to Monte Carlo analysis (Palmer et al., 1991) in order to estimate errors associated with optimized parameters. For some residues, a small fraction of Monte Carlo simulations for the S^2, τ_f model produced τ_f values that were significantly larger than average, leading to overestimates of the error. Thus, values of τ_f that exceed the mean τ_f by 2.5 σ were removed from the Monte Carlo data sets, and the resulting statistical analysis produced errors closer to expected values.

The notebook initially generates per residue plots of R_1, R_2, and $\{^1H\}$–^{15}N NOE, and their errors (Figure 5), creates tables listing residue number, optimized parameters, and AIC values for each of the five models. Following Monte Carlo analysis, the notebook tabulates lists of residue, optimized parameters and their errors, and the selected model. The notebook generates per residue plots of S^2, τ_f, τ_s, R_{ex}, and their errors for the selected model (Figure 6), and exports these results as text files.

Notebook for reduced spectral density analysis

For unfolded proteins, or proteins that weakly self-associate, it is not straightforward to calculate reliable values of S^2, and it is common to perform a reduced spectral density analysis (Farrow et al., 1995). As in the case for the model free notebook, text files that are created by the ^{15}N-T_1 and T_2 notebook are used as input for this notebook. Other input parameters include a flag to activate Monte Carlo simulations, the number of Monte Carlo simulations to conduct, a flag to analyze data for only the first residue, or all residues, and a variable to set spectrometer frequency. The notebook uses the Solve command to solve the reduced spectral density equations for R_1, R_2, and NOE in terms of $J(0)$, $J(\omega_N)$, and $J(0.87\ \omega_H)$. These equations are then used to calculate reduced spectral density values on a per residue basis.

The notebook creates per residue plots of R_1, R_2, and $\{^1H\}$–^{15}N NOE, and their errors. A list of residue, $J(0)$, $J(\omega_N)$, and $J(0.87\ \omega_H)$ is then tabulated. If Monte Carlo simulations are requested, a table of residue, $J(0)$, error, $J(\omega_N)$, error, $J(0.87\ \omega_H)$, error is also created. The notebook creates individual plots of $J(0)$, $J(\omega_N)$, and $J(0.87\ \omega_H)$ as a function of residue (Figure 7). If Monte Carlo simulations are conducted, the errors are also plotted. Finally, the notebook will output a text file of the spectral density analysis containing a list of residue, $J(0)$, error, $J(\omega_N)$, error, $J(0.87\ \omega_H)$, error. Errors are only included if Monte Carlo simulations are carried out.

Results

Analysis of $^{15}N - T_1$ and T_2

The notebooks for analysis of ^{15}N-T_1 and T_2 decays were tested on ^{15}N-T_1 data collected previously at a ^1H Larmor frequency of 600 MHz (Spyracopoulos et al., 2005). As previously reported, the average value for ^{15}N-T_1 is 772 ms, with an average error of 22 ms. The average T_1 is typical for a protein of 150 residues. For the χ^2 goodness-of-fit test, 61 residues from a total of 106 residues have χ^2 less than the critical χ^2 of 11.5 using a noise scale factor of 1.0, whereas for a noise scale factor of $\sqrt{2}$, 98 residues have $\chi^2 <$ 11.5, indicating that a two-parameter exponential decay is likely to adequately represent the relaxation data.

Ubiquitin ^{15}N relaxation test data

The notebooks for the inertia tensor, diffusion tensor, model-free analysis, and spectral density analysis were tested on ^{15}N-T_1, -T_2, and $\{^1H\}$–^{15}N NOE data collected in duplicate at a ^1H Larmor

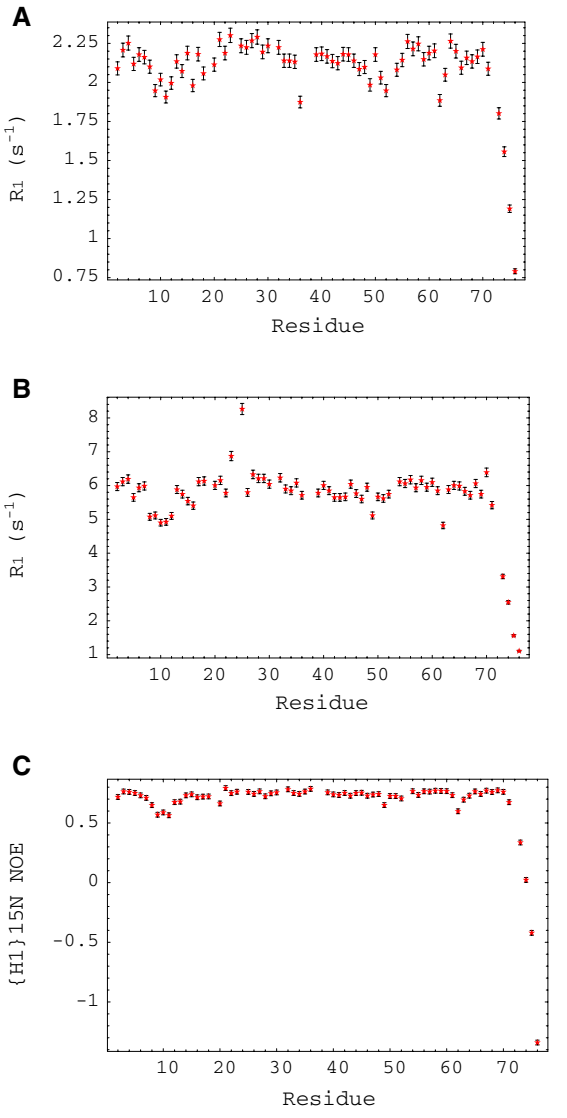

Figure 5. Graphical output generated within the notebook for model free analysis. ^{15}N-R_1 (A), -R_2 (B) and {^{1}H}–^{15}N NOE (C) for ubiquitin (BMRB accession number 6470) are displayed within the model-free notebook prior to numerical optimization.

frequency of 600 MHz by Bax and co-workers for the protein ubiquitin (BMRB accession number 6470, PDB 1UBQ) (Tjandra et al., 1995). Values for ^{15}N-T_1, -T_2, and {^{1}H}–^{15}N NOE data used in the notebooks were the average of the duplicate measurements. Per residue errors for ^{15}N-T_1 and -T_2 were set to $\pm 2\%$, and per residue NOE errors were assumed to be ± 0.02. For analysis of rotational diffusion anisotropy, main chain amide protons were added to 1UBQ using the xleap utility and the *ff03* force field from the Amber 8

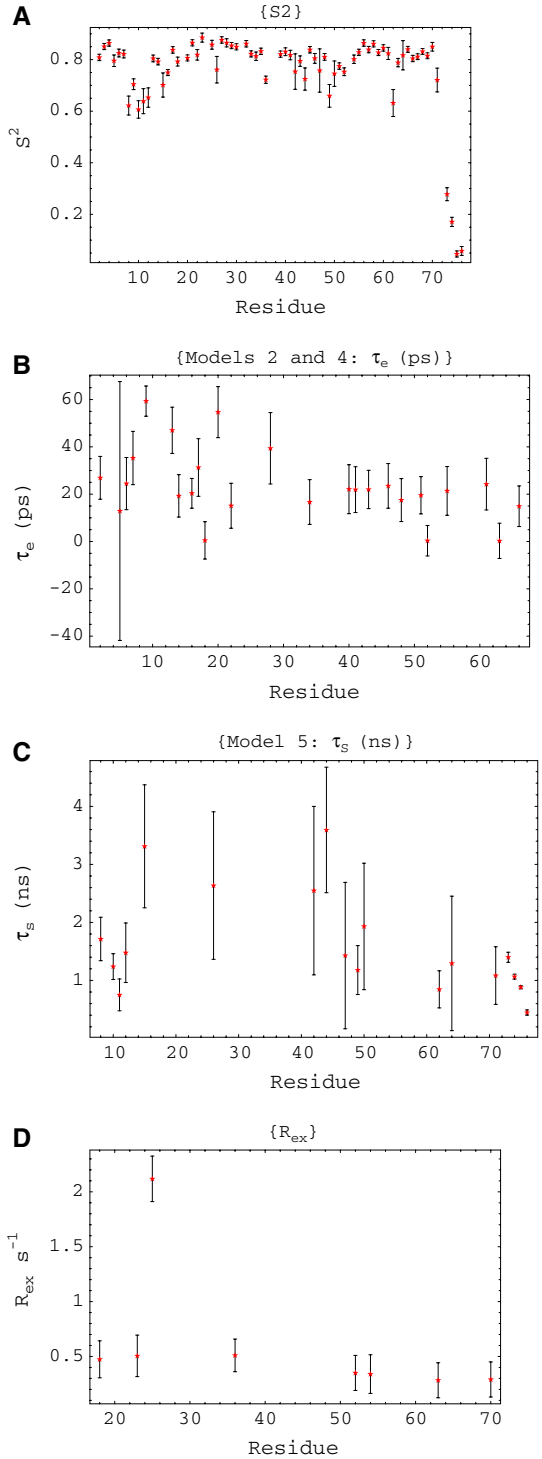

Figure 6. Graphical output for S^2, τ_f, τ_s, and R_{ex} parameters generated within the model-free notebook. Optimized model-free parameters S^2 (A), τ_f (B), τ_s (C), and R_{ex} (D) determined from ^{15}N-R_1, -R_2 and {^{1}H}–^{15}N NOE relaxation data for ubiquitin (BMRB accession number 6470).

suite of programs (Cornell et al., 1995, Case et al., 2005).

Inertia tensor

The values for the normalized principal components of the inertia tensor were calculated to be 1.0:0.90:0.64 using the coordinates from PDB 1UBQ, and are shown in Figure 3. These values are identical to those previously reported (Tjandra et al., 1995).

Isotropic correlation time, rotational diffusion tensor, and R_{ex} terms

For analysis of the rotational diffusion of ubiquitin, the notebook selected 55 residues on the basis of NOE and conformational exchange. This is one more residue than previously selected (Tjandra et al., 1995), and though the source of the discrepancy is not known, the difference does not have a large impact the analysis. The τ_c value for isotropic rotational diffusion was determined to be 4.09 ns, identical to the previously reported value. The AIC values were 85.9, 67.2, and 82.2 for the isotropic, axially symmetric prolate, and axially symmetric oblate rotational diffusion models. Thus, the data are consistent with an axially symmetric, prolate diffusion model, as previously reported, with $D_{||}/D_{\perp} = 1.14$, $1/6D = 4.13$ ns, $\theta = 41°$, and $\phi = 50°$. These values correspond closely to the previously determined values of $D_{||}/D_{\perp} = 1.17$, $1/6D = 4.12$ ns, $\theta = 40°$, and $\phi = 46°$ (Tjandra et al., 1995).

The notebook also estimated R_{ex} values of 1.3 ± 0.3, 2.5 ± 0.2, 0.7 ± 0.2, 0.6 ± 0.2, and 0.6 ± 0.2 s^{-1} for residues 23, 25, 27, 28, and 70, respectively from ubiquitin using the method of Bracken and co-workers (Kneller et al., 2002). The identification of these residues using this method compares favorably with the observation that residues 23, 25, 55, and 70 are involved in an exchange process, as characterized using ^{15}N NMR $R_{1\rho}$ relaxation measurements (Massi et al., 2005).

Model free analysis

The optimized parameters for the model-free analysis of previously reported ^{15}N relaxation data for ubiquitin are shown in Figure 6 (BMRB accession code 6470 and PDB 1UBQ). S^2 and τ_f values for ubiquitin in the presence of axially symmetric prolate rotational diffusion for model 2 (S^2, τ_f) are also included in BMRB deposition 6470, and a comparison (ΔS^2) between BMRB 6470 S^2 values and those calculated using the *Mathematica* notebook described herein is shown in Figure 8. In general, ΔS^2 values are not statistically different from zero, except for residues whose relaxation data were fit by models 3 (S^2, R_{ex}) and 5 (S^2, τ_s) in the current notebook.

Reduced spectral density analysis

Results for the reduced spectral density analysis using published ^{15}N relaxation data from ubiquitin are shown in Figure 7. The average values of $J(0)$, $J(\omega_N)$, and $J(0.87\ \omega_H)$ are 1.7, 0.5, and 0.01 ns rad^{-1}, respectively. These values are comparable to those of the 70-residue protein eglin *c*, determined using the original spectral density mapping approach (Peng and Wagner, 1992), for which the average values for the core residues are $J(0) = 1.3$ ns rad^{-1}, $J(\omega_N) = 0.5$ ns rad^{-1}, $J(\omega_N + \omega_H) = 0.02$ ns rad^{-1}, $J(\omega_H) = 0.05$ ns rad^{-1}, and $J(\omega_N - \omega_H) = 0.08$ ns rad^{-1}.

Discussion

Though there are a number of computer programs for analysis of protein main chain ^{15}N relaxation data, several unique features for the integrated suite of *Mathematica* notebooks presented herein lend to their general utility. The *Mathematica* environment contains a unified symbolic and numerical computational engine, programming language, and graphics system, and importantly, is available for a wide variety of computer platforms.

The unified features of *Mathematica* were fully utilized in the suite of notebooks presented herein. For example, two parameter exponential decay functions and model free spectral density equations describing ^{15}N NMR relaxation are programmed as symbolic functions; numerical optimizations and statistical analyses are performed using standard *Mathematica* packages.

Data input and output in *Mathematica* is easily managed, the notebooks can read output from the 'rh' package of the NMR spectral analysis program Sparky to calculate T_1 and T_2 values, and

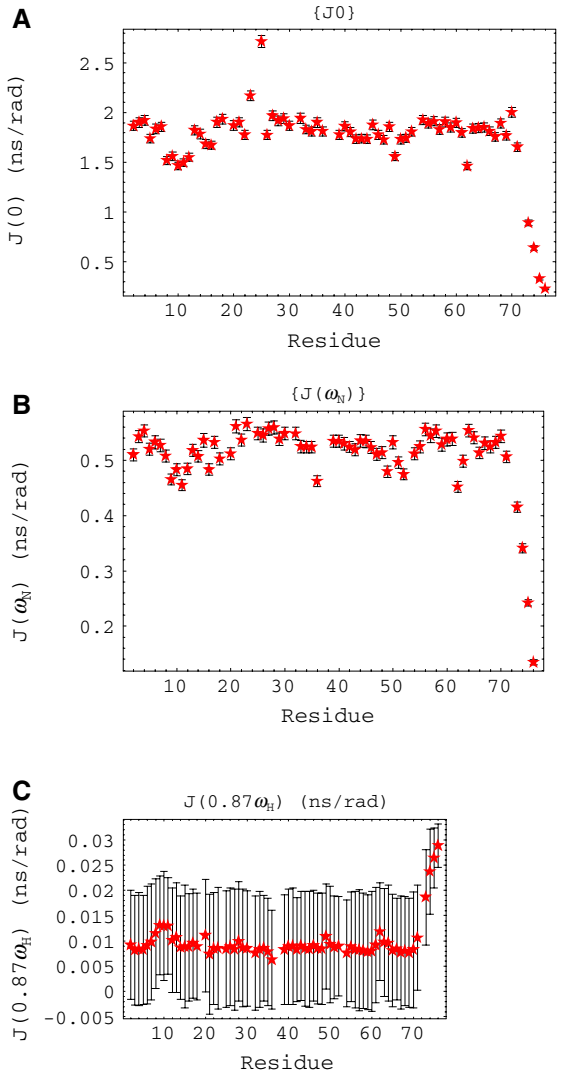

Figure 7. Graphical output for $J(0)$, $J(\omega_N)$, and $J(0.87\ \omega_H)$ generated within the spectral density analysis notebook. Reduced spectral density values $J(0)$ (A), $J(\omega_N)$ (B), and $J(0.87\ \omega_H)$ (C), determined from ^{15}N-R_1, -R_2 and $\{^1H\}$–^{15}N NOE relaxation data for ubiquitin (BMRB accession number 6470).

read PDB coordinate files to calculate the components of the inertia and diffusion tensors. Additionally, results from T_1 and T_2 curve fitting, calculation of direction cosines, results from model-free analysis, and results from reduced spectral density analysis are exported as text files.

The *Mathematica* graphics system is a key factor in the utility of the notebooks. Visualization of input and results generated automatically within a notebook greatly facilitates analysis and interpretation by the user. Input and output such as T_1 and T_2 peak intensity decays and their associated errors and curve fits are plotted, as well as statistical analysis of the χ^2 distribution, allowing for rapid visual assessment of the quality of the relaxation decays. This assessment is critical given the impact of signal to noise ratio on model-free analyses. For model-free and spectral density analysis notebooks, input data such as T_1, T_2, and NOE are easily assessed as they are plotted as functions of residue number within notebooks. The components of the inertia tensor, and a Cα trace through the protein backbone are plotted in three dimensions and can be rotated in real time. Optimized parameters from model free analyses, and values of the reduced spectral density are plotted as functions of residue number. These graphics can be directly incorporated into publications, or easily imported as vector graphics in the PostScript language into a graphics program for aesthetic enhancement.

The fidelity of the inertia tensor, diffusion tensor, and model-free notebooks was evaluated by analyzing previously published ^{15}N-T_1, -T_2, and $\{^1H\}$–^{15}N NOE relaxation data acquired at 600 MHz by Bax and co-workers (reference (Tjandra et al., 1995) and BMRB accession number 6470). The notebooks presented herein were found to accurately reproduce previously published results. Additionally, the diffusion tensor notebook was used to estimate R_{ex} parameters from R_1R_2 (Kneller et al., 2002). Residues 23 and 25 displayed the largest R_{ex} parameters, with values exceeding 1.0 s^{-1}. These residues also display the largest absolute chemical shift changes for the underlying chemical exchange process in ubiquitin, as determined through ^{15}N $R_{1\rho}$ measurements (Massi et al., 2005). Finally, the spectral density analysis notebook was found to calculate $J(\omega)$ values that closely match those from eglin *c*, a 70 residue protein that is similar in size to ubiquitin.

The *Mathematica* notebooks described herein represent an integrated, easy to use package for analysis of main chain amide ^{15}N relaxation data in proteins. The package has true cross-platform computer compatibility, produces high quality graphics, and demonstrates fidelity with respect to previously published results. In addition, the *Mathematica* code in the notebooks is transparent, allowing the user the possibility of adapting the notebooks to a specific function, or modifying the

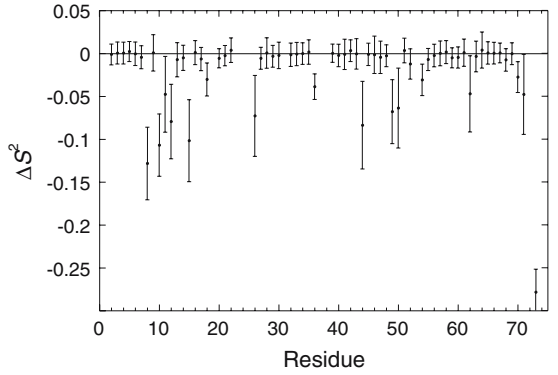

Figure 8. Comparison of ubiquitin model-free S^2 values. S^2 values for ubiquitin calculated with model 2 and axially symmetric prolate diffusion (BMRB accession number 6470) are compared to S^2 values determined using the model-free notebook presented herein. The value of the nitrogen chemical shift anisotropy was set to −160 ppm to facilitate comparison to the values reported in the BMRB. Values of ΔS^2 that deviate from zero include residues whose relaxation data were fit with different models (model 3: residues 36, 54, 70, and model 5: residues 8, 10–12, 15, 26, 44,49, 50, 62, 71, and 73).

notebooks for instructional purposes. Therefore, it is expected that this collection of *Mathematica* notebooks will be generally useful. The notebooks are available for download at www.bionmr.ualberta.ca/∼lspy.

Acknowledgments

This research was supported by grants from the Canadian Institutes of Health Research (CIHR) and the Alberta Heritage Foundation for Medical Research (AHFMR). L.S. is an AHFMR Medical Research Senior Scholar. The author thanks Matt Crump (University of Bristol) for assistance with the spectral density mapping notebook. The *Mathematica* code for reading PDB coordinates was kindly provided by Brian Higgins (www.higgins.ucdavis.edu).

References

Akke, M., Bruschweiler, R. and Palmer, A.G. (1993) *J. Am. Chem. Soc.*, **115**, 9832–9833.

Barbato, G., Ikura, M., Kay, L.E., Pastor, R.W. and Bax, A. (1992) *Biochemistry*, **31**, 5269–5278.

Blackledge, M., Cordier, F., Dosset, P. and Marion, D. (1998) *J. Am. Chem. Soc.*, **120**, 4538–4539.

Cantor, C.R. and Schimmel, P.R. (1980) *Techniques for the Study of Biological Structure and Function*, W. H. Freeman, San Francisco.

Case, D.A., Cheatham, T.E., Darden, T., Gohlke, H., Luo, R., Merz, K.M., Onufriev, A., Simmerling, C., Wang, B. and Woods, R.J. (2005) *J. Comput. Chem.*, **26**, 1668–1688.

Clore, G.M., Gronenborn, A.M., Szabo, A. and Tjandra, N. (1998) *J. Am. Chem. Soc.*, **120**, 4889–4890.

Clore, G.M., Szabo, A., Bax, A., Kay, L.E., Driscoll, P.C. and Gronenborn, A.M. (1990) *J. Am. Chem. Soc.*, **112**, 4989–4991.

Cordier, F., Caffrey, M., Brutscher, B., Cusanovich, M.A., Marion, D. and Blackledge, M. (1998) *J. Mol. Biol.*, **281**, 341–361.

Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W. and Kollman, P.A. (1995) *J. Am. Chem. Soc.*, **117**, 5179–5197.

d'Auvergne, E.J. and Gooley, P.R. (2003) *J. Biomol. NMR*, **25**, 25–39.

de la Torre, J.G., Huertas, M.L. and Carrasco, B. (2000) *J. Magn. Reson.*, **147**, 138–146.

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.W., Szabo, A., Torchia, D.A. and Kay, L.E. (1995) *J. Biomol. NMR*, **6**, 153–162.

Foote, J. and Raman, A. (2000) *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 978–983.

Ghose, R., Fushman, D. and Cowburn, D. (2001) *J. Magn. Reson.*, **149**, 204–217.

Idiyatullin, D., Daragan, V.A. and Mayo, K.H. (2003) *J. Phys. Chem. B*, **107**, 2602–2609.

Jarymowycz, V.A. and Stone, M.J. (2006) *Chem. Reviews*, **106**, 1624–1671.

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

Kneller, J.M., Lu, M. and Bracken, C. (2002) *J. Am. Chem. Soc.*, **124**, 1852–1853.

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, A.G. (1995) *J. Mol. Biol.*, **246**, 144–163.

Massi, F., Grey, M.J. and Palmer, A.G. (2005) *Protein Sci.*, **14**, 735–742.

Nicholson, L.K., Yamazaki, T., Torchia, D.A., Grzesiek, S., Bax, A., Stahl, S.J., Kaufman, J.D., Wingfield, P.T., Lam, P.Y.S., Jadhav, P.K., Hodge, C.N., Domaille, P.J. and Chang, C.H. (1995) *Nat. Struct. Biol.*, **2**, 274–280.

Orekhov, V.Y., Nolde, D.E., Golovanov, A.P., Korzhnev, D.M. and Arseniev, A.S. (1995) *Appl. Magn. Reson.*, **9**, 581–588.

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

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

Spyracopoulos, L., Lewis, M.J. and Saltibus, L.F. (2005) *Biochemistry*, **44**, 8770–8781.

Tjandra, N., Feller, S.E., Pastor, R.W. and Bax, A. (1995) *J. Am. Chem. Soc.*, **117**, 12562–12566.

Woessner, D.E. (1962) *J. Chem. Phys.*, **37**, 647–654.

Yang, D.W. and Kay, L.E. (1996) *J. Mol. Biol.*, **263**, 369–382.