An algebraic solution for determining overall rotational correlation times from cross-correlated relaxation rates

Scott A. Robson and Joshua J. Ziarek

Department of Molecular and Cellular Biochemistry, Indiana University, 212 S. Hawthorne Ave, Bloomington, IN 47405

Keywords
Model-free, order parameters, TRACT, NMR

Summary:

An accurate rotational correlation time is critical for quantitative analysis of fast timescale NMR dynamics. As molecular weights increase, the classic derivation using transverse and longitudinal relaxation rates becomes increasingly unsuitable due to the non-trivial contribution of dipole-dipole and chemical exchange processes. Derivation using cross-correlated relaxation experiments, such as TRACT, overcome these limitations but are erroneously calculated in at approximately 50% of the citing literature. The goals of this study are to 1) investigate the potential sources of the error, 2) provide an algebraic solution, and 3) highlight that future inaccuracies can be minimized by requiring publication of sufficient raw data and computational routes for re-evaluation.

Introduction:

A particle's rotational Brownian diffusion is characterized by the average time to rotate one radian aka the rotational correlation time (τ_c). It is related to the size and shape of a molecule, and in the case of a rigid, spherical particle, can be estimated from the Stokes-Einstein relation[1]. The rotational correlation time is frequently used in biophysics to gauge molecular aggregation and solvent viscosity, however it's particularly prevalent in solution NMR through its relation to longitudinal (R_1) and transverse (R_2) relaxation rates (Eq. 1, where ν_N is the frequency of ^{15}N in Hz)[2]. Rotational correlation time estimates are used to optimize interscan recycling delays, magnetization transfer delays in correlation experiments, and indirect dimension evolution times in multidimensional experiments[3]. Perhaps most significantly, τ_c is the critical parameter for any quantitative dynamics analyses in which separation of overall and internal motion are required such as 'model-free' formalism[4,5].

There are several important caveats to the determination of rotational correlation times from NMR relaxation data. Firstly, there is a circular logic problem in that experimental measurements of ^{15}N R_1 and R_2 are often used to estimate τ_c using Equation 1[2,6]. In this model, measured R_2 and R_1 rates are assumed to result from only the auto-relaxation rate. Under this assumption the τ_c value can then be used to calculate relaxation rates for other nuclei where it may prove useful. However, ^{15}N R_2 and R_1 relaxation rates may have significant contributions from additional dipole interactions with other nearby ^1H atoms, an effect that becomes significant when τ_c itself becomes large[7]. Chemical exchange also adds to measured relaxation rates[8] as nuclei become de-phased due to stochastically exchanging between two or more local magnetic fields at fixed rates. As a result, relaxation rates are often a complicated function of not only local atomic geometry and τ_c, but also chemical exchange rates and dipole interactions with covalently remote but spatially close ^1H nuclei, the latter being more and more relevant as molecular dynamics data is collected on higher and higher molecular weight systems.

$$\tau_c = \frac{1}{4\pi\nu_N}\sqrt{6\frac{R_2}{R_1} - 7} \quad (1)$$

Measurement of relaxation mechanisms independent of remote dipolar couplings and chemical exchange greatly simplify this problem. Transverse cross-correlated relaxation (CCR, η_{xy}) results from the coordinated rotation of two nuclei in a magnetic field[9] and is influenced only by the internuclear dipole-dipole (DD) coupling, each nuclei's chemical shift anisotropy (CSA), and the molecular rotational correlation time. The sign of $\pm\eta_{xy}$ depends on the spin state of the coupled nucleus. Cross-correlated relaxation only contributes to R_2 and R_1 in coupled systems because the opposing signs would cancel during decoupling[9]. The TROSY effect exploits this property by selecting signals only from the spin state with relaxation interference in ^{1}H-^{15}N spin systems[10], ^{1}H-^{13}C aromatic spin systems[11], or ^{13}C methyl groups[12]. There are a few ways to measure η_{xy} in the ^{1}H-^{15}N spin system. The most common approach is a set of two experiments that record the differential relaxation rates ^{15}N alpha (R_α) and beta (R_β) spin states. These rates are the sum of the auto-relaxation rate (R_{auto}), remote ^{1}H dipole interactions (R_D) , chemical exchange (R_{ex}) and η_{xy} (Eqs. 2 and 3); η_{xy} is extracted from subtraction of R_α and $R_\beta\eta_{xy}$[7,13,14] (Eqs. 4 and 5). It is also possible to directly measure DD/CSA CCR in constant-time experiments[15], but spectral interpretation is complicated by movement in the ^{15}N dimension of up-field TROSY peaks towards the downfield anti-TROSY position.

$$R_\alpha = R_{auto} + R_D + R_{ex} - \eta_{xy} \quad (2)$$
$$R_\beta = R_{auto} + R_D + R_{ex} + \eta_{xy} \quad (3)$$

$$R_\beta - R_\alpha = 2\eta_{xy} \quad (4)$$

$$\eta_{xy} = \frac{R_\beta - R_\alpha}{2} \quad (5)$$

Goldman[9] showed that τ_c can be calculated from η_{xy} for a given magnetic field and internuclear bond vector (Eq. 6).

$$\eta_{xy} = p\delta_N\big(4J(0) + 3J(\omega_N)\big)(3cos^2\theta - 1) \quad (6)$$

where:

$$p = \frac{\mu_0\gamma_H\gamma_N h}{16\pi^2\sqrt{2}r^3} \quad (7)$$

$$\delta_N = \frac{\gamma_N B_0 \Delta\delta_N}{3\sqrt{2}} \quad (8)$$

$$J(\omega) = \frac{2\tau_c}{5[1+(\tau_c\omega)^2]} \quad (9)$$

and:

$$h = 6.62607004 \cdot 10^{-34}\, J \cdot s$$

$\mu_0 = 1.25663706 \cdot 10^{-6}\ H\ \cdot m^{-1}$
$\gamma_H = 267.52218744 \cdot 10^6\ radians \cdot s^{-1} \cdot T^{-1}$
$\gamma_N = -27.116 \cdot 10^6\ radians \cdot s^{-1} \cdot T^{-1}$
$r\ =\ 1.02\ \text{Å}\ =\ 1.02 \cdot 10^{-10}\ m$
$\Delta\delta_N = 160\ ppm = 160 \cdot 10^{-6}$
$\theta = 17\ ^\circ = \frac{17\pi}{180}\ radians$

The relationship between η_{xy} and τ_c was first exploited experimentally by Lee and colleagues in the highly-cited [^{15}N,^{1}H]-TRACT (TROSY for rotational correlation times) pulse sequence[7], but the manuscript does not explicitly detail how the Goldman equation is used to solve for τ_c from η_{xy} and magnetic field strength using Equations 6-9. Despite what appears to be a relatively straightforward mathematical process, we noted inconsistent τ_c calculations in the original paper while analyzing our own TRACT data. This discrepancy appears to us to be too large to be the simple result of rounding errors or slightly different physical and geometric constants. We wondered whether this error was commonplace in calculations of τ_c using cross-correlated relaxation data. To examine this we determined an analytical, algebraic solution to Equation 6, which we used to conduct a survey of literature siting the original TRACT paper and their calculations. We found a significant number of similar errors in the body of literature that site the original TRACT paper. We also found a significant number of results that agree with our algebraic solution. We discuss how the original error may have propagated into the literature and propose some ideas on how to avoid this sort of problem in the future.

Results and Discussion:

The errors in the original TRACT paper appear to overestimate τ_c by approximately 6.6%. Specifically, Figures 3 and 4 in the original TRACT paper give a field strength of 750 MHz and τ_c = 21 ns and 24 ns, respectively, from which we calculate η_{xy} rates of 27.1 Hz and 30.9 Hz using Equations 6-9. This is inconsistent with the reported R_α and R_β which yield η_{xy} rates of (64-13)/2 = 25.5 Hz and (80-22)/2 = 29 Hz, respectively, using Equations 4-5 and a field of 750 MHz. However, we did note that the error was approximately, but not exactly equal to, a difference in field strength between 750 and 700 MHz (an error of 7.1%) suggesting that calculations were performed using a field of 700 MHz. We discounted the idea that the field used in the calculations was 800 MHz (an error of 6.7%) because an increase in field should underestimate τ_c, not overestimate it. Simulating τ_c as a function of magnetic field (using Eq. 6; Fig. S1) illustrates that errors in field strength would scale almost but not quite linearly – supporting that the observed error is within expectations and suggesting a numerical or clerical oversight in the original manuscript. To ensure an error did not result from a computational problem with an optimization/minimization algorithm, we generated an exact solution to Equation 6 with respect to τ_c given magnetic field strength and η_{xy}.

We start by expanding Equation 6 with the spectral density function (J) and substituting η_{xy} with $(R_\beta - R_\alpha)/2$ which gives Equation 10:

$$4\frac{2\tau_c}{5} + 3\frac{2\tau_c}{5[1+(\tau_c\omega_N)^2]} = \frac{R_\beta - R_\alpha}{2p\delta_N(3cos^2\theta - 1)} \qquad (10)$$

The righthand side of this equation is a constant once a measurement has been made for the relaxation rates at a specific field. We therefore replace this side with the symbol ‘c’.

$$4\frac{2\tau_c}{5} + 3\frac{2\tau_c}{5[1+(\tau_c\omega_N)^2]} = c \tag{11}$$

To solve for τ_c we entered Equation 11 into Wolfram Alpha[16] and requested an algebraic solution for τ_c. The result is Equation 12 below.

$$\tau_c = \frac{\sqrt[3]{125c^3\omega_N^6+24\sqrt{3}\sqrt{625c^4\omega_N^{10}-3025c^2\omega_N^8+21952\omega_N^6}+1800c\omega_N^4}}{(24\omega_N^2)} - \frac{(336\omega_N^2-25c^2\omega_N^4)}{24\omega_N^2\sqrt[3]{125c^3\omega_N^6+24\sqrt{3}\sqrt{625c^4\omega_N^{10}-3025c^2\omega_N^8+21952\omega_N^6}+1800c\omega_N^4}} + \frac{5c}{24} \tag{12}$$

where,

$$c = \frac{R_\beta - R_\alpha}{2p\delta_N(3\cos^2\theta-1)} = \frac{\eta_{xy}}{p\delta_N(3\cos^2\theta-1)} \tag{13}$$

And ω_N is the frequency of the ^{15}N nucleus in radians per second.

Using this algebraic solution, we confirmed that substituting a field of 700 MHz reproduced the reported τ_c values in the original TRACT paper for the given relaxation rates. However, we cannot conclude whether the error was the result of quoting an incorrect field in the manuscript or using an incorrect field in their calculations. Our analysis assumes that the two, independent cross-correlated relaxation rates reported in the original manuscript were accurately used in calculations. This is more plausible than both relaxation rates were incorrectly used in calculations such that they appear to be from a magnet at 700 MHz rather than the quoted 750 MHz.

In addition to the miscalculation detailed above we note that while the mathematics behind Equation 6 are simple in principle, there are multiple points where an inexperienced researcher may make an error or attempt to simplify. We wondered whether this oversight may have led to other miscalculations by users of the TRACT experiment to determine τ_c from η_{xy}. As of late 2020, the original manuscript had been cited 120 times in PubMed. We surveyed each of these manuscripts looking for experimental data and calculated τ_c values to determine the accuracy of the community's calculations. Table 1 shows that out of 120 manuscripts, three referenced the original TRACT paper without performing any cross-corelated relaxation experiments and an additional six were reviews of NMR methodologies. Seventy-eight used the TRACT methodology but did not provide enough information to verify their τ_c calculation. Excluding those papers that did not present τ_c values, only 30% (33:111 papers) reported sufficient data to verify their results (including supplementary information). In these thirty-three papers we found a total of sixty-five τ_c calculations. Further, 50% of all citing publications appeared in the years 2015-2020, indicating in increasing interest in this methodology 10-15 years after its' original publication.

We derived an error measure by dividing the published τ_c value by our algebraically determined value. For example, the original paper reported τ_c = 21 and 24 ns, while we calculated values of 19.7 and 22.5 ns, giving error ratios of 1.066 and 1.067, respectively. A histogram of our total analysis is given in Figure 1A. There are several things to note from this histogram. Firstly, there is a significant cluster of results that are accurate (ratio ≈ 1.0) with 35% (23/65) of results clustering to within a 2% error interval of 0.99 and 1.01. This supports the accuracy of our exact solution and suggests that others have previously established similar approaches. Secondly, 23% (15/65) of results cluster around 1.067 ± 0.03; this strongly suggests the presence of a systematic error in the NMR literature that was propagated from the original TRACT paper. Finally, we note that 29% (19/65) of calculations have error ratios greater than 1.1 (>10% error). Extrapolating our results to all citing literature implies that ~65% (over 70 calculations) incorrectly estimate the rotational correlation time using the TRACT methodology, and that errors in the literature have so far not started to self-correct with time (Fig. 1B).

How could this happen? A possible explanation is that researchers have taken the erroneous results from the original paper to estimate solutions. Equation 6 can be restated as Equation 14, which reduces the bulk of the physical constants to a single 'k' value:

$$\eta_{xy} = k\,B_0\,(3\cos^2\theta - 1)\left(4J\left(B_0, 0, \tau_c\right) + 3J(B_0, \gamma_N, \tau_c)\right) \tag{14}$$

where,

$$k = \frac{\mu_0 \gamma_H \gamma_N h}{16\pi^2\sqrt{2}r^3} \cdot \frac{\gamma_N \Delta\delta_N}{3\sqrt{2}} \tag{15}$$

Numerically equating the left- and right-hand sides of Equation 14 with values from the original TRACT paper generates a value for 'k' that can be used for subsequent new calculations; that is, one does not need to work with the detailed computations that involve physical constants in Equation 15. Once relaxation rates and η_{xy} are determined, the calculation simply involves numerically finding values for τ_c that equate the left- and right-hand sides of Equation 14. Such a method would usually involve the use of numerical optimization/minimization algorithms. This approach would explain a large number of errors uncovered in our literature survey that approximate the original error of 6.6% (Fig. 1A). We note that this approach, even if calibrated on correct parameters, would fail to estimate rotational correlation times that deviate from those in the original TRACT paper because η_{xy} is not a linear function of τ_c (Fig. S2). This may explain the larger variance of results around the 1.067 versus 1.00 error ratio. Additionally, η_{xy} is not a linear function of magnetic field strength, so 'k' calculated at one field strength would not be accurate at another (Fig. S1).

We also noted that researchers are more likely to overestimate the rotational correlation time, especially for low τ_c values. A scatter plot of τ_c versus error ratio demonstrates an inverse trend with highly erroneous values when τ_c < 2 ns (Fig. 2). Figure S2 suggests that the approach of Equation 14 to solve for τ_c would especially fail in this range of τ_c values as the relationship between τ_c and η_{xy} is exceptionally non-linear in this region for typical NMR field strengths.

Finally, we tested the robustness of an optimization approach to this problem, by exploring minimization results using several common algorithms. We reposed Equation 10 so all variables can be used to optimize (minimize) a numerical expression, in this case the LHS of Equation 16:

$$\left|4\frac{2\tau_c}{5} + 3\frac{2\tau_c}{5[1+(\tau_c\omega_N)^2]} - \frac{R_\beta - R_\alpha}{2p\delta_N(3cos^2\theta - 1)}\right| = 0 \tag{16}$$

Although a number of the 'off the shelf' optimization algorithms available in the NumPy library[17] give results that are essentially the same as our algebraic solution, some were significantly different and/or failed to converge (Table 1; Listing S2). It is clear that optimization methods can't always be trusted to give accurate results even for this relatively simple problem.

Conclusion:

We believe we have found that a flawed approach, similar to Equation 14, has become a common method to calculate rotational correlation times from cross-correlated relaxation data. This apparent approach was particular conspicuous as the original TRACT work contains a clerical or calculation error, but it further fails because of underlying assumptions about linear relationships between cross-correlated relaxation, rotational correlation time, and field strength. The problems discovered above are especially significant when calculating τ_c values in the low nanosecond range but also impact calculations of τ_c values in any range. This problem is not without potential consequences for dynamics analysis of molecules by NMR. Correlation time plays a significant role in the analysis of fast motions by the model-free formalism of Lipari and Szabo. Several studies have shown that incorrect τ_c values impact this analysis[6,18-20]. Clearly, errors in measurements would lead to errors in molecular parameters and therefore misinterpretation of molecular behaviors. In this work we provide an algebraic solution to the Goldman equation (Eq. 6), which provides a more straightforward and easier to implement method for calculating τ_c from η_{xy} and magnetic field strength.

While preparing this manuscript we could not find a detailed discussion of the presence or impact of computational errors in deriving molecular parameters from NMR data. It goes without saying that inaccurate results can substantially impact the direction of future research. Declaring numerical errors as trivial has the potential to lead to substantial wastes of time and resources. Given the breadth and complexity of typical dynamic NMR (model-free, relaxation dispersion) analyses, we feel it is timely to discuss of the importance of validating the accuracy of these calculations. We propose that the following should be expected to take place when publishing numerical results from NMR data.

1) A complete record of all numerical points used in calculations be made available at the time of submission of a manuscript to reviewers.

2) Computer code for calculations should be made available at review time so reviewers can verify calculations and robustness of the code. In accord with this, already published methods with poor documentation on how data is processed need to be made more explicit with open source, highly readable computer code.

3) Preferably computational code should be in a language that is as human-readable as possible (Python, Mathematica or MATLAB for example) to enable non-expert coders to follow the steps of calculation. Source code should be published along with the manuscript or linked to at an open source repository such as GitHub.

Accurate conclusions require accurate calculations. Our results highlight the importance of attention to detail when performing these calculations; and unfortunately, suggest that errors are easily propagated into the published literature. We hope this work will draw the necessary attention for prevention of future errors.

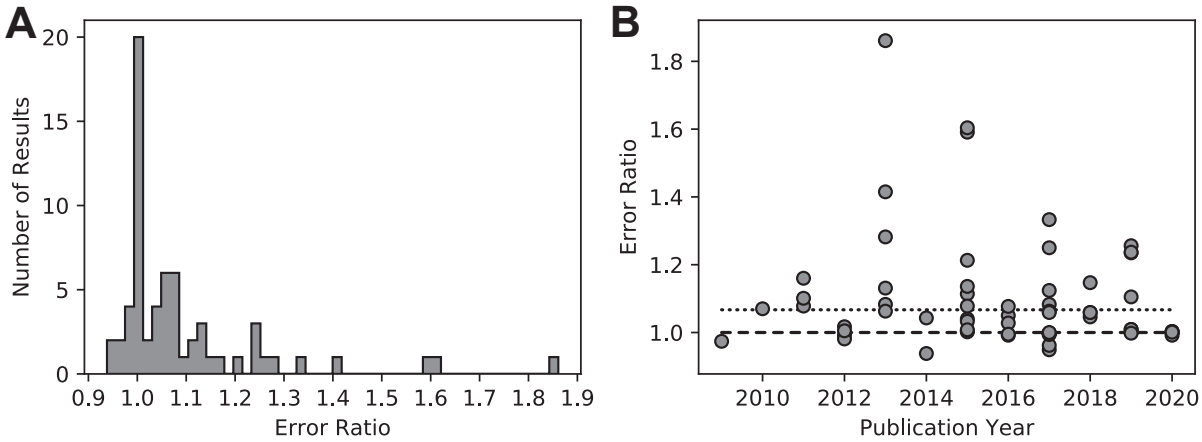

Figure 1: Analysis of errors in τ_c calculations that appear in published papers. A) Histogram of error ratios for τ_c. There are two significant clusters of errors; one is narrowly distributed around the accurate result (1.0); a second cluster is close the error (1.067) that appears in the TRACT paper with a slightly wider distribution. There are also a number of significantly more inaccurate results that tend to be overestimates. B) Scatter plot of errors versus year of publication, showing that there is no clear trend showing that errors in τ_c calculations are diminishing. The dashed line (-) indicates where accurate calculations (error ratio of 1.0) should be. The dotted line (…) indicates an error ratio of 1.067.

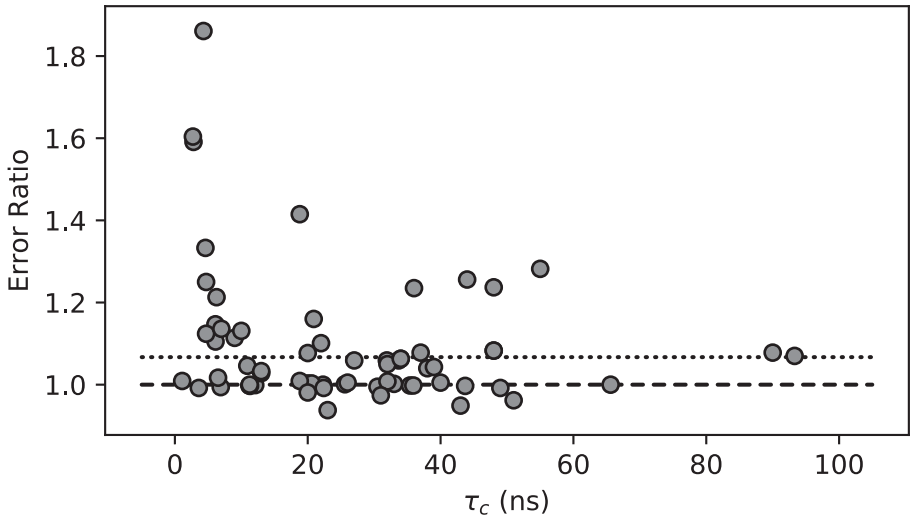

Figure 2: A scatter plot of τ_c versus error ratio demonstrating there is generally no trend between τ_c and error ratio, apart from a group of large errors for small τ_c values. The dashed line (-) indicates where an error ratio of 1.0 is. The dotted line (…) indicates an error ratio of 1.067 lies.

Table 1: Table of published manuscripts that site the original TRACT paper broken down into four categories: Papers that reference the TRACT paper but provided no τ_c calculations; reviews of NMR methods that included the TRACT experiment; papers that determined $\tau_c$with the TRACT method but did not provide enough data for τ_c verification; papers that did provide enough data for verification of τ_c calculations.

Paper Category	Number of Articles	Percentage of Articles (%)
Only Referenced	3	2.5
Reviews	6	5.0
Insufficient Data for Analysis	78	65.0
Sufficient Data for Analysis	33	27.5

Table 2: Calculations of τ_c using Algebraic and Numerical Minimization Methods with Equation 6 (Algebraic) or Equation 16 (Minimization) and data from the original TRACT paper.

Panel	Method	Convergence	τ_c (ns) 750 MHz	τ_c (ns) 700 MHz
A	Algebraic	N/A	19.746	21.156
A	Brent Minimization	Yes	19.749	21.154
A	BFGS Minimization	No	13.220	14.916
A	Powell Minimization	Yes	19.746	21.153
A	TNC Minimization	No	19.670	21.114
B	Algebraic	N/A	22.498	24.105
B	Brent Minimization	Yes	22.502	24.106
B	BFGS Minimization	No	16.455	18.341
B	Powell Minimization	Yes	22.498	24.105
B	TNC Minimization	No	22.436	24.049

References

1. Einstein, A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik* **322**, 549-560 (1905).
2. Kay, L.E., Torchia, D.A. & Bax, A. Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972-9 (1989).
3. Rovnyak, D., Hoch, J.C., Stern, A.S. & Wagner, G. Resolution and sensitivity of high field nuclear magnetic resonance spectroscopy. *J Biomol NMR* **30**, 1-10 (2004).
4. Lipari, G. & Szabo, A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *Journal of the American Chemical Society* **104**, 4546-4559 (1982).
5. Lipari, G. & Szabo, A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *Journal of the American Chemical Society* **104**, 4559-4570 (1982).
6. Lee, A.L. & Wand, A.J. Assessing potential bias in the determination of rotational correlation times of proteins by NMR relaxation. *Journal of Biomolecular NMR* **13**, 101-112 (1999).
7. Lee, D., Hilty, C., Wider, G. & Wuthrich, K. Effective rotational correlation times of proteins from NMR relaxation interference. *J Magn Reson* **178**, 72-6 (2006).
8. McConnell, H.M. Reaction Rates by Nuclear Magnetic Resonance. *The Journal of Chemical Physics* **28**, 430-431 (1958).
9. Goldman, M. Interference effects in the relaxation of a pair of unlike spin-12 nuclei. *Journal of Magnetic Resonance (1969)* **60**, 437-452 (1984).
10. Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. Attenuated T₂ relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences* **94**, 12366 (1997).
11. Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. Transverse Relaxation-Optimized Spectroscopy (TROSY) for NMR Studies of Aromatic Spin Systems in 13C-Labeled Proteins. *Journal of the American Chemical Society* **120**, 6394-6400 (1998).
12. Tugarinov, V., Sprangers, R. & Kay, L.E. Line Narrowing in Methyl-TROSY Using Zero-Quantum 1H-13C NMR Spectroscopy. *Journal of the American Chemical Society* **126**, 4921-4925 (2004).
13. Vasos, P.R., Hall, J.B. & Fushman, D. Spin-state selection for increased confidence in cross-correlation rates measurements. *J Biomol NMR* **31**, 149-54 (2005).
14. Hall, J.B., Dayie, K.T. & Fushman, D. Direct measurement of the 15N CSA/dipolar relaxation interference from coupled HSQC spectra. *J Biomol NMR* **26**, 181-6 (2003).
15. Liu, Y. & Prestegard, J.H. Direct measurement of dipole-dipole/CSA cross-correlated relaxation by a constant-time experiment. *J Magn Reson* **193**, 23-31 (2008).
16. Wolfram|Alpha. (Wolfram Alpha LLC, 2009).
17. Harris, C.R. et al. Array programming with NumPy. *Nature* **585**, 357-362 (2020).

18. Korzhnev, D.M., Orekhov, V.Y. & Arseniev, A.S. Model-Free Approach beyond the Borders of Its Applicability. *Journal of Magnetic Resonance* **127**, 184-191 (1997).
19. Korchuganov, D.S. et al. Determination of protein rotational correlation time from NMR relaxation data at various solvent viscosities. *Journal of Biomolecular NMR* **30**, 431-442 (2004).
20. Jin, D., Figueirido, F., Montelione, G.T. & Levy, R.M. Impact of the Precision in NMR Relaxation Measurements on the Interpretation of Protein Dynamics. *Journal of the American Chemical Society* **119**, 6923-6924 (1997).