

An in-cell diffusion method to characterize the size, abundance and permeability of cells

Gogulan Karunanithy¹, Richard J. Wheeler², Louise R. Tear³, Nicola J. Farrer³,
Stephen Faulkner³ and Andrew J. Baldwin¹

1. Department of Chemistry, Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom.
2. Peter Medawar Building for Pathogen Research, Nuffield Department of Medicine, University of Oxford, Oxford OX1 3SY, United Kingdom. Present address: Sir William Dunn School of Pathology, University of Oxford OX1 3RE Oxford, United Kingdom.
3. Chemistry Research Laboratory, University of Oxford, 12 Mansfield Road, Oxford OX1 3TA, United Kingdom.

Correspondence to: andrew.baldwin@chem.ox.ac.uk

Abstract

NMR and MRI diffusion experiments contain information describing the shape, size, abundance, and membrane permeability of cells although extracting this information can be challenging. Here we present the INDIANA (IN-cell Diffusion ANALysis) method to simultaneously and non-invasively measure cell abundance, effective radius, permeability and intrinsic relaxation rates and diffusion coefficients within the inter- and intra- cellular populations. The method couples an experimental dataset comprising stimulated-echo diffusion measurements, varying both the gradient strength and the diffusion delay, together with software to fit a model based on the Kärger equations to robustly extract the relevant parameters. A detailed error analysis is presented by comparing the results from fitting simulated data from Monte Carlo simulations, establishing its effectiveness. We note that for parameters typical of mammalian cells the approach is particularly effective, and the shape of the underlying cells does not unduly affect the results. Finally, we demonstrate the performance of the experiment on systems

of suspended yeast and mammalian cells. The extracted parameters describing cell abundance, size, permeability and relaxation are independently validated.

Introduction

Magnetic resonance experiments provide a powerful method to non-invasively analyse molecules within cells and tissues¹⁻⁵, and diffusion methods are frequently employed⁶⁻⁸. Applications range from the medical, where diffusion based MRI measurements can be used diagnostically^{9,10}, through to characterising protein folding^{11,12} and localising biomolecules in cellular mixtures¹³. There continues to be significant development of diffusion methodology including the FEXSY method¹⁴, triple quantum experiments¹⁵ and improved diffusion tensor imaging methods¹⁶. For biological systems, the high abundance and slow relaxation of water make it an excellent probe. The restricted motion of water restrained within internal cavities of cells and tissues, the transport of water across permeable membranes via channels such as aquaporin¹⁷⁻¹⁹ and variations in intrinsic relaxation can all complicate the analysis of data (Fig. 1A).

In both NMR and MRI diffusion experiments, magnetisation is typically first encoded with a position dependent phase by application of a pulsed field gradient of strength g applied for duration δ . After a delay Δ , where magnetisation is held either transverse in a spin-echo (SE) or longitudinal in stimulated echo (STE) experiment, a decoding pulsed field gradient is applied. Signals originating from molecules that have moved away from their original position are attenuated. Where the particle of interest undergoes translational self-diffusion, signal will decay according to²⁰:

$$S(b) = S(0)e^{-bD} \quad (1)$$

where in the limit $\Delta > \delta$, $b = (\Delta - \delta/3)q^2$, $q = g\gamma\delta$ and γ is the gyromagnetic ratio of the nucleus under study.

The restrictive effects of a cavity on translational diffusion has been treated by Neuman²¹ under the assumption that the displacement probability distribution will remain Gaussian. This assumption breaks down at high gradients and in the long-time limit^{22,23}, where distinctive ‘diffusion-diffraction’ patterns are expected whose profile is dominated by shape of the cavity, not the local diffusion coefficient. The size of mono-disperse polystyrene spheres²⁴ has been measured using this method. When the sample under study is heterogeneous, as is typically the case for biological samples²², this effect is rarely observed experimentally.

Reflecting the need to combine both restricted diffusion and permeability, Kärger developed a model²⁵⁻²⁷ where water can exchange between an interior, and an exterior pool. The displacement probability distributions are implicitly assumed to be Gaussian and so Kärger equations are compatible with the results of Neuman.

The heterogeneity of human tissue means that the assumption of two pools can be inappropriate. Monte Carlo simulations have been used to examine the effects of water exchanging between 2 and 3 pool models with differing underlying geometries²⁸⁻³⁰ tailored to specific applications. These have been followed by applications to study the diffusive behaviour in perfused glial cells³¹, erythrocytes³²⁻³⁴, rat brain tissue³⁵, blood³⁶ and bovine optic nerve³⁷. More recently, a method has been proposed for analysing cancer tissue diffusion data that includes restricted diffusion but neglects exchange³⁸.

For in-cell magnetic resonance experiments, the geometry is less complex than inside living tissue. In such applications, it is desirable to have a single model

that can be applied to give useful information that is insensitive to shape of the cavity. Moreover, when applying such a model, it is important to have a reasonable expectation for the accuracy of fitted parameters obtained, and where the systematic errors are likely to be.

To address these challenges, here we present the INDIANA (IN-cell Diffusion ANALysis) experimental protocol and software package. Water diffusion data from an STE diffusion experiment (Fig 1B) is acquired, varying both g and Δ . The data is analysed by a model that treats the cells as spheres following the approach of Neuman²¹, and allows intra-cellular water to exchange with a freely moving exterior population using Kärger^{25–27} equations, that have been modified to include the effects of intrinsic relaxation. The outputs from the model are cell count (interior population), cell radius, permeability (cell interior to exterior exchange rate), as well as the intrinsic relaxation rates and diffusion coefficients from the interior and exterior populations.

Regions where we expect acceptable parameters are identified by fitting data simulated using Monte-Carlo models to our model (Fig 5). We demonstrate that accounting for a physically reasonable distribution of cavity sizes and the expected variation in gradient strengths in typical NMR spectrometers³⁹ in simulations is necessary to resemble real data, and effectively extend the region of validity of the approach. The accuracy of the fitting parameters will be dependent on the choice of g and Δ , but for regions of parameter space typical for mammalian cells, the accuracy is remarkably high. Moreover, the results are relatively insensitive to the specific geometry of the underlying cell.

Experimental data can be acquired in as little as 4 minutes and we show in applications to budding yeast (*Saccharomyces cerevisiae*), 3T3 and HeLa

mammalian cells the parameters we extract are consistent with those published previously (Fig 6). Overall, the experimental and analysis tools are well suited for in-cell NMR studies to quickly and efficiently determine properties of the underlying cells in a manner that is not overly dependent on the underlying geometry. Software for the INDIANA analysis is written in Python and can be freely downloaded.

Results

We first introduce an analysis that demonstrates the need to use a Kärger-type model when analysing water diffusion data. Temporarily neglecting relaxation, in the case where two freely diffusing, but non-interacting populations are present, A and B, signal decay will follow a bi-exponential form:

$$S(b) = S_A(0)e^{-bD_A} + S_B(0)e^{-bD_B}. \quad (2)$$

Datasets acquired with varying g and Δ will show identical attenuation at a particular b value, independent of Δ (Fig 1 Ci,ii). Data acquired on a sample of suspended 3T3 cells (Fig 1D) show more complex attenuation, demonstrating the need for more sophisticated analysis. Individual datasets at constant Δ with varying g decay typically take a bi-exponential form although at smaller values of b , signal decay can appear to be mono-exponential (Fig 1Ci). Restricted diffusion effects, cell permeability and differential relaxation are all expected to significantly affect the measured signal (Fig 2,3). Following the approach of Kärger, for a stimulated echo diffusion experiment, the differential equation for the attenuation of signal S from the interior (*in*) and external (*out*) populations due to diffusion, relaxation and exchange will be:

$$\frac{d}{dt} \begin{pmatrix} S_{in} \\ S_{out} \end{pmatrix} = \begin{pmatrix} -q^2 D_{in} - k_{IO} - R_1^{in} & k_{OI} \\ k_{IO} & -q^2 D_{out} - k_{OI} - R_1^{out} \end{pmatrix} \begin{pmatrix} S_{in} \\ S_{out} \end{pmatrix} \quad (3)$$

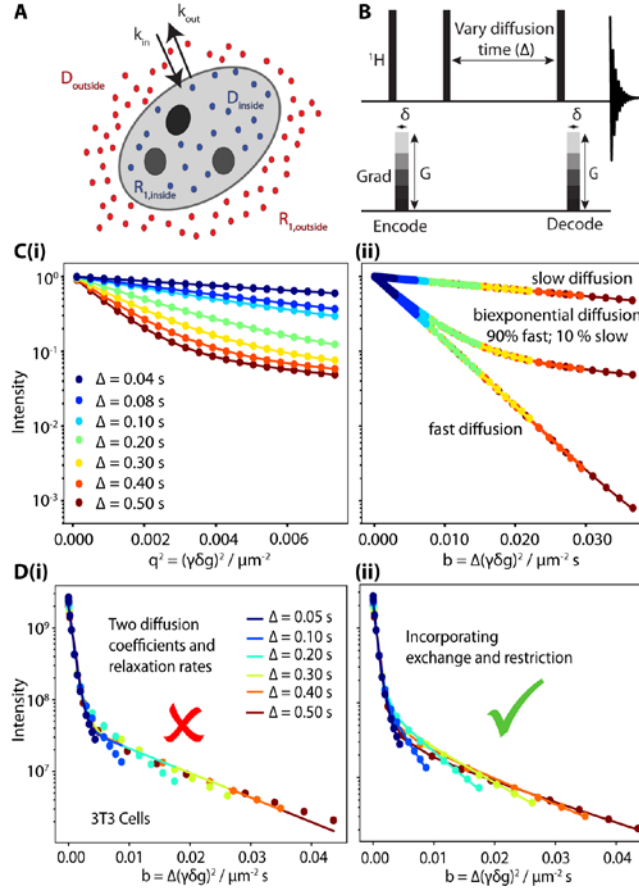


Figure 1 **A** Factors that need to be considered when analysing water diffusion experiments in cellular suspensions. The cell abundance, radius and permeability, as well as intra- and extra-cellular diffusion coefficients and relaxation rates will be significant. **B** STE pulse sequence used to measure diffusion. As magnetisation is stored longitudinally in this experiment, longer times can be experimentally accessed, that allows the cellular spaces to be more fully explored. **C(i)** Simulated signal intensities from diffusion experiments when the signal comprises two independent, non interacting components as a function of q^2 . The slow/fast diffusion coefficients were set to 200/2000 $\mu m^2 s^{-1}$ in a 1:9 population ratio. **C(ii)** the same data plotted as a function of b together with the cases where there is no exchange. In the cases where there is no exchange and constant diffusion coefficients, all curves show identical attenuation at the same b value. **D(i)** demonstrates the failure of a simple biexponential model to explain the diffusion data observed for a real system of mammalian 3T3 cells. The curves show significantly different attenuation depending on the

diffusion time at a particular b value, and there are significant systematic deviations between the data and the best fit. **D(ii)** Inclusion of exchange and restriction effects using the model set out in this paper (equations (8) and (10)) results in good fits to the data, and physically reasonable fitted parameters (Fig 5).

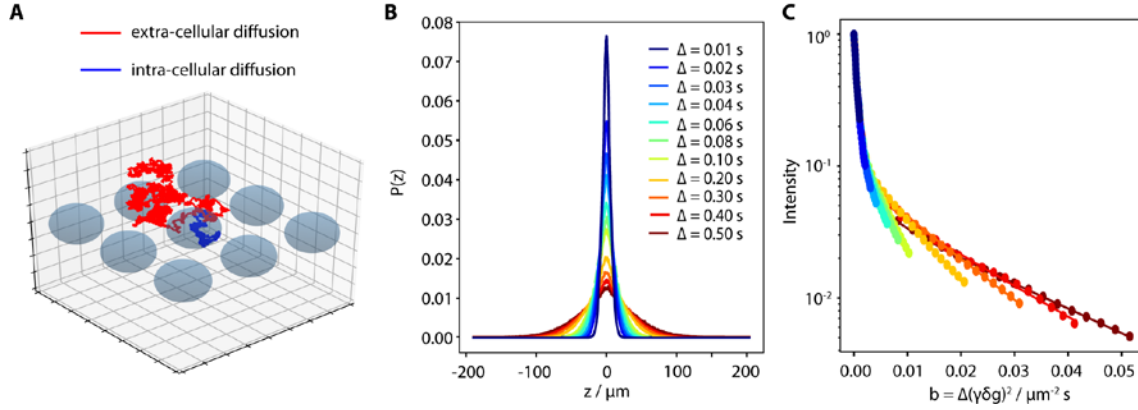


Figure 2 **A** Visual representation of a single Monte-Carlo single particle trajectory (Methods, A). The simulations account for distinct intra- and extra-cellular environments characterised by separate diffusion coefficients as well as exchange between the two environments and differential relaxation in the two phases. **B** A probability distribution of displacements for different diffusion times can be determined. **C** The trajectories can be combined to simulate experimental data (equation (21)).

Where D_{in} and D_{out} are the intra- and extra-cellular diffusion coefficients, k_{IO}/k_{OI} are the in-to-out/out-to-in exchange rates, and R_1^{in}/R_1^{out} are the intra-/extra-cellular longitudinal relaxation rates. For a spin-echo diffusion experiment, the relevant relaxation rates will be transverse (R_2) rather than longitudinal (R_1). Since the magnetisation is also held transverse during the period of gradient application, the final signal will also be affected by the transverse relaxation rate. However, in all experimental applications described here, this duration is held constant so its effect on the signal attenuation is uniform and can be excluded.

The differential equation is of the form $\frac{dS}{dt} = \rho S$, whose general solution is given by $S(q, \Delta) = e^{\rho(q)\Delta} S(0,0)$. The matrix exponential can be evaluated by constructing a matrix with Eigenvalues along the diagonal (D) together with a matrix of Eigenvalues (P). In the case where the chemical shift in the two compartments is identical such as for water in typical cellular samples, a detection operation that combines the signals is required:

$$S(q, \Delta) = P e^{D\Delta} P^{-1} \begin{pmatrix} p_{in} \\ p_{out} \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix} S(0,0) \quad (4)$$

where p_{in} and p_{out} and the equilibrium populations. The Eigenvalues of $-\rho$ are

$$\lambda^{\pm} = \frac{1}{2}(\xi^{\pm} \pm \phi), \quad (5)$$

expressed in terms of the following substitutions:

$$\begin{aligned} \xi^{\pm} &= q^2 D^{\pm} + R_1^{\pm} + k_{ex}(p_{out} \pm p_{in}), \\ \phi &= \sqrt{(\xi^{-})^2 + 4k_{ex}^2 p_{in} p_{out}} \quad \text{and} \\ D^{\pm} &= D_{in} \pm D_{out}, R_1^{\pm} = R_1^{in} \pm R_1^{out} \end{aligned} \quad (6)$$

and the overall exchange rate is given by $k_{ex} = k_{OI} + k_{IO}$, the reciprocal of the lifetime where $k_{OI} = p_{in} k_{ex}$ and $k_{IO} = p_{out} k_{ex}$. In the case where the resonances of the interior and exterior populations can be spectroscopically resolved then:

$$\begin{aligned} \frac{S_{in}(q, \Delta)}{S_{in}(0,0)} &= p_{in} \left[e^{\lambda^{+}\Delta} + \left(\phi + k_{ex} - q^2 D^{+} - R_1^{+} \right) \frac{e^{-\lambda^{-}\Delta} - e^{-\lambda^{+}\Delta}}{2\phi} \right] \\ \frac{S_{out}(q, \Delta)}{S_{out}(0,0)} &= p_{out} \left[e^{-\lambda^{-}\Delta} + \left(\phi - k_{ex} - q^2 D^{-} - R_1^{-} \right) \frac{e^{-\lambda^{+}\Delta} - e^{-\lambda^{-}\Delta}}{2\phi} \right] \end{aligned} \quad (7)$$

In the case where the chemical shift is identical in the interior and exterior partitions, as is typically the case for water in cellular systems, then the combined signal will be:

$$\frac{S(q, \Delta)}{S(0, 0)} = e^{-\lambda^- \Delta} + \left(\phi - k_{ex} + (p_{in} - p_{out}) (q^2 D^- + R_1^-) \frac{e^{-\lambda^+ \Delta} - e^{-\lambda^- \Delta}}{2\phi} \right) \quad (8)$$

When the exchange rate is slow when compared to the difference in diffusion coefficients $|q^2 D^- + R_1^-| \gg k_{ex}$ (Fig 3A, equation (27)), the diffusion curves retain a strong bi-exponential character and the attenuation at a given b -value is largely independent of diffusion time, Δ . As the diffusion time increases to an intermediate regime (Fig 3A), the signal attenuation becomes dependent on Δ for a given b value, and increased attenuation is observed for higher diffusion times. Physically, this corresponds to slower diffusing population having an increased probability to move into the fast diffusing phase at longer times, where it experiences greater signal attenuation. As the exchange rate continues to increase into the fast regime, $|q^2 D^- + R_1^-| \ll k_{ex}$, the attenuation for a given b -value becomes independent of Δ again and fits well to a single exponential whose apparent diffusion coefficient (ADC) is equal to the population-weighted average diffusion coefficient (Fig 3A, equation (28)). The Kärger equations²⁵⁻²⁷ are recovered by setting $R_1^{in} = R_1^{out} = 0$.

Following the results of Murday, Cotts⁴⁰ and Neuman²¹, if the intra-cellular displacement probability distribution for particles within an impermeable spherical cavity of radius r is approximately Gaussian, then:

$$S(b) = S(0) e^{-b D_{in}^{app}} \quad (9)$$

Where the apparent internal diffusion coefficient is:

$$D_{in}^{app} = \frac{2\Phi(D_{in}, \delta, \Delta, r)}{D_{in}^2 \delta^2 (\Delta - \delta/3)} \quad (10)$$

where

$$\Phi(D, \delta, \Delta, r) = \sum_{n=1}^{\infty} \frac{1}{\alpha_n^6 (r^2 \alpha_n^2 - 2)} \left(2\alpha_n^2 D \delta - 2 + 2e^{-\alpha_n^2 D \delta} + 2e^{-\alpha_n^2 D \Delta} - e^{-\alpha_n^2 D (\Delta - \delta)} - e^{-\alpha_n^2 D (\Delta + \delta)} \right) \quad (11)$$

is a complicated sum, requiring α_n , the n^{th} root of:

$$(\alpha_n r) J_{3/2}'(\alpha_n r) - \frac{1}{2} J_{3/2}(\alpha_n r) = 0 \quad (12)$$

where J is a Bessel function of the first kind and J' is its first derivative. The roots of the Bessel function equation are readily obtained numerically. Typically, the first twenty roots are sufficient for the sum to get an error of approximately 0.1% in the applications described here.

With $\zeta = \frac{D\Delta}{r^2}$, then in the limit $\Delta \gg \delta$ and $\zeta \ll 1$, it follows that

$$\sum_{n=1}^{\infty} \frac{1}{r^2 \alpha_n^2 - 2} = \frac{1}{2}, \text{ and we recover equation (1), reflecting the free diffusion limit}$$

where particles rarely hit the cavity walls:

$$D_{app} = \frac{2D_{in}^3 \delta^2}{D_{in}^2 \delta^2} \sum_{n=1}^{\infty} \frac{1}{(r^2 \alpha_n^2 - 2)} = D_{in} \quad (13)$$

In the limit $\Delta \gg \delta$ and $\zeta \gg 1$, it follows that $\sum_{n=1}^{\infty} \frac{1}{\alpha_n^2 (r^2 \alpha_n^2 - 2)} = \frac{r^2}{10}$ and therefore:

$$D_{app} = \frac{2}{\Delta} \sum_{n=1}^{\infty} \frac{1}{\alpha_n^2 (r^2 \alpha_n^2 - 2)} = \frac{r^2}{5\Delta} \quad (14)$$

In this limit, the decay of signal intensity will be independent of Δ reflecting the limit where the results are independent of the internal diffusion coefficient.

To determine the practical validity of the formula, Monte-Carlo simulations were performed on particles held within a spherical cavity (Fig 3B, C). Data was simulated with variable values of g and Δ (Fig 3B, points) and fitted to equation (10). At relatively short times, where a molecule has relatively few collisions with the cavity walls, and the Neuman approximation is excellent (Fig 4A). At longer times where molecules are colliding with the surface of the cavity many times, the expected distribution is no longer Gaussian (Fig 4A inset, equation (29)) and the theory breaks down, evidenced by negative curvature in the diffusion plots (Fig 4A). When $qr \sim 4.49$, dS/dq reaches a minimum with $S=0$, leading to the characteristic 'diffusion diffraction' pattern^{8,24} (Fig. 4). A series expansion of the exact result reveals that the Neuman formula is exact to 2nd order in qr , and will deviate substantially from the expected result when $qr \gg \sqrt{10}$, (equation (31)). Empirically we find the Neuman formula to be numerically accurate to within a few percent in the range $qr \leq 2$ (Fig 4C).

Diffusion diffraction effects, first theoretically predicted by Stejskal and Tanner⁴¹, are not typically observed in experimental data analysing tissue and cells, though they are expected from simulations of monodisperse systems (Fig 4). Biological samples contain a heterogeneous distribution of particle sizes and the diffraction patterns associated with different confinement lengths can interfere and be averaged out⁴². HeLa cells, for example, are expected to have a variation in radius of $\pm 20\%$ in a typical population^{43,44}. Exceptions to this may occur in specific cases: for example, diffusive-diffraction effects have been observed in erythrocytes

that align in a magnetic field³⁴. While the overall displacement probability distributions in simulations from uniform spheres are not Gaussian when $\zeta \gg 1$, including a distribution of sphere sizes in simulations results in distributions that are close to Gaussian (Fig 4B). Moreover, the gradients executed in typical NMR spectrometers are not expected to be uniform. When both of these physically reasonable effects are taken into account in Monte-Carlo simulations, diffusion diffraction effects are largely averaged away. Comparing simulations including these effects to the Neuman result reveal that the region of validity is extended $qr \leq 4$, and the accuracy of radii obtained from fitting the data is improved (Fig 4H). For an experiment looking at water restrained within cells of radius 10 μm , typical experimental values with a maximum gradient $g = 60 \text{ G cm}^{-1}$ and $\delta = 2 \text{ ms}$, the product qr is 3.3, which is outside the useful limit of the Neuman result. By including the distribution of sphere sizes in simulations, the Neuman formula nevertheless continues to provide an excellent description of restriction effects.

Overall, our model is expressed by equation (8), where D_{in} is replaced by an apparent diffusion coefficient D_{in}^{app} in equation (10). The fitting parameters can be rearranged to give the cell density in cells per unit volume,

$$\frac{N_{cell}}{V_{tot}} = \frac{3p_{in}}{4\pi r^3} \quad (15)$$

and the permeability, a measure of net water flux, in m s^{-1}

$$\kappa = \frac{k_{IO}r}{3} = \frac{p_{out}k_{ex}r}{3} \quad (16)$$

Determining the accuracy of fitted parameters

To quantitatively determine the regions of parameter space where we expect to obtain reliable fitting parameters, an extensive set of Monte-Carlo

simulations were performed systematically varying all parameters, which were then fitted globally to the model (equations (8) and (10)). Experimentally accessible values of g and Δ were used, remaining in the limit $qr < 4$. The returned parameters were compared to those employed in the simulation to determine accuracy. The uncertainties in the numerical fitting process were established by a boot-strapping procedure where a set of diffusion time curves are fit whose total number is equal to the number of simulated diffusion times but selected at random with replacement, a process repeated 1000 times⁴⁵. The fitted parameters are then collated and the standard deviation of each is calculated. Overall, this analysis reveals regions of parameter space where the theory is expected to return reliable parameters and where to expect systematic errors.

The following baseline values were used in the simulation, chosen to reflect typical values expected for an NMR/MRI sample containing mammalian cells: $D_{in} = 500 \mu\text{m}^2 \text{ s}^{-1}$, $D_{out} = 2000 \mu\text{m}^2 \text{ s}^{-1}$, $k_{ex} = 3.55 \text{ s}^{-1}$, $R = 10 \mu\text{m}$ and $P_{in} = 0.15$. Individual parameters were then systematically varied over a range with the others fixed (r Fig 5A, p_{in} 5B k_{ex} Fig 5C, aspect ratio of the cavity Fig 5D R_1^m Fig 5E, R_1^{out} Fig 5F).

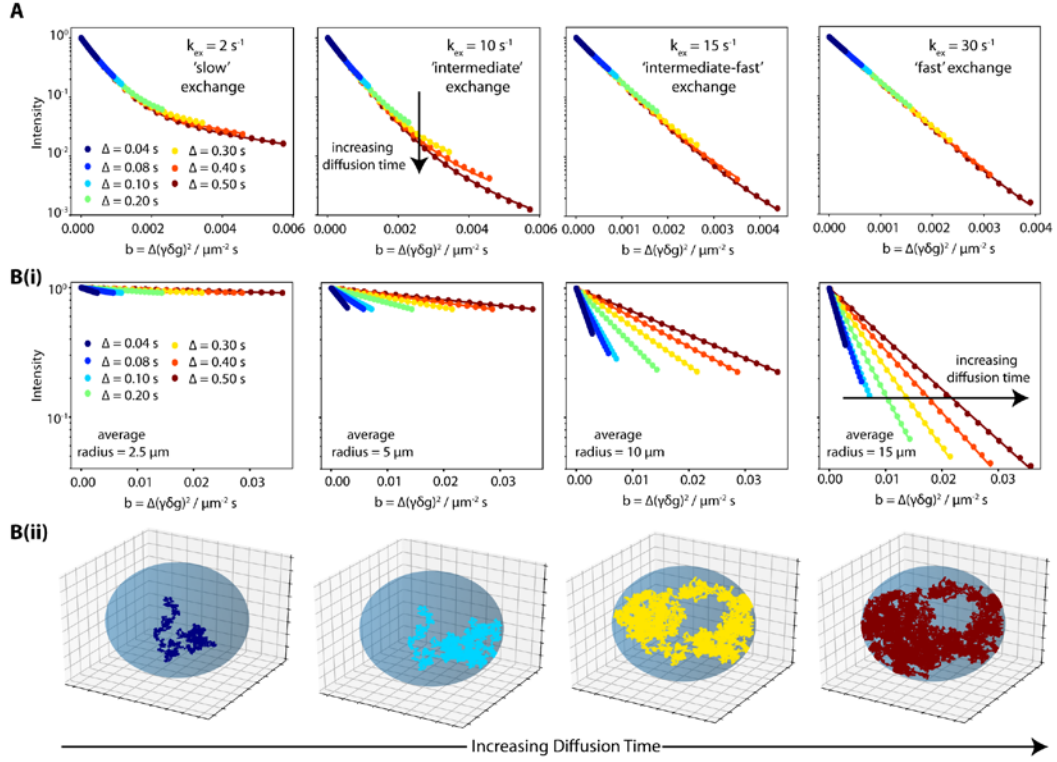


Figure 3 **A** Simulated data neglecting relaxation were fitted to equation (8) showing the effect of exchange between freely diffusing populations on the diffusion attenuation. The slow and fast diffusion coefficient were 200 and 2000 $\mu\text{m}^2 \text{ s}^{-1}$, weighted with a 1:9 ratio. In general, for a given b -value we see differential attenuation depending on the diffusion time, Δ . The solid lines show fits to the Kärger equations. **B(i)** Simulated data showing the effect of restricted diffusion. In each case, particles are diffusing within an impermeable sphere with the diffusion coefficient set to 500 $\mu\text{m}^2 \text{ s}^{-1}$. In all cases, the sphere sizes are normally distributed: for each subplot the average radius is stated and the standard deviation is set to 20% of this value, reflecting size distributions observed in cellular systems. Data simulated for variable values of Δ do not show identical attenuation at a given b -value. Data are fitted to the Neuman equation²¹ (solid lines equation (10)). **B(ii)** Schematic showing single particle trajectories for case when average radius is set to 10 μm . The number of collisions with the membrane increases with the diffusion time and the diffusion attenuation increasingly becomes a property of the restricting geometry rather than the viscosity.

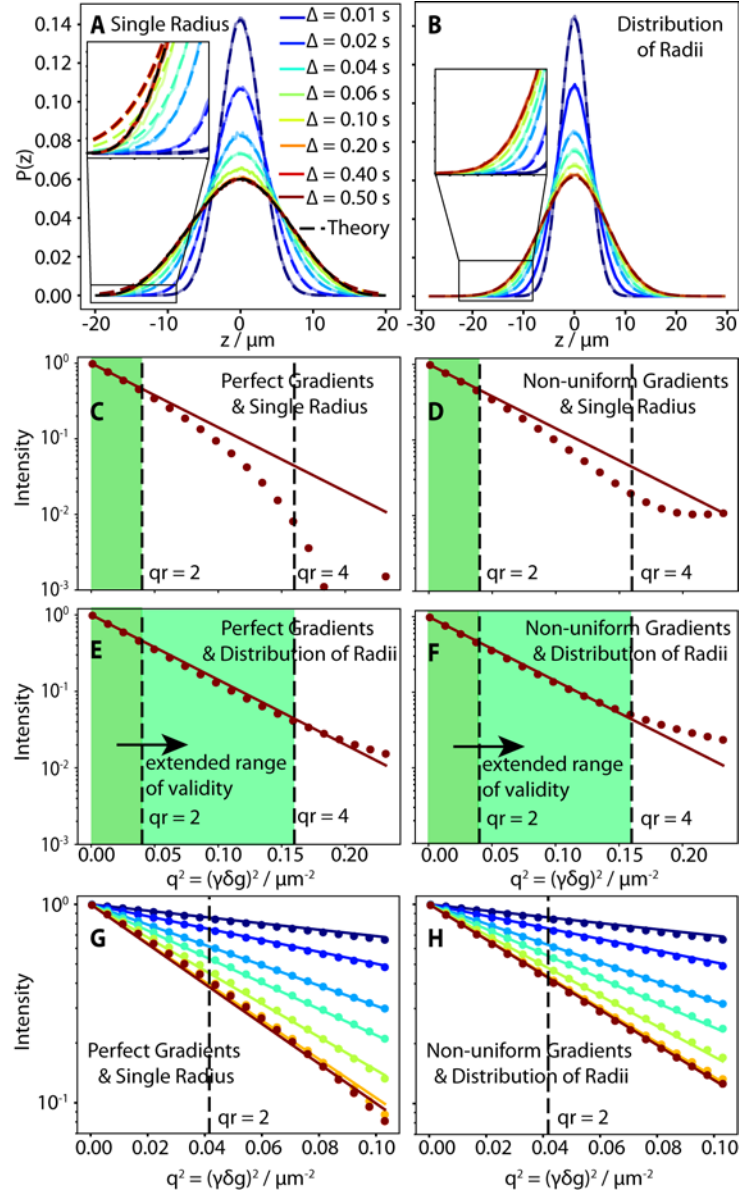


Figure 4 **A** Simulated displacement distributions, $P(z)$, for different diffusion times calculated from 500,000 simulated single particle trajectories where each particle is started in a random position in an impermeable spherical cavity with a $10 \mu\text{m}$ radius and internal diffusion coefficient $500 \mu\text{m}^2 \text{s}^{-1}$. The coloured dashed lines indicate Gaussian fits to the simulated distributions. The distributions are Gaussian at short times, but at long times, there are significant deviations (inset), indicating a breakdown in the model proposed by Neuman. The simulations tend towards the exact expected probability distribution expected at long times (equation(29)). **B** Simulated displacement distributions as for A, for a Gaussian distribution of impermeable spherical particles with average radii $10 \mu\text{m}$ and standard deviation $2 \mu\text{m}$, as expected for a preparation of HeLa cells.⁴⁴ The distributions are Gaussian (inset). **C-F** Calculated NMR signal at $\Delta = 0.5$ s with average gradient

strengths (g) varying from 4 to 80 G cm⁻¹ and $\delta = 3$ ms. For **C** and **D** the NMR signal is calculated using the displacement probability distribution in **A**, assuming a uniform radius whereas **E** and **F** are calculated using the distribution in **B** that includes a range of cell sizes. For **C** and **E** the signal is calculated assuming perfectly uniform gradients whereas in **D** and **F** the effects of non-uniform gradients are explicitly included (Fig S1). The solid lines in **C-F** show the signal attenuation predicted by the Neuman equation. The combined effect of including a distribution, and non-uniform gradients is to extend the limit of the theory from $qr \leq 2$ to $qr \leq 4$. **G-H** Simulated NMR decay curves for typical experimental parameters: varying both g from 4 to 60 Gcm⁻¹ and $\Delta = 0.01$ to 0.5 s with $\delta = 2$ ms. The data is fitted to the Neuman result (equation (10)) assuming perfect gradients and the displacement distribution in **A** for **G** and non-uniform gradients and the distribution in **B** for **H**. The fitted diffusion coefficients are 483 ± 7 (**G**) and 480 ± 10 $\mu\text{m}^2 \text{s}^{-1}$ (**H**) and radii are 10.82 ± 0.03 (**G**) and 10.15 ± 0.05 μm (**H**). Error bars in model parameters are determined by a bootstrapping analysis, fitting datasets made up of randomly chosen Δ values with replacement, a process repeated 1000 times. The fitted parameters are then collated and the standard deviation of each is calculated. Including the effects of non-uniform gradients and accounting for an expected range of sphere sizes effectively extends the region of validity of the Neuman result (equation (10)).

Overall, fitted parameters fall within 50% of the true value (Fig. 5, green) indicating that the model can be used with confidence when analysing water diffusion data from eukaryotic cells (Fig. 5). The analysis identifies regions of parameter space where the error from bootstrapping is smaller than the actual error. This indicates regions of parameter space where systematic errors are expected. Relatively large discrepancies in parameters are observed when the exchange rate tends towards the fast regime (equation (28)), where $k_{ex} > 40 \text{ s}^{-1}$ (Fig 5C). Previously determined exchange rates from cells tend to be $< 10 \text{ s}^{-1}$ ^{43,46-49}, an area where the model is accurate. Notably, the bulk diffusion coefficient in

the interior of the cell D_{in} is often poorly determined (Fig 5A-E, orange), a factor likely due to our specific choice of g , δ and Δ values.

By contrast, the effective cell radius is typically well determined (Fig 5A), and the fitted parameters are tolerant to both large changes in cell abundance (Fig 5B), and substantial changes in the geometry of the cell (Fig 5D). The radius was underestimated when cells exceed 30 μm , as the 0.5 s maximum value of Δ used here is insufficient to lead to significant restriction of particles. The fitted radius corresponds to the effective radius of the equivalent sphere whose volume matches that of the cells. Thus, while we expect a large variation in cell shape, the fitted parameters nevertheless provide a robust measurement of the overall cell volume (Fig 5D).

The effects of differential longitudinal relaxation can also be significant (Fig 5E,F), as has been noted previously⁵⁰. The R_1 of bulk water is small ($\sim 0.25 \text{ s}^{-1}$) and will not have a major impact on the on the data. Within biological tissue or in the presence of paramagnetic agents the longitudinal relaxation rate can be significantly higher and failure to account for this can result in significant errors in the fitted values for other parameters. R_1^{out} is typically well determined (Fig 5G,H), but R_1^{in} is poorly determined unless its value becomes larger than $\sim 0.75 \text{ s}^{-1}$. Below this limit, the effects of R_1^{in} are imperceptible in the data. For example in Fig 5F R_1^{in} shows consistent systematic errors, and the fitted value tends to zero. In these simulations, its maximum contribution to the data is ca. 1% ($R_1^{in} 0.25 \text{ s}^{-1}$). When R_1^{in} is greater than 0.75 s^{-1} , the contribution to the raw data is substantially larger, and the parameter can be reliably measured free of systematic errors (Fig. 5E).

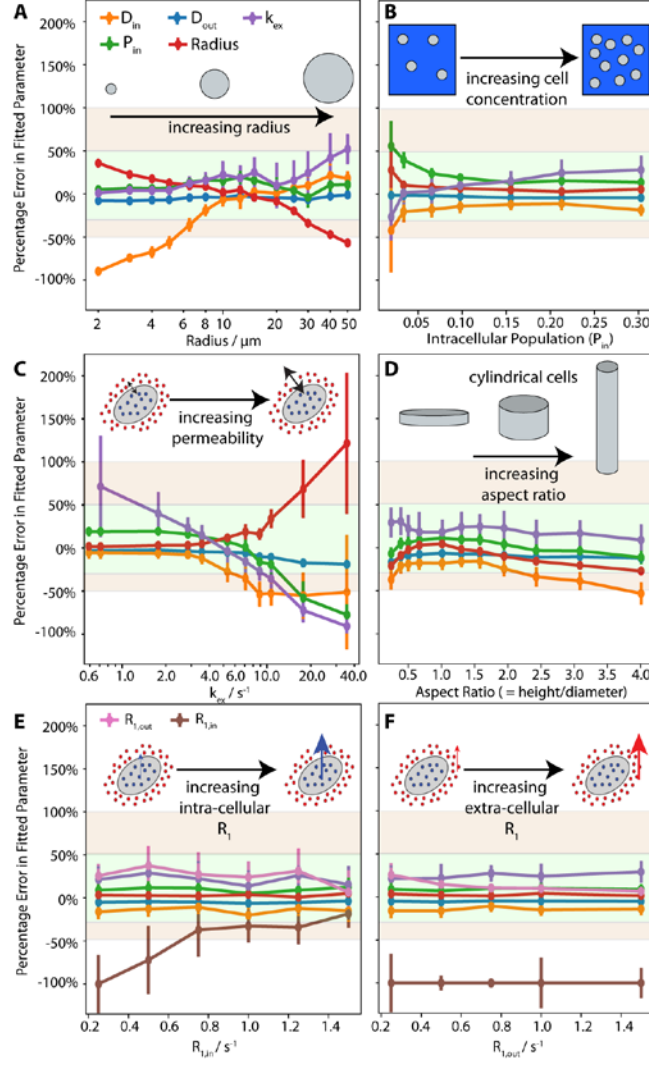


Figure 5 Error plots made by comparing fitted parameters using our model (equations (8) and (10)) to values set in the simulations. Error bars in model parameters are determined by a bootstrapping analysis, where a subset of curves is selected whose total number is equal to the number simulated but selecting from values of Δ at random with replacement, a process repeated 1000 times. Where a parameter's value is not varied within a particular plot its value is fixed as follows: $D_{in} = 500 \mu m^2 s^{-1}$, $D_{out} = 2000 \mu m^2 s^{-1}$, $k_{ex} = 3.55 s^{-1}$, $R = 10 \mu m$ and $P_{in} = 0.15$. In Figures **A-D** the values of $R_1^{in} = R_1^{ex} = 0 s^{-1}$. In Figures **E** and **F** the values of $R_1 = 0.25 s^{-1}$ when they are not being varied.

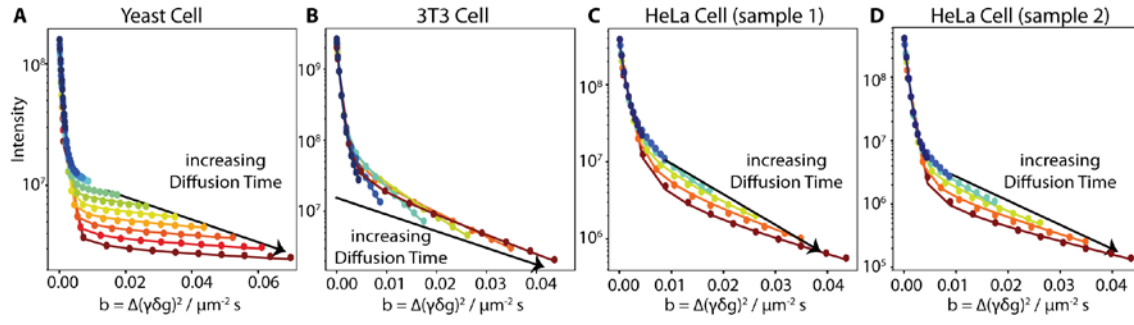


Figure 6 NMR data collected on real systems of **A** yeast, **B** 3T3 and **C**, **D** HeLa cells with varying cell densities fitted to the model as described in section 4.1. The data is well explained by the model, and the extracted parameters are in good agreement with those obtained independently (**Tables 1**, **2**).

Table 1 Derived parameters from NMR analysis. Error bars are taken from 1000 bootstrap runs.

Sample	$D_{in}/\mu\text{m}^2\text{s}^{-1}$	$D_{out}/\mu\text{m}^2\text{s}^{-1}$	k_{ex}/s^{-1}	Radius/ μm	P_{in}	$R_{1,in}/\text{s}^{-1}$	$R_{1,out}/\text{s}^{-1}$	Cell density/ ml^{-1}	$\kappa/\mu\text{m}\text{s}^{-1}$
Yeast (<i>S. cerevisiae</i>)	20 ± 4	1250 ± 15	1.23 ± 0.15	3.57 ± 0.18	0.092 ± 0.03	0.86 ± 0.13	0.81 ± 0.07	$5.3\times 10^8\pm 8\times 10^7$	1.33 ± 0.09
3T3	826 ± 65	1980 ± 44	4.45 ± 0.24	12.42 ± 0.11	0.086 ± 0.03	0.00 ± 0.00	0.29 ± 0.03	$1.1\times 10^7\pm 4\times 10^5$	16.84 ± 0.79
HeLa Sample 1	291 ± 62	1310 ± 69	7.39 ± 0.66	10.38 ± 0.18	0.169 ± 0.012	0.00 ± 0.09	0.51 ± 0.07	$3.95\times 10^7\pm 6.6\times 10^6$	21.25 ± 1.78
HeLa Sample 2	344 ± 86	1750 ± 61	6.43 ± 0.49	10.69 ± 0.22	0.046 ± 0.003	0.00 ± 0.02	0.28 ± 0.06	$8.99\times 10^6\pm 1.25\times 10^6$	21.85 ± 1.10

Table 2 Radius and exchange parameters determined in this study compared to literature measurements of these values.

Sample	Radius (here) / μm	Radius (literature) / μm	k_{io} (here) / s^{-1}	k_{io} (literature) / s^{-1}	Cell density (NMR) / cells/ml	Cell density (independent) / cells/ml
Yeast (<i>S. cerevisiae</i>)	3.57	$3.97^{51}, 2.88^{52}$ (average sizes)	1.16	$1.49^{\ddagger 53}, \sim 1.80^{46}$	5.3×10^8	$\sim 2.1\times 10^9$
3T3	12.42	5.45 to 14.33^{44}	4.07	---	1.1×10^7	----
HeLa (sample 1)	10.38	$10.5\pm 2.2^{43,44}$	6.14	$8.4^{\S 43}$	3.95×10^7	$\sim 3.5\times 10^7$
HeLa (sample 2)	10.69	$10.5\pm 2.2^{43,44}$	6.13	$8.4^{\S 43}$	8.99×10^6	$\sim 1.1\times 10^7$

[^]cell density estimated from OD600 measurement for yeast, and in a coulter counter (HeLa).

[‡] Temperature of measurement is not stated.

[§] Measurement is made at 310 K rather than 298 K so exchange rate/permeability is expected to be higher.⁴⁶

Experimental application to cell samples

To obtain experimental validation of our analysis, suspensions of yeast cells and mammalian 3T3 and HeLa cells were prepared (Methods, H). STE diffusion data with variable g and Δ were acquired (Methods, G). The data were analysed by the INDIANA software, and the reported parameters from fitting were physically reasonable (Table 1). The fitted radii, permeabilities and cell counts were consistent with the expected values (Table 2). In the case of the smaller yeast cells, the radius will be systematically slightly overestimated (Fig. 5A), which leads to a slight overestimate of the cell count. For the mammalian cells, the agreement is improved. Similarly, the intra-cellular diffusion coefficients and the small longitudinal relaxation values are likely to be systematically inaccurate (Fig. 5).

Discussion

INDIANA is a method to characterise the cell density, size, permeability, as well as intrinsic relaxation and diffusion coefficients of cell suspensions. Experimentally, STE water diffusion spectra can be obtained for a range of g and Δ values on either an MRI or an NMR spectrometer. The data are then globally analysed using equations (8) and (10). A freely downloadable software package is provided to facilitate this analysis.

The advantage of the STE over the SE experiment is that as $R_1 \ll R_2$, longer values of b can be accessed, a regime where restricted diffusion effects are more prominent and slower exchange rates can be reliably fitted. The need to analyse the data using a model that contains restricted diffusion and exchange can be identified if analysis of the data are not well explained using equation (2). More generally, as permeability can significantly impact data, a multi-exponential model that neglects exchange is unlikely to be suitable⁵⁴.

Following the approaches of Kärger and Neuman, our model is valid when the displacement distributions of water are approximately Gaussian. Where a sample comprises uniform spheres, simulations reveal the model is reasonable in the regime $qr \leq 2$. In the case where there is a distribution of particle sizes, as expected in biological samples, the average radius can still be determined robustly, up to an extended range of $qr \leq 4$. By comparing the parameters obtained from fitting our model to an extensive set of simulations, we determine the region of parameter space where we expect the regions of parameter space effectively encompassed by our theory. This region encompasses the regions likely to be encountered when studying eukaryotic cells by NMR.

Our approach assumes that cells are spherical which in principle presents a significant limitation. By simulating data with non-spherical geometries, and fitting our spherical model, we nevertheless obtain an ‘effective’ radius (Fig 5D). The volume associated with this effective sphere is in excellent agreement with the volume of simulated non-spherical cavity suggesting that the effective radius returned from our analysis provides generally useful information even in the case where cells are non-spherical, a finding that establishes the generality of the method.

There are also several inherent assumptions in the Kärger diffusive exchange model that might appear inappropriate. Notably, from a microscopic perspective, all particles within a phase are assumed to have an equal probability of changing phase, whereas in cellular systems exchange can only occur at membranes. Nevertheless, our simulations and experimental validation show the theory maintains an excellent agreement within the ranges identified (Figure 5). As we are using a stimulated echo experiment in this work, and $\Delta \gg \delta$, it is

appropriate to use the ‘short gradient pulse’ approximation inherent in the Kärger model.

Another feature of the model is that the intra- and extra-cellular water mobility is modelled using unique diffusion coefficients. Microscopically, diffusion will involve a variety of collisions between the probe molecule, and the myriad of other components molecular of the system. Further complicating the picture, transient complexes might be formed by interactions between sticky molecules and ‘crowding’ type interactions that locally limit the passage of molecules through the medium. These interactions for water will typically be on ps-μs timescales^{55,56} and so from the perspective of the diffusion experiment they are all in fast exchange, $k_{ex} \gg q^2 \Delta D$. All these effects will then effectively average to give a single ensemble averaged diffusion coefficient, represented here by D_{in} and D_{out} . Any change in the cellular interior that substantially alters the passage of the probe molecule will alter these values.

A clear distinction should be made between these diffusion constants, which are ‘true’ values, in the sense that they describe the translational diffusion of a particle throughout a medium according to Fick’s laws, and apparent diffusion coefficients (ADC) as typically reported in the NMR and MRI literature. In the latter, an ADC is often derived from 2 measurements to give a first order description of the loss of signal intensity. The resulting value is only physically reasonable in the case of the motion of the probe molecule in a single isolated phase. While this measurement reveals changes in a sample and can be obtained quickly, it is not directly physically meaningful. In our INDIANA method, by performing a larger number of experiments with varying δ and g , we are seeking to

measure the ‘true’ diffusion coefficients of the probe molecule in the relevant phase.

While we have focused here on the water signal, where both interior and exterior portions are detected at the same chemical shift value, this analysis could in principle be carried out on a resonance from any molecule that can be spectroscopically resolved. Expressions are produced for the case where the interior and exterior chemical shifts differ (equation (7)). This approach will be particularly powerful where a resonance has been isotopically enriched, allowing background-free measurements. For MRI applications where there are limitations on total experiment time, it will be possible to apply compressed sensing algorithms to reduce the number of q and Δ values required.

With the tremendous growth of in-cell NMR and MRI applications, our method provides a straightforward route to determine the abundance, permeability, size and relaxation characteristics of cellular preparations, in a non-invasive fashion. The methodology could also be applied for example to monitor cell preservation in perfusion system: gross changes in the cell permeability, size or number should be quantifiable, provided they are bigger than uncertainties in the parameters themselves. There is also clear potential for using these experiments in the context of diffusion-MRI to obtain information on cell suspensions.

Acknowledgements

The authors wish to thank Kevin Ray, Nicola Sibson, Michael Chappell and James Kennedy for helpful discussion. G.K. thanks the EPSRC for funding. The work was part funded by the BBSRC grant BBSRCBB/J014346/1.

Methods

A Monte Carlo simulations

Monte-Carlo simulations were performed similarly to those described previously^{28,57}. Space is discretised into cubes each with either a semi-permeable sphere or a cylinder at its centre, with identical cubes extending in all directions (Fig 2). A particle is initialised at a random point in space such that the average initial positions of the particles reflects the intra- and extra-cellular mole fractions. The extra- and intra-cellular spaces are each characterised by a single diffusion coefficient, D_{out} and D_{in} respectively. The current displacement, and an intensity factor are then updated for each step of duration δt :

$$\begin{aligned}
 x(t + \delta t) &= x(t) + \delta x_{step} \\
 y(t + \delta t) &= y(t) + \delta y_{step} \\
 z(t + \delta t) &= z(t) + \delta z_{step} \\
 I(t + \delta t) &= \begin{cases} \text{inside: } I(t)e^{-R_1^{\text{in}}\delta t} \\ \text{outside: } I(t)e^{-R_1^{\text{out}}\delta t} \end{cases}
 \end{aligned} \tag{17}$$

The displacement step sizes δx_{step} , δy_{step} , and δz_{step} are determined by independently drawing a random number from a Gaussian distribution of standard deviation $\sqrt{2D\delta t}$, where the value of D selected depends on whether the particle is inside, or outside the cell. Total simulation times were as specified (typically 0.5 s) and δt was typically 10 μ s, a value that provides a compromise between overall simulation time and accuracy, values that necessitate 50,000 individual steps. The final signal intensity is obtained by summing the final values over a set of independent particle trajectories:

$$\frac{S(q, \Delta)}{S(0, 0)} = \text{Re} \left(\sum_{\text{Particles}} I(\Delta) e^{-iqz} \right) = \sum_{\text{Particles}} I(\Delta) \cos(qz) \tag{18}$$

Typically, δ was set to 2 ms and g was varied quadratically in the range 4 to 60 G cm⁻¹. mimicking commonly used experimental values. Under conditions where attenuation occurs more quickly these values are adjusted such that at the highest applied gradient field strength the signal is attenuated by approximately two orders of magnitude.

After each step, the new position of the particle is checked. If particles have just crossed the cell membrane, we determine if we will accept this jump. The in-to-out or out-to-in probabilities are calculated^{28,58}:

$$P_{io} = \frac{2s\kappa}{D_{in}} \quad P_{oi} = \frac{2s\kappa}{D_{out}} \quad (19)$$

where s is the distance between the particle and cell surface prior to the jump, and κ , is the cell's permeability, which is in turn is related to the in-to-out rate constant k_{io} ⁴⁶:

$$\kappa = k_{io} \frac{V}{A}$$

(20)**Error!**
Bookmark
not
defined.

where V is the cell volume and A is the cell surface area. In the case of spherical cells $\kappa = k_{io} \frac{r}{3}$ where r is the radius. A uniform random number between 0 and 1 is drawn. If the relevant probability is less than this number, the crossing is accepted and if not, the jump is rejected, and the particle is translated back into the compartment from which it came. Treating the collisions as specular reflections gives no discernible advantage in terms of accuracy, as reported previously^{59,60}.

The Monte-Carlo simulations were implemented in Python, and the source code is available. Simulations are composed of 500000 single particle trajectories.

To account for variable sizes of cells, in each trajectory the cell radius is randomly drawn from a normal distribution with a specified average value and the standard deviation set to 20% the average value (Fig 4). As the single particle trajectories are all independent, the calculation can be easily run in parallel.

The Monte Carlo simulations were performed over a wide range of conditions parametrised by D_{in} , D_{out} , k_{ax} , r , p_{in} (in all cases) and R_1^{in} and R_1^{out} (in Figures 5E and 5F only), to validate the robustness of the theoretical model (Fig 5). The time required to simulate a single set of parameters, parallelised over 20 2.4 GHz Intel Xeon CPUs was approximately 50 minutes.

To model the effects of non-uniform gradient profiles expected in typical instruments³⁹, we evaluate the following (Fig S1):

$$\frac{S(q, \Delta)}{S(0, 0)} = \sum_{G=G_{min}}^{G_{max}} \sum_{Particles} I(\Delta) P(G) e^{iq(G)z(\Delta)} \quad (21)$$

where $P(G)$ is the probability of a given gradient strength.

B Fitting simulations and experimental data to model

Actual and simulated NMR data were fit to the model (equations (8) and (10)) with a Python script making use of the Lmfit module⁶¹ (Fig 3C). Uncertainties in individual signal intensities were estimated from repeated measurements, and found to be approximately a constant percentage over several orders of magnitude of signal (Fig S3). To account for this weighting, fitting was performed on the logarithm of the signal intensity. No prior information is fed into the fitting procedure and the only restraint placed on the free parameters is that their values must be ≥ 0 . The fitted values were compared to those specified in the simulations (Fig 5) to determine their accuracy. The values obtained on ‘real’ experimental systems (Fig 6) were in excellent agreement with independently validated values.

A global fitting strategy is adopted to extract model parameters from the data to avoid overfitting the data. F-tests are also used to ensure that increasing the complexity of the models result in statistically significant improvements in fits to the data.

C Detailed derivation of equation (8)

To calculate the detected signal intensity, we need to evaluate the following

$$S(q, \Delta) = e^{\rho \Delta} S(0, 0) O$$

Where ρ is the evolution matrix from equation (3), $S(0)$ contains the equilibrium signal intensity, and O is an operator that combines the populations to give a single observed signal, appropriate for when the chemical shift of the interior and exterior populations cannot be distinguished. The matrix exponential can be evaluated by taking a square matrix with the eigenvalues on the diagonal (D), the matrix of eigenvectors P and its inverse:

$$S(q, \Delta) = P e^{-D \Delta} P^{-1} \begin{bmatrix} p_{in} \\ p_{out} \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} S(0, 0) \quad (22)$$

With the eigenvalues λ^\pm are defined in terms of the substitutions ξ^\pm and ϕ , defined earlier (equation (6)):

$$P = \begin{pmatrix} 1 & 1 \\ \frac{-2k_{ex}p_{out}}{\xi^- + \phi} & \frac{-2k_{ex}p_{out}}{\xi^- - \phi} \end{pmatrix} \quad (23)$$

The matrix exponential is given by:

$$P e^{-D \Delta} P^{-1} = \frac{1}{2\phi} \begin{pmatrix} -(\xi^- - \phi)e^{-\lambda^- \Delta} + (\xi^- + \phi)e^{-\lambda^+ \Delta} & 2k_{ex}p_{in}(e^{-\lambda^- \Delta} - e^{-\lambda^+ \Delta}) \\ 2k_{ex}p_{out}(e^{-\lambda^- \Delta} - e^{-\lambda^+ \Delta}) & (\xi^- + \phi)(e^{-\lambda^- \Delta} - e^{-\lambda^+ \Delta}) \end{pmatrix}$$

Multiplying by the equilibrium populations is required for the final result (equation (8)). An alternative form is:

$$\frac{S(q, \Delta)}{S(0, 0)} = \frac{1}{4k_{ex}\phi} \left(\left((q^2 D^- + R^-)^2 - (k_{ex} - \phi)^2 \right) e^{-\lambda^+ \Delta} - \left((q^2 D^- + R^-)^2 - (k_{ex} + \phi)^2 \right) e^{-\lambda^- \Delta} \right)$$

D Fast and slow exchange limits

The expression reduces to the expected form in the fast and slow exchange limits.

Rearranging ϕ allows for fast and slow exchange definitions:

$$\phi = \sqrt{(q^2 D^- + R_1^-)^2 + 2(q^2 D^- + R_1^-)k_{ex}(p_{out} - p_{in}) + k_{ex}^2}$$

Slow exchange is where $|q^2 D^- + R_1^-| \gg k_{ex}$, and fast exchange is where

$|q^2 D^- + R_1^-| \ll k_{ex}$. In these two limits:

$$\begin{aligned} \phi_{fast} &= k_{ex} + (p_{out} - p_{in})(q^2 D^- + R_1^-) \\ \phi_{slow} &= k_{ex}(p_{out} - p_{in}) + (q^2 D^- + R_1^-) \end{aligned} \quad (24)$$

which leads to the following Eigenvalues:

$$\begin{aligned} \lambda_{slow}^- &= q^2 D_{out} + R_1^{out} + k_{OI} \\ \lambda_{slow}^+ &= q^2 D_{in} + R_1^{in} + k_{IO} \\ \lambda_{fast}^- &= p_{in}(q^2 D_{in} + R_1^{in}) + p_{out}(q^2 D_{out} + R_1^{out}) = q^2 \bar{D} + \bar{R}_1 \\ \lambda_{fast}^+ &= k_{ex} \end{aligned} \quad (25)$$

Notably, λ_{fast}^- tends to the combined population averaged value of D and R_1 , and the slow exchange values tend towards the values expected for the isolated species. In the slow exchange limit, signal from the interior, exterior and total become:

$$\begin{aligned}
\frac{S_{in}(q, \Delta)}{S_{in}(0, 0)} &= p_{in} e^{-\lambda_{slow}^+ \Delta} + \frac{k_{ex} p_{in} p_{out}}{k_{ex} (p_{out} - p_{in}) + (q^2 D^- + R_1^-)} \left(e^{-\lambda_{slow}^- \Delta} - e^{-\lambda_{slow}^+ \Delta} \right) \\
\frac{S_{out}(q, \Delta)}{S_{out}(0, 0)} &= p_{out} e^{-\lambda_{slow}^- \Delta} + \frac{k_{ex} p_{in} p_{out}}{k_{ex} (p_{out} - p_{in}) + (q^2 D^- + R_1^-)} \left(e^{-\lambda_{slow}^- \Delta} - e^{-\lambda_{slow}^+ \Delta} \right) \\
\frac{S_{slow}(q, \Delta)}{S_{slow}(0, 0)} &= p_{in} e^{-\lambda_{slow}^+ \Delta} + p_{out} e^{-\lambda_{slow}^- \Delta} + \frac{2k_{ex} p_{in} p_{out}}{k_{ex} (p_{out} - p_{in}) + (q^2 D^- + R_1^-)} \left(e^{-\lambda_{slow}^- \Delta} - e^{-\lambda_{slow}^+ \Delta} \right)
\end{aligned} \tag{26}$$

In the limit where $k_{ex}=0$, we recover the expected population weighted bi-exponential decay for the total signal

$$\begin{aligned}
\frac{S_{in, k_{ex}=0}(q, \Delta)}{S_{in}(0, 0)} &= p_{in} e^{-(q^2 D_{in} + R_1^{in}) \Delta} \\
\frac{S_{out, k_{ex}=0}(q, \Delta)}{S_{out}(0, 0)} &= p_{out} e^{-(q^2 D_{out} + R_1^{out}) \Delta} \\
\frac{S_{k_{ex}=0}(q, \Delta)}{S(0, 0)} &= p_{in} e^{-(q^2 D_{in} + R_1^{in}) \Delta} + p_{out} e^{-(q^2 D_{out} + R_1^{out}) \Delta}
\end{aligned} \tag{27}$$

Finally, in the fast exchanging limit we obtain the following expressions:

$$\begin{aligned}
\frac{S_{fast, in}(q, \Delta)}{S_{in}(0, 0)} &= p_{in} \left(e^{-\lambda_{fast}^- \Delta} + \frac{p_{out} (q^2 D^- + R_1^-) (e^{-k_{ex} \Delta} - e^{-\lambda_{fast}^- \Delta})}{k_{ex} + (p_{out} - p_{in}) (q^2 D^- + R_1^-)} \right) \\
\frac{S_{fast, out}(q, \Delta)}{S_{out}(0, 0)} &= p_{out} \left(e^{-\lambda_{fast}^- \Delta} - \frac{p_{in} (q^2 D^- + R_1^-) (e^{-k_{ex} \Delta} - e^{-\lambda_{fast}^- \Delta})}{k_{ex} + (p_{out} - p_{in}) (q^2 D^- + R_1^-)} \right) \\
\frac{S_{fast}(q, \Delta)}{S(0, 0)} &= e^{-\lambda_{fast}^- \Delta} = e^{-(q^2 \bar{D} + \bar{R}) \Delta}
\end{aligned} \tag{28}$$

Where the total signal decays with a single exponential whose decay constant is the population average of the diffusion and relaxation rates.

E Long time limit for restricted diffusion within a sphere

In the case of restricted diffusion, we can calculate the probability distribution in the long time limit. The displacement within a sphere in spherical polar coordinates will be $dz = r(\cos \theta_1 - \cos \theta_2)$. The probability of a given displacement

will depend on the area $(r \sin \theta)^2$. The un-normalised displacement distribution will be:

$$P(z) = \int_0^\pi \int_0^\pi \sin^3 \theta_1 \sin^3 \theta_2 \delta(r(\cos \theta_1 - \cos \theta_2) - z) d\theta_1 d\theta_2$$

Changing variables such that $u = \cos \theta_1$ and $v = \cos \theta_2$:

$$P(z) = \int_{-1}^1 \int_{-1}^1 (1-u^2)(1-v^2) \delta(r(u-v) - z) dudv$$

Following a final change of variables $\kappa = \frac{|z|}{r}$ and we arrive at the following symmetric normalised distribution in the range $-2 > \kappa > 2$, and zero otherwise:

$$P(\kappa) = \frac{3}{160} (2 - \kappa)^3 (4 + \kappa(6 + \kappa)) \quad (29)$$

This can be converted to expected NMR signal intensity^{41,62}:

$$S(q) = \int_{-\infty}^{\infty} P(z) e^{iqz} dz = \frac{9(qr \cos(qr) - \sin(qr))^2}{(qr)^6} \quad (30)$$

We note that the Neuman approximation for restricted diffusion does not converge on this limit. For shorter times and specific gradient strengths encountered on typical NMR and MRI spectrometers, the Neuman approximation is sufficiently accurate. This can be proven from taking the expansion of equation (30) to 6th order:

$$\frac{S_i}{S_0} = 1 - \frac{(qr)^2}{5} + \frac{3(qr)^4}{175} - \frac{4(qr)^6}{4725} \dots \quad (31)$$

reveals that the distribution will be effectively Gaussian providing the quartic term is smaller than the quadratic term. The Neuman formula is exact in this limit to 4th order in qr (equation (14)). The formal requirement for this limit will be $10 \gg (qr)^2$. In more practical terms, the Neuman result is numerically accurate to

within a per cent providing $qr \leq 2$. This is a challenge for experiments, as $10 \mu\text{m}$ spheres, analysed with a gradient of 60 G cm^{-1} applied for 2 ms , looking at water where γ is $26,700 \text{ G}^{-1} \text{ rads s}^{-1}$ exceed this limit. The limit of applicability is extended to $qr \leq 4$ when we account for a distribution of sphere sizes, and non-uniform gradients inherent in all spectrometers is taken into account (Fig. 4).

F Continuous diffusion theory

The results from Monte-Carlo simulations were cross-validated with calculations conducted with a macroscopic perspective. In one dimension, the concentration in a region of space will be updated according Fick's second law of diffusion:

$$\frac{dc}{dt} = D \frac{d^2c}{dx^2}$$

Which can be numerically integrated using the following scheme:

$$c(i, t + \delta t) = c(i, t) + D \frac{\delta t}{\delta x^2} (c(i+1, t) + c(i-1, t) - 2c(i, t)) \quad (32)$$

which is readily expanded to 2 and 3 dimensions. A spherical cavity was maintained in the calculation by establishing boundaries where $x_i^2 + y_j^2 + z_k^2 < r^2$. A desired total concentration was placed into a single square at a specific location f along the z-axis, which was then allowed to evolve for the desired amount of time according to equation (32). The concentration in all cubes inside the sphere, $c_{ijk}(f)$ were saved at desired times Δ . Starting locations f were varied uniformly from 0 to r , and the spherical average was calculated by determining the displacement probabilities along the z axis by rotating the initial position coordinates positions. For a given rotation of θ about the x axis, the probability distribution will be:

$$P(\delta z, f, \theta) = \sum_{ijk} c_{ijk}(f) \delta(y_j \sin \theta + z_k \cos \theta - \delta z)$$

Which can be spherically averaged to yield the final distribution:

$$P(\delta z) = \int_0^r \int_0^\pi P(\delta z, f, \theta) f^2 \sin \theta d\theta df$$

Which is converted to NMR signal intensity using

$$\frac{S(q, \Delta)}{S(0, 0)} = \int_{-2R}^{2R} P(z) \cos(qz) dz \quad (33)$$

The ratio $D\delta t/\delta x^2$ was set to 0.15, and space was split into cubic boxes of side length δx 1 μm , the sphere's radius was set to 20 μm , $\delta = 2$ ms, $\gamma = 26,700$ G rads s^{-1} and 100 gradient strengths in the range 0-60 G cm^{-1} , varied quadratically. The interior diffusion coefficient was set to 10^{-9} $\text{m}^2 \text{s}^{-1}$, a value expected for water free in solution. With this value, the time step δt becomes 1.5 ms and for 10,000 steps, the total simulation time was 1.5 s which takes approximately 5 minutes on a single Intel i7 processor. The distributions from this method were in excellent agreement with those obtained from Monte-Carlo simulations (Fig S2) demonstrating the equivalence of the two approaches.

G NMR experimental methods

All NMR experiments were carried out at 298 K on a 14.1 T (600 MHz) spectrometer equipped with a 5 mm z-axis gradient triple resonance room temperature probe. Gradients are calibrated using the known diffusion coefficient of residual HDO ($1902 \mu\text{m}^2 \text{s}^{-1}$) in a sample of D_2O at 298 K⁶³. The spectrometer temperature was calibrated with methanol- d_4 . An STE pulse sequence with variable diffusion delays was used in all diffusion experiments. In all experiments $\delta = 2$ ms and τ_2 , the time for which the magnetisation is transverse, is always set to

10 ms. For yeast sample, 11 quadratically spaced gradient strengths from 8-60 G cm⁻¹ were used and $\Delta = 20, 40, 60, 80, 100, 200, 300, 400, 500, 600, 700$ and 800 ms. For 3T3 sample, 11 quadratically spaced gradient strengths were used from 1-60 G cm⁻¹, $\Delta = 50, 100, 200, 300, 400$ and 500 ms. For HeLa cell samples, 11 quadratically spaced gradient strengths from 8-60 G cm⁻¹ were used and $\Delta = 50, 100, 200, 300, 400$ and 500 ms. All spectra were Fourier transformed and phased using NMRPipe⁶⁴. Data are analysed in phase-sensitive mode, with the lowest g value spectrum used determine phase parameters that are then applied to all other spectra. In each spectrum the water peak is integrated using Python scripts, making use of the nmrglue module⁶⁵, and water peak intensities were carried forward for analysis as described for the simulated datasets above.

H Cell preparation methods

Yeast cell samples were prepared according to previously published protocols^{46,66}. Briefly, dried yeast was mixed with tap water in a ratio of approximately 1:3 and the mixture was left for three days during which it was shaken periodically to allow for the release of carbon dioxide bubbles. The mixture is subsequently transferred to a 5 mm NMR tube with 5% D₂O added and forms a viscous homogeneous suspension that fills the active volume.

In the case of 3T3 cells, cells were grown in T-175 flasks and detached from their substrate by incubation with trypsin-EDTA for 5 min. Cells were then re-suspended in 250 μ L Dulbecco's phosphate buffered saline (DPBS) with 5% D₂O and transferred to a Shigemi tube for NMR experiments.

For HeLa cells, cells were grown in DMEM with 10% FBS (foetal bovine serum) in petri dishes and detached from plates using mechanical scraping (using petri dishes gives access to maximal area for scraping cells). Cells were

subsequently re-suspended in 250 μ L FBS, with 5% D₂O and transferred to a Shigemi tube for NMR experiments.

References

- ¹ F.X. Theillet, A. Binolfi, B. Bekei, A. Martorana, H.M. Rose, M. Stuiver, S. Verzini, D. Lorenz, M. Van Rossum, D. Goldfarb, and P. Selenko, *Nature* **530**, 45 (2016).
- ² G. Lippens, E. Cahoreau, P. Millard, C. Charlier, J. Lopez, X. Hanouille, and J.C. Portais, *Analyst* **143**, 620 (2018).
- ³ E. Luchinat and L. Banci, *Acc. Chem. Res.* **51**, 1550 (2018).
- ⁴ R. Hänsel, L.M. Luh, I. Corbeski, L. Trantirek, and V. Dötsch, *Angew. Chem. Int. Ed. Engl.* **53**, 10300 (2014).
- ⁵ L.A. Baker, T. Sinnige, P. Schellenberger, J. de Keyzer, C.A. Siebert, A.J.M. Driessen, M. Baldus, and K. Grünewald, *Structure* **26**, 161 (2018).
- ⁶ D. Le Bihan and H. Johansen-Berg, *Neuroimage* **61**, 324 (2012).
- ⁷ W.S. Price, *NMR Studies of Translational Motion* (Cambridge University Press, 2009).
- ⁸ P.T. Callaghan, *Principles of Nuclear Magnetic Resonance Microscopy* (Oxford University Press, 1993).
- ⁹ A.E. Baird and S. Warach, *J. Cereb. Blood Flow Metab.* **18**, 583 (1998).
- ¹⁰ P.D. Schellinger, G. Thomalla, J. Fiehler, M. Köhrmann, C.A. Molina, T. Neumann-Haefelin, M. Ribo, O.C. Singer, O. Zaro-Weber, and J. Sobesky, *Stroke* **38**, 2640 (2007).
- ¹¹ J.A. Jones, D.K. Wilkins, L.J. Smith, and C.M. Dobson, *J. Biomol. NMR* **10**, 199 (1997).
- ¹² A.J. Baldwin, J. Christodoulou, P.D. Barker, C.M. Dobson, and G. Lippens, *J. Chem. Phys.* **127**, 114505 (2007).
- ¹³ C.A. Waudby, M.D. Mantle, L.D. Cabrita, L.F. Gladden, C.M. Dobson, and J. Christodoulou, *J. Am. Chem. Soc.* **134**, 11312 (2012).
- ¹⁴ N. Shemesh, E. Özarslan, M.E. Komlosh, P.J. Basser, and Y. Cohen, *NMR Biomed.* **23**, 757 (2010).
- ¹⁵ R. Huang, J.P. Brady, A. Sekhar, T. Yuwen, and L.E. Kay, *J. Biomol. NMR* **68**, 249 (2017).
- ¹⁶ D. Le Bihan, J.-F. Mangin, C. Poupon, C.A. Clark, S. Pappata, N. Molko, and H. Chabriat, *J. Magn. Reson. Imaging* **13**, 534 (2001).
- ¹⁷ A.S. Verkman, A.N. van Hoek, T. Ma, A. Frigeri, W.R. Skach, A. Mitra, B.K. Tamarappoo, and J.

- Farinas, Am J Physiol **270**, C12 (1996).
- ¹⁸ A.S. Verkman, J. Membr. Biol. **173**, 73 (2000).
- ¹⁹ M. Palmgren, M. Hernebring, S. Eriksson, K. Elbing, C. Geijer, S. Lasič, P. Dahl, J.S. Hansen, D. Topgaard, and K. Lindkvist-Petersson, J. Membr. Biol. **250**, 629 (2017).
- ²⁰ E.O. Stejskal and J.E. Tanner, J. Chem. Phys. **42**, 288 (1965).
- ²¹ C.H. Neuman, J. Chem. Phys. **60**, 4508 (1974).
- ²² D.S. Grebenkov, *Diffusion NMR of Confined Systems* (Royal Society of Chemistry, Cambridge, 2016).
- ²³ D.S. Grebenkov, Rev. Mod. Phys. **79**, 1077 (2007).
- ²⁴ P.T. Callaghan, A. Coy, D. MacGowan, K.J. Packer, and F.O. Zelaya, Nature **351**, 467 (1991).
- ²⁵ J. Kärger, Ann. Phys. **479**, 1 (1969).
- ²⁶ J. Kärger, Adv. Colloid Interface Sci. **23**, 129 (1985).
- ²⁷ J. Kärger, H. Pfeifer, and W. Heink, Adv. Magn. Opt. Reson. **12**, 1 (1988).
- ²⁸ E. Fieremans, D.S. Novikov, J.H. Jensen, and J.A. Helpert, NMR Biomed. **23**, 711 (2010).
- ²⁹ C. Meier, W. Dreher, and D. Leibfritz, Magn. Reson. Med. **50**, 500 (2003).
- ³⁰ A. Szafer, J. Zhong, and J.C. Gore, Magn. Reson. Med. **33**, 697 (1995).
- ³¹ J. Pfeuffer, U. Flögel, W. Dreher, and D. Leibfritz, NMR Biomed. **11**, 19 (1998).
- ³² J. Andrasko, Biochim. Biophys. Acta **428**, 304 (1976).
- ³³ J. Andrasko, J. Magn. Reson. **21**, 479 (1976).
- ³⁴ P.W. Kuchel, A. Coy, and P. Stilbs, Magn. Reson. Med. **37**, 637 (1997).
- ³⁵ C. Meier, W. Dreher, and D. Leibfritz, Magn. Reson. Med. **50**, 510 (2003).
- ³⁶ G.J. Stanisz, J.G. Li, G.A. Wright, and R.M. Henkelman, Magn. Reson. Med. **39**, 223 (1998).
- ³⁷ G.J. Stanisz, A. Szafer, G.A. Wright, and R.M. Henkelman, Magn. Reson. Med. **37**, 103 (1997).
- ³⁸ E. Panagiotaki, S. Walker-Samuel, B. Siow, S.P. Johnson, V. Rajkumar, R.B. Pedley, M.F. Lythgoe, and D.C. Alexander, Cancer Res. **74**, 1902 (2014).
- ³⁹ M.A. Connell, P.J. Bowyer, P. Adam Bone, A.L. Davis, A.G. Swanson, M. Nilsson, and G.A. Morris, J. Magn. Reson. **198**, 121 (2009).
- ⁴⁰ J.S. Murday and R.M. Cotts, J. Chem. Phys. **48**, 4938 (1968).
- ⁴¹ J.E. Tanner and E.O. Stejskal, J. Chem. Phys. **49**, 1768 (1968).
- ⁴² N.N. Yadav and W.S. Price, Diffus. Fundam. **6**, (2007).

- ⁴³ L. Zhao, C.D. Kroenke, J. Song, D. Piwnica-Worms, J.J.H. Ackerman, and J.J. Neil, *NMR Biomed.* **21**, 159 (2008).
- ⁴⁴ R. Milo, P. Jorgensen, U. Moran, G. Weber, and M. Springer, *Nucleic Acids Res.* **38**, D750 (2010), BNID: 103788, 108913.
- ⁴⁵ W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes: The Art of Scientific Computing*, Third Edition (Cambridge University Press, 2007).
- ⁴⁶ I. Åslund, A. Nowacka, M. Nilsson, and D. Topgaard, *J. Magn. Reson.* **200**, 291 (2009).
- ⁴⁷ X. Tian, H. Li, X. Jiang, J. Xie, J.C. Gore, and J. Xu, *J. Magn. Reson.* **275**, 29 (2017).
- ⁴⁸ H. Li, X. Jiang, J. Xie, J.C. Gore, and J. Xu, *Magn. Reson. Med.* **77**, 2239 (2017).
- ⁴⁹ D.M. Yang, J.E. Huettner, G.L. Bretthorst, J.J. Neil, J.R. Garbow, and J.J.H. Ackerman, *Magn. Reson. Med.* **79**, 1616 (2018).
- ⁵⁰ S. Eriksson, K. Elbing, O. Söderman, K. Lindkvist-Petersson, D. Topgaard, and S. Lasič, *PLoS One* **12**, 1 (2017).
- ⁵¹ M. Zakhartsev and M. Reuss, *FEMS Yeast Res.* **18**, 1 (2018).
- ⁵² T. Srinorakutara, *Ferment. Bioeng. J.* **86**, 253 (1998).
- ⁵³ C. Labadie, J.H. Lee, G. Vetek, and C.S. Springer, *J. Magn. Reson. Ser. B* **105**, 99 (1994).
- ⁵⁴ P.C.M. Van Zijl, C.T.W. Moonen, P. Faustino, J. Pekar, O. Kaplan, and J.S. Cohen, *Proc. Natl. Acad. Sci.* **88**, 3228 (1991).
- ⁵⁵ V.A. Makarov, B.K. Andrews, P.E. Smith, and B.M. Pettitt, *Biophys. J.* **79**, 2966 (2000).
- ⁵⁶ E. Persson and B. Halle, *Proc. Natl. Acad. Sci. U. S. A.* **105**, 6266 (2008).
- ⁵⁷ B. Balinov, B. Jönsson, P. Linse, and O. Söderman, *J. Magn. Reson. - Ser. A* **104**, 17 (1993).
- ⁵⁸ J.G. Powles, M.J.D. Mallett, G. Rickayzen, and W.A.B. Evans, *Proc. R. Soc. A Math. Phys. Eng. Sci.* **436**, 391 (1992).
- ⁵⁹ C.A. Waudby and J. Christodoulou, *J. Magn. Reson.* **211**, 67 (2011).
- ⁶⁰ H. Xing, F. Lin, Q. Wu, and Q. Gong, *Magn. Reson. Med.* **70**, 1167 (2013).
- ⁶¹ M. Newville, T. Stensitzki, D.B. Allen, and A. Ingargiola, (2014).
- ⁶² P.T. Callaghan, A. Coy, T.P.J. Halpin, D. MacGowan, K.J. Packer, and F.O. Zelaya, *J. Chem. Phys.* **97**, 651 (1992).
- ⁶³ M. Holz and H. Weingartner, *J. Magn. Reson.* **92**, 115 (1991).
- ⁶⁴ F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, and A. Bax, *J. Biomol. NMR* **6**, 277 (1995).

⁶⁵ J.J. Helmus and C.P. Jaroniec, J. Biomol. NMR **55**, 355 (2013).

⁶⁶ I. Åslund and D. Topgaard, J. Magn. Reson. **201**, 250 (2009).