



HAL
open science

A Unified Description of Intrinsically Disordered Protein Dynamics under Physiological Conditions Using NMR Spectroscopy

Wiktor Adamski, Nicola Salvi, Damien Maurin, Justine Magnat, Sigrid Milles, Malene Ringkjøbing Jensen, Anton Abyzov, Christophe J Moreau, Martin Blackledge

► To cite this version:

Wiktor Adamski, Nicola Salvi, Damien Maurin, Justine Magnat, Sigrid Milles, et al.. A Unified Description of Intrinsically Disordered Protein Dynamics under Physiological Conditions Using NMR Spectroscopy. *Journal of the American Chemical Society*, 2019, 141 (44), pp.17817-17829. 10.1021/jacs.9b09002 . hal-02414561

HAL Id: hal-02414561

<https://hal.univ-grenoble-alpes.fr/hal-02414561>

Submitted on 23 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Unified Description of Intrinsically Disordered Protein Dynamics under Physiological Conditions using NMR Spectroscopy

Wiktor Adamski, Nicola Salvi, Damien Maurin, Justine Magnat, Sigrid Milles, Malene Ringkjøbing Jensen, Anton Abyzov, Christophe J. Moreau, Martin Blackledge*

Institut de Biologie Structurale,

Université Grenoble Alpes-CEA-CNRS

71, Avenue des Martyrs

Grenoble

France

ABSTRACT

Intrinsically disordered proteins (IDPs) are flexible biomolecules whose essential functions are defined by their dynamic nature. Nuclear magnetic resonance (NMR) spectroscopy is ideally suited to the investigation of this behaviour at atomic resolution. NMR relaxation is increasingly used to detect conformational dynamics in free and bound forms of IDPs under conditions approaching physiological, although a general framework providing a quantitative interpretation of these exquisitely sensitive probes as a function of experimental conditions is still lacking. Here, measuring an extensive set of relaxation rates sampling multiple-timescale dynamics over a broad range of crowding conditions, we develop and test an integrated analytical description that accurately portrays the motion of IDPs as a function of the intrinsic properties of the crowded molecular environment. In particular we observe a strong dependence of both short-range and long-range motional timescales of the protein on the friction of the solvent. This tight coupling between the dynamic behavior of the IDP and its environment allows us to develop analytical expressions for protein motions and NMR relaxation properties that can be accurately applied over a vast range of experimental conditions. This unified dynamic description provides new insight into the physical behavior of IDPs, extending our ability to quantitatively investigate their conformational dynamics under complex environmental conditions, and accurately predicting relaxation rates reporting on motions on timescales up to tens of nanoseconds, both *in vitro* and *in cellulo*.

INTRODUCTION

Intrinsically disordered proteins (IDPs) occur throughout biology,¹ where they represent a high proportion of proteomes from all domains of life.^{2,3} IDPs are characterized by the high flexibility of their functional state, encoded by their primary amino acid sequence, that allows the protein to sample a quasi-continuum of rapidly interconverting conformations.⁴⁻⁶ Extreme conformational heterogeneity endows IDPs with considerable advantages over their folded counterparts, for example allowing them to interact with multiple partners, and to exploit transient and local disorder-to-order transitions with rapid association and dissociation rates that allow for efficient processing. These characteristics highlight the importance of dynamics for the function of IDPs. The relationship between conformational dynamics and molecular function of IDPs in their physiological environment remains however difficult to characterize experimentally.

Nuclear magnetic resonance (NMR) spectroscopy is the method of choice for studying protein dynamics at atomic resolution. In particular spin relaxation is sensitive to fluctuations of backbone and side-chain moieties and provides unique insight into dynamics throughout the molecular chain. Although considerable effort has been devoted to accurately describing the conformational space sampled by IDPs in solution,⁷⁻¹² mainly using NMR spectroscopy, small angle scattering or single molecule FRET,¹³⁻¹⁵ the development of a physical understanding of the dynamic modes and timescales characterizing interconversion between the distinct sub-states remains challenging.¹⁶⁻²² We recently investigated the temperature dependence of the dynamics of a 125 amino acid IDP, N_{TAIL}, using an extensive set of temperature and magnetic field dependent ¹⁵N relaxation measurements.²³ Using an Arrhenius-type relationship to relate dynamic correlation times to temperature

$$\tau_k(T) \propto e^{\frac{E_{a,k}}{RT}} \quad (1)$$

we analyzed up to 62 ¹⁵N relaxation rates per ¹⁵N-¹H bond, over a temperature range of 30K, to determine individual activation energies (E_a) of distinct dynamic modes for each amino acid in the protein. This allowed us to identify three distinct dynamic contributions to ¹⁵N relaxation, sampling dynamics on timescales from tens of picoseconds to tens of nanoseconds. All three components were necessary to adequately describe the dynamic properties of each backbone amide site in the protein. The physical origin of these three modes were attributed to fast (~50 ps) motions within a flat potential whose timescales are independent of temperature, intermediate motions (in the nanosecond range) reporting on backbone dihedral angle fluctuations, and slower motions (around 10 ns), with higher activation energies, that report on segmental, chain-like dynamics. This model was shown to accurately reproduce independent data sets and further inspired the development of analytical approaches to simulate IDP dynamics using MD-based approaches.^{24,25}

Despite this progress, it was not clear how to transfer this framework into the complex physiological environments that are increasingly exploited to investigate IDP function, for example *in cellulose*.²⁶ In cells IDPs operate in extremely crowded environments,²⁷⁻²⁹ with molecular concentrations reaching 400

g/L, considerably modifying viscosity on different length scales, and therefore very likely strongly affecting the timescales of IDP dynamics.³⁰⁻³² Although single molecule fluorescence has been used to investigate the relative importance of internal and solvent friction on the dynamics of IDPs and denatured states and on the kinetics of protein folding,³³⁻³⁵ as well as the impact of the cellular environment on the dynamics of IDPs³⁶ and protein function,³⁷ the relationship between primary sequence and protein flexibility in complex crowded environments remains poorly understood. This is particularly important for IDPs that exhibit extreme solvent accessibility and for which the physiological environment found in complex multi-component environments, or even live cells, can be expected to strongly impact dynamic modes and timescales.

In-cell NMR has also been successfully applied to the study of both folded and disordered proteins,³⁸⁻⁴⁴ allowing for the investigation of structure, folding, maturation and dynamics of proteins in live cells. To interpret NMR relaxation measured in physiological environments it is necessary to incorporate the impact of molecular crowding into any analytical model describing the relationship between measured rates and dynamic modes and timescales.^{45,46} In the current work we develop a unified model of IDP dynamics as a function of environmental conditions that can be used to interpret relaxation data measured in complex molecular settings. Systematically modifying viscosity, using inert crowding agents, as well as temperature, we measure an extensive set of NMR relaxation rates over a broad range of environmental conditions. This allows us to map the dynamics of two independent IDPs, comprising over 210 amino acids, as a function of environment, incorporating both temperature and nanoscale viscosity under a range of crowding conditions. On this basis we develop and test an integrated analytical description that relates the motional modes and timescales characterizing IDP dynamics as a function of the intrinsic properties of the molecular milieu, as reported via the dynamics of the surrounding solvent. This model not only allows for the highly accurate analysis of relaxation rates measured in different crowded environments *in vitro*, but also reproduces experimental ¹⁵N relaxation rates measured in living cells, suggesting a promising method for assessing functionally important motional modes and dynamic timescales of IDPs in very different complex environments.

RESULTS

¹⁵N backbone relaxation rates of the 125 amino acid intrinsically disordered C-terminal domain of the nucleoprotein from Sendai virus (N_{TAIL}), comprising a partially folded helical element between residues 478 and 490,⁴⁷ were measured as a function of crowding conditions. The polysaccharide dextran with average molecular weight of 40 kDa, which was inert in admixtures of N_{TAIL} (supplementary figure S1), was used as a molecular crowder. Longitudinal (R_1), transverse (R_2), cross-relaxation (heteronuclear nOe) and transverse cross-correlated dipole-dipole/CSA (η_{xy}) ¹⁵N relaxation rates were measured at 288 and 298K, over a range of dextran concentrations from 50-195 g/L and at three magnetic field strengths (figure 1, supplementary figure S2). The differential effects on R_1 and R_2 along the sequence testify to the heterogeneous response of the different regions of the protein to crowding.

Solvent nanoscale viscosity was probed by measuring longitudinal relaxation of water protons.⁴⁸ Although exchange effects between free water and water molecules bound to large polymers and translational diffusion effects in the vicinity of surfaces may give rise to relaxation dispersion effects at low magnetic fields,^{49–51} at the high magnetic fields used here rotational diffusion is expected to dominate R_1 . Changes in viscosity as a function of crowding, were gauged using the following expression for solvent friction:

$$\rho(C) = (\eta_C - \eta_0)/\eta_0 = (R_{1,C} - R_{1,0})/R_{1,0} \quad (2)$$

where $R_{1,0}$ and η_0 are the longitudinal relaxation of water and the viscosity in the absence of viscogen respectively, and η_C is the viscosity of the sample of interest (measured at concentration C).

Admixtures containing up to 145 g/L of dextran at any measured temperature (278-298K) exhibit similar $\rho(C)$ curves (figure 2b). Beyond 145g/L the slope begins to deviate from a linear dependence, signifying transition to the semi-dilute regime, where the concentration becomes sufficiently high that the solute molecules interact with, and permeate each other's hydrodynamic volume. Such regimes may be of interest in, for example, membraneless organelles formed by IDPs in solution, or complex cellular environments.

Analytical definitions of the measured ¹⁵N relaxation rates are given in the methods section. The spectral density function is defined using the ‘model-free’ approach,⁵² as the sum of $k=(1,2,3)$ Lorentzian contributions reporting on motions occurring on distinct timescales τ_k :

$$J(\omega) = \sum_k \frac{A_k \tau_k}{1 + \omega^2 \tau_k^2} \quad (3)$$

where A_k represents the amplitudes of the three dynamic components previously characterized as a function of temperature.²³ Our earlier study of N_{TAIL} revealed that τ_1 reports on very fast (~ 50 ps) librational motions that are independent of temperature and were treated here in a similar way (A_1 was optimised and τ_1 fixed at 50 ps). Analysis of the individual experimental data sets measured at three different concentrations of viscogen (50, 95 and 145 g/L) reveals that the amplitudes of the three components did not vary significantly with respect to crowding conditions (supplementary figure S4), indicating that the presence of the crowder does not significantly affect the conformational sampling of the peptide chain.

Determination of the slow and intermediate timescales at each viscogen concentration reveals a strong dependence on viscosity for the slowest component (τ_3), with a weaker dependence of the correlation time of the intermediate component (τ_2) that mirrors the correlation time dependence of the water (figure 2c, d and supplementary figure S3) (the average concentration dependence of τ_3 is approximately three times that of τ_2 (figure 2c, d)). This distinct dependence on the concentration of viscogen can be understood on the basis of well-documented relationship of measured effective viscosity on the dimensions of the probe (r_p), in the limit when r_p is significantly smaller than the dimensions of the

viscogen.^{53–59} In this respect the observation of different viscosity coefficients for τ_2 and τ_3 suggests that local and longer-range motional modes report on structural elements exhibiting different dimensions.

In order to interpret all of the measured data from N_{TAIL} simultaneously as a function of viscosity, we exploited the Einstein equation:⁶⁰

$$\tau_k(C) = \tau_{k,\infty}(\varepsilon_k \rho(C) + 1) \quad (4)$$

where $\tau_{k,\infty}$ is the correlation time at infinite dilution. The solvent friction is defined by the measurement of R_1 of water as a function of viscogen concentration (see above) and ε_k is a friction coefficient, relative to $\rho(C)$, that allows us to define the dependence of each correlation time on viscosity. Not surprisingly, given the distinct dependences of τ_2 and τ_3 (figure 2c, d), a simple model using a common ε for all three motional modes fails to adequately reproduce experimental data, showing underestimation of the relaxation rates that are dominated by slower motions, and inaccurate reproduction of rates dominated by intermediate motion (data not shown). We therefore propose a model allowing for distinct friction coefficients ε_k for the different motional modes exhibited by the protein backbone, so that scaling parameters are defined for each of the contributions to the total spectral density function (equation 3).

The correlation times shown in equation 4 were incorporated into the model-free analysis using equation 3. The amplitudes of the motion were held constant (*vide supra*) over the crowding conditions, but allowed to vary freely within the constraints of $\sum_k A_k = 1$, so that in total, 8 parameters (A_2 , A_3 , τ_2 , τ_3 , θ , ε_1 , ε_2 and ε_3) were initially optimized for each site in the protein. Values of ε_2 and ε_3 (corresponding to intermediate and slow timescale motions respectively) cluster around 1.0 ± 0.3 and 2.5 ± 0.8 respectively, while ε_1 shows very poorly defined distribution across the sequence. The viscosity-dependent analysis was repeated using a value of $\varepsilon_1 = 0$, resulting in indistinguishable data reproduction and effectively identical values for ε_2 and ε_3 , suggesting that the fastest component is not detectably dependent on concentration. All analysis described below therefore individually optimizes the parameters $\varepsilon_{k=2,3}$ and assumes a value of $\varepsilon_1 = 0$, meaning that the fast timescale does not evolve significantly with viscogen concentration and is fixed at 45 ps.

Importantly, the dynamic parameters (A_k and τ_k) determined throughout the protein using data acquired as a function of molecular crowding (here) or as a function of sample temperature (previous study)²³ are remarkably similar (supplementary figure S5), when compared under a common set of conditions between the two studies (absence of viscogen and 298K).

A General Expression for Rotational Correlation Times in IDPs as a Function of Environmental Conditions

The similarity of the dynamic parameters resulting from temperature- and viscogen-dependent analyses at zero viscogen concentration indicates a compatibility of the two approaches, and encouraged us to propose a combined expression that simultaneously describes the correlation times of the dynamic modes as a function of both temperature and viscosity. As shown in figure 2, the correlation times of

both intermediate and slow dynamic components appear to be correlated to solvent friction, based on the observation that the two components respond with distinct sensitivities to viscocon concentration (around 1 and 3 times the solvent friction). We therefore propose the following combined expression:

$$\tau_k(C, T) = \tau'_{k,\infty} (\varepsilon_k \rho(C) + 1) e^{\frac{E_{a,k}}{RT}} \quad (5)$$

where $E_{a,k}$ is the activation energy of intermediate and slow motions. In this case $\tau'_{k,\infty}$ represents the correlation time at infinite dilution and infinite temperature. We use this expression to simultaneously analyze 105 relaxation data sets measured over a temperature range of 274-298K and 0-195 g/L dextran concentration for each site in the protein. Note that solvent friction is non-linear with respect to C beyond 150 g/L and that this behavior is mirrored by backbone relaxation rates. Linking the protein dynamic components to solvent friction rather than viscocon concentration allows us to analyse ^{15}N relaxation data beyond the linear range, which represents a key advantage for studies carried out in more complex environments or *in cellulo* where the effective crowder concentration may not be known.

Fifteen parameters are optimized per amino acid ($\varepsilon_3, \varepsilon_2, \tau'_{3,\infty}, \tau'_{2,\infty}, E_{a,3}, E_{a,2}, \theta$ and A_2 and A_3 at each temperature), with a maximum of 90 degrees of freedom. The results are shown in figure 3. All data are well reproduced using this combined description of the temperature and viscosity dependent dynamic behavior of N_{TAIL} . When 30% of relaxation rates were randomly removed they were predicted with comparable accuracy with respect to the direct fit (supplementary figure S6). The activation energies of the two processes are well separated in the central part of the protein, and converge at the termini, as previously seen in the uniquely temperature-dependent analysis.²³ ε_2 and ε_3 sample distinct ranges that again differ by a factor of approximately 3, with ε_3 showing a marked sequence dependence, for example the dip around residue 415 corresponding to a potentially highly flexible region comprising three sequential glycines. The friction coefficient corresponding to local backbone dihedral motions ε_2 is relatively flat across the primary sequence, again suggesting that it predominantly reports on motions of individual peptide planes.

Experimental data fitted to equation 5 therefore allow us to map the dynamic behavior of N_{TAIL} , in particular the intermediate backbone and slow segmental motion of each amino acid in the chain, as a function of temperature and solvent friction. Figure 4 shows two amino acids from the fully disordered and partially helical regions respectively.

Unified Expression is Quantitatively Predictive of Data Measured Under Independent Conditions

In order to test the predictive nature of equation 5 to describe the dynamics of IDPs as a function of sample crowding and temperature, we removed an entire set of relaxation rates, measured under conditions where none of the remaining fitted data were measured. 12 relaxation rates (R_1, R_2, nOe and η_{xy} measured at 700 MHz from N_{TAIL} at 288K at dextran concentrations of 50, 95 and 145 g/L) were calculated on the basis of the 15 parameters determined from the analysis described above (this time only comprising crowding data from N_{TAIL} at 298K and non-crowded solutions at temperatures from

274-298K). The prediction of the experimental data is essentially indistinguishable from the reproduction of the same data when actively fitted (supplementary figure S7).

We have also tested the transferability of the model by measuring R_1 , R_2 , nOe and η_{xy} for N_{TAIL} in 90 g/L PEG (10 kDa) (measurements made at 1H frequency of 600 MHz and 298K) and 135 g/L (measurements made at 1H frequency of 850 MHz and 298K) PEG (10 kDa). Data were predicted using equation 5, incorporating friction coefficients (ϵ_2 and ϵ_3), activation energies ($E_{a,3}$, $E_{a,2}$) and motional amplitudes (A_2 , A_3) determined using dextran as viscogen (figure 3), and $\rho(C)$ determined from $R_{1,water}$ for both sample conditions. Remarkably, experimental data are reproduced effectively within the estimated uncertainty for all four relaxation rates at both concentrations and magnetic field strengths (figure 5), and are quite distinct from rates predicted under dilute conditions.

These results testify to the general applicability of equation 5 to describe the rotational correlation times affecting ^{15}N spin relaxation in IDPs, and more generally their dynamics, over a broad range of environmental conditions, accurately capturing the impact of temperature and viscosity on the timescale of intermediate and slow backbone motions, and suggest that the strong coupling between IDP dynamic timescales and the dynamic properties of the solvent implicit in equation 5 is experimentally verified across distinct crowding conditions.

Dynamics of the N-terminal IDR of MKK4 as a Function of Environmental Conditions

We further tested the general applicability of equation 5 by investigating the dynamic behavior of the 86 amino-acid disordered N-terminal domain of mitogen-activated kinase kinase (MKK4). The conformational sampling of this domain was recently characterized by NMR,⁶¹ revealing a highly flexible N-terminal region, rich in glycines and serines, followed by an extended p38 α interaction motif and a C-terminal region comprising more bulky side-chains. The disordered domain contains no significantly populated secondary structure and comprises segments of primary sequence exhibiting distinct dynamic characteristics.

Experimental relaxation rates were again measured at four magnetic field strengths (600, 700, 850 and 950 MHz) and temperatures ranging from 273 to 288K (figure 6a, c). Twelve relaxation rates (R_1 , nOe , η_{xy} and η_z) were measured under similar crowding conditions as N_{TAIL} , using dextran as crowding agent at 50 and 95 g/L and two temperatures (273 and 278K) (figure 6b). $R_{1,water}$ was measured as a function of viscogen concentration (supplementary figure S8) showing a very different (more than 50% greater) solvent friction coefficient compared to N_{TAIL} . Initial analysis of the temperature-dependent dilute data following the procedure applied for N_{TAIL} confirmed that three distinct contributions are again required to correctly model the experimental data at each temperature. 30% of data removed from the analysis are accurately predicted when three contributions to equation 3 are used but not with two (supplementary figure S9). The fastest motion again showed negligible temperature dependence and was fixed to 45 ps at all temperatures. MKK4 is the second IDP for which temperature-dependent ps-ns dynamics have been described at amino-acid specific resolution using an Arrhenius-type approach to couple the

motional modes over the temperature range, substantiating previous conclusions concerning the three observed modes.²³

The data measured under crowding conditions were combined with the temperature-dependent relaxation rates in the absence of viscogen (in total 68 ¹⁵N relaxation rates) and simultaneously analysed using equation 5, again optimizing the same 15 parameters that were determined for N_{TAIL}. The resulting dynamic parameters (amplitudes, timescales and activation energies) (figure 7) are very similar to those determined using only temperature-dependent data (supplementary figure S10), exhibiting a clear separation of E_a of the slow and intermediate components. E_a of the intermediate timescales fall in a similar range for backbone dihedral angle fluctuations (8-10 kJ mol⁻¹) while the slowest timescale again falls in the range 20-25 kJ mol⁻¹ (except in the glycine/serine rich region where the activation energies are closer together, a phenomenon also seen at the flexible termini of N_{TAIL}). The amplitude of the slowest motion shows a bell-shaped distribution with respect to primary sequence, reaching a maximum at the p38 α interaction site and the bulky region. The intermediate timescale motion is much flatter than the slower component.

Length-scale Dependent Viscosities are Closely Related to Solvent Friction

Similarly to N_{TAIL}, the friction coefficients of the slow and intermediate dynamic modes ϵ_2 and ϵ_3 are found to sample distinct ranges around 1, again mirroring the solvent friction, and 3.5 respectively (figure 7), despite absolute differences in the value of $\rho(C)$ in the two experimental systems (more than 50% greater for MKK4). These differences are likely related to buffer composition (pH 6 and 500 mM NaCl for N_{TAIL} in phosphate buffer compared to pH 7 and 150 mM NaCl for MKK4 in HEPES), and possibly solvent-solute interactions, but strikingly the friction coefficients of the protein backbone fall in very similar ranges.

Prediction of the Dynamic Properties of Intrinsically Disordered Proteins in Live Cells

In addition to the improved insight into the dependence of polypeptide dynamics on the intrinsic properties of the solution experienced by solvent molecules, the consequences of the general applicability of equation 5 are potentially far-reaching. Using this framework it should be possible to predict dynamic correlation times in any particular medium for which ¹⁵N relaxation have been measured *in vitro* (i.e. where sequence-specific values of ϵ_k and E_{a,k} have been determined), assuming that the solvent friction $\rho(C)$ can be determined from $R_{1,water}$ under the more complex conditions.

To further test this idea, we have recorded η_{xy} and heteronuclear nOe relaxation rates in MKK4 microinjected into *Xenopus laevis* oocytes (292K and 700 and 600 MHz respectively) (supplementary figure S11). Although a number of residues are absent due to line-broadening, probably due to slowing down of segmental motions in particular between the central region of the domain and the bulky C-terminal strand, and magnetic field inhomogeneity typical of in-cell NMR experiments, the quality of the spectrum is nevertheless sufficient to allow determination of reproducible relaxation rates for over half of the protein. Cell viability was monitored by comparing ¹⁵N edited ¹H spectra of the sample

throughout the measurement period, and of the supernatant immediately before and directly after acquisition (supplementary figure S12). Relaxation measurements were acquired during the period when the spectra were found to be the most stable (24-50 hours after injection). The dynamic modes of MKK4 *in cellulo* were predicted using equation 5, with $\rho(C)$ estimated from experimental measurements considering the measured ^1H R_1 in buffer and *in cellulo* ($\rho(C)=1.00\pm 0.08$). η_{xy} experiments were repeated twice using independent batches of cells, whereas heteronuclear nOe were performed with one batch. Activation energies, $\tau_{k,\infty}$ and ε_k were taken directly, and amplitudes linearly extrapolated to 292K from the analysis described above *in vitro*. The experimental measurements are compared to prediction in figure 8, and compared to the values predicted at the same temperature in standard buffer. The reproduction of the experimental data is remarkable considering the complexity of the experimental system and the approximations implicit in the model. These results suggest a general applicability of this approach for the study of protein dynamics using NMR relaxation in complex mixtures and *in cellulo*.

DISCUSSION

IDPs are highly dynamic, and this flexibility is expected to play an essential role in their function. In order to understand the relationship between dynamics and function it is important to develop approaches to accurately describe their motion under physiological conditions. Numerous studies have demonstrated the feasibility of measuring NMR parameters, and even heteronuclear relaxation, *in cellulo*, and this exciting tool promises to gain in popularity following the pioneering work of a number of groups over the last two decades.^{40,43,41,62} In addition, a growing number of experimental systems allow for the study of flexible IDPs in the presence of significant molecular crowding, for example in large reconstituted complexes, underlining the importance of understanding the dependence of NMR parameters on protein dynamics under both crowded and *in vivo* conditions.

Using NMR relaxation measured from two long IDPs (more than 210 amino acids in total) exhibiting highly different primary sequence characteristics and dynamic fingerprints, we develop a general relationship between experimental measurement and IDP dynamics over a broad range of environmental conditions. By systematically varying viscosity and sample temperature and measuring an extensive set (113 and 68 independent rates per residue for N_{TAIL} and MKK4 respectively) of ^{15}N relaxation rates that are sensitive to correlation times spanning nearly three orders of magnitude, we develop and test a general relationship between experimental measurement and protein dynamics over a broad range of environmental conditions. The dependence on concentration of crowder and temperature are found to be complementary and coherent, resulting in consistent descriptions of IDP dynamics at zero viscogen concentration whether independently analyzing data that vary as a function of temperature or viscogen concentration. This allows us to propose a unified model of IDP dynamics in complex environments, incorporating both temperature and nano- and sub-nanoscale viscosity into a single expression for the rotational correlation times of the different components to the motion (equation 5) for each amino acid in the protein. Fitting the experimental data to this expression results in close reproduction of

experimental data from both proteins, and accurately estimates temperature ($E_{a,k}$) and viscosity (ε_k) coefficients, as well as the timescales and amplitudes (A_k) of the motions under the different environmental conditions.

The intermediate and slow dynamic contributions measured on the protein backbone respond very differently to changes in viscosity, resulting in distinct ranges of friction coefficients, in line with expected behavior from viscosity probes of different dimensions.⁵³⁻⁵⁷ As noted above, the friction coefficients (ε) would be expected to decrease for smaller probes, as described in recent studies that have parameterized analytical relationships between microscopic viscosity and probe dimensions in complex mixtures and across a range of length scales.^{58,59} Assuming that the dynamic modes associated with timescales τ_2 and τ_3 originate from motions of groups of atoms, or segments having different dimensions (a single peptide plane or multiple planes respectively), these contributions may be expected to be characterized by distinct length scales and therefore distinct friction coefficients.

The friction coefficient of the intermediate timescale backbone motion, ε_2 is very similar for MKK4 and N_{TAIL}, closely mirroring solvent friction in both cases, even though the absolute values of solvent friction are very different in the two experimental systems. This suggests that backbone dihedral angle motions report on nanoscale viscosity (the dimensions of the water molecule) and indicates that their associated length-scales are similar and probably on the level of the individual amino acid. Friction coefficients associated with the slower motion, ε_3 , lie in a range three times greater than the intrinsic solvent friction for both proteins, appearing to report on concomitantly longer effective length-scales, likely involving concerted motions of segments of the backbone comprising more than one amino acid. In relation to this, ensemble-based molecular dynamics simulation^{25,63} and analysis of RDCs in N_{TAIL}⁶⁴, reveal that segments of three or more amino acids exhibit coupled motional modes throughout the protein backbone on timescales (ns) that contribute to measured ¹⁵N relaxation rates. Interestingly, the intrinsic friction of the slower motions shows a clear sequence dependence, with for example three sequential glycines in N_{TAIL}, that were predicted from MD simulation to form the shortest segments, also exhibiting the lowest local ε_3 . The viscosity-dependent analysis therefore clearly assigns intermediate and slow backbone components of ¹⁵N relaxation as reporting on the dynamic properties of individual peptide planes and collective segments respectively, substantiating our previous assignment of these modes on the basis of temperature-dependent analysis.

The expression is remarkably robust, predicting entirely independent experimental data sets measured under conditions that are not included in the analysis. Once the site-specific parameters ($E_{a,k}$, ε_k and A_k) have been determined, equation 5 is transferable between solutions comprising viscogens of very different nature (dextran and PEG), even including the complex mixture of microenvironments found *in cellulo*, simply by measuring ¹H relaxation rates of water that report on the nanoscale viscosity of the sample. This approach accurately predicts relaxation rates that are exquisitely sensitive to dynamics occurring on timescales spanning nearly 3 orders of magnitude. The apparently general nature of the description also highlights the importance of solvent behavior for the accurate description of IDP

dynamics, revealing a remarkably tight coupling between the dynamics of the protein and solvent.⁶⁵

CONCLUSIONS

In summary, we present a simple expression relating characteristic timescales, activation energies, amplitudes and viscosity coefficients that is capable of simultaneously describing the dynamics of IDPs at vastly different viscosities using a minimal set of physical parameters, providing unique insight into the dynamic processes that give rise to NMR relaxation in IDPs. These physical parameters are derived from ¹⁵N relaxation *in vitro*, allowing the prediction of relaxation rates for an IDP of interest under any given conditions where the intrinsic solvent friction of the experimental environment can also be estimated. The method was applied to over 150 sites from two long IDPs, exhibiting distinct sequence and structural characteristics, and is found to be robust and transferable between different experimental conditions, predicting independent data sets measured under conditions of crowding and temperature that were entirely absent from the active fitted data set and reporting on timescales spanning three orders of magnitude. Remarkably, when applied to relaxation measured over different samples of MKK4 injected into xenopus oocytes, the relationship also reproduces the relaxation rates measured in the protein *in cellulo*.

This unified description of IDP dynamics under complex environmental conditions provides a physical interpretation of NMR relaxation rates measured in crowded environments *in vitro* and *in vivo* and reveals a strong and quantifiable dependence of IDP dynamics on the properties of the solvent in the same environment. The framework will advance the application of NMR spectroscopy to the study of the dynamics of IDPs in increasingly complex environments, for example in membraneless organelles, reconstituted multi-component assemblies or *in cellulo*.

MATERIALS AND METHODS

Sample preparation

Protein expression and purification, N_{TAIL}. The C-terminal domain of Sendai virus nucleoprotein (N_{TAIL}) comprising residues 401–524 strain Harris corresponding to SeV Fushimi strain (UniProtKB accession number Q07097) except for mutation E410K, were expressed and purified as previously described.⁶⁶ The NMR experiments were performed in 50 mM phosphate buffer, pH 6.0, 500 mM NaCl, (further referred to as N_{TAIL} buffer).

Protein expression and purification, MKK4. The intrinsically disordered regulatory domain of MKK4 (residues 1-86, UniProtKB accession number P45985), was expressed and purified as described.⁶¹ NMR experiments were performed in 50 mM HEPES pH 7.0, 150 mM NaCl, 2 mM dithiothreitol (DTT) further referred to as MKK4 buffer.

Preparation of crowder solutions. Concentration of Dextran40 (BioChemica, AppliChem A2249) was estimated by measuring absorbance at 270 nm (A_{270}) on a ThermoScientific Genesys 10uv Scanning spectrophotometer in UV-Cuvettes micro 70 μ l. The extinction coefficient of Dextran40 was estimated by probing A_{270} of different dilutions of Dextran40 stock with known concentration

of the polymer. Briefly, 3g of the polymer was dissolved in 25 ml buffer of interest in volumetric flask over night at approximately 318K and subsequently diluted to 120 g/L. The obtained stock was diluted to reach concentrations in the range of 12 to 120 g/L of Dextran40. Estimated extinction coefficient yielded $4.96 \text{ cm}^{-1}\text{mLg}^{-1}$, so that 100 g/L of Dextran40 corresponded to approximately A_{270} of 0.5 (supplementary figure S13). Stocks of Dextran40 were prepared by dissolving 0.5 g of the polymer in 1 ml of buffer of interest in a water bath (318K) followed by vigorous mixing and over-night equilibration at room temperature.

An analogous dissolution procedure was applied to PEG 10 kDa (Sigma 81280), however as PEG does not have a clear absorption band above 240 nm, the concentration of equilibrated stock of polymer could not be determined exactly and was estimated to be 400 g/L.

Xenopus laevis oocyte injections. *Xenopus* oocytes were prepared as reported previously.⁶⁷ Briefly, after surgical retrieval, oocytes were defolliculated enzymatically with type 1A collagenase (2 mg/ml) for approximately 2 hours in modified Barth's solution without calcium (1 mM KCl, 0.82 mM MgSO₄, 88 mM NaCl, 2.4 mM NaHCO₃, 16 mM Hepes and pH 7.4). Cells were then washed several times and for at least 20 min with the same buffer without collagenase and stored in the modified Barth's solution with calcium (0.41 mM CaCl₂ and 0.3 mM Ca(NO₃)₂). Oocytes in stage V and VI were manually selected and stored overnight at 292K in the modified Barth's solution with calcium supplemented with 100 U ml⁻¹ of penicillin, streptomycin and 100 mg/L of gentamycin. The following day, selected oocytes were placed in a small petri dish containing a grid for injection. Oocytes were oriented with the animal pole upward for injection of labelled proteins in this pole. Micro-injection was performed manually using a Nanoject II (Drummond) and 250 oocytes were each injected with 50 nl of protein solution. The injected oocytes were pooled in a 15-ml tube and gently washed once with the same buffer before storage at 292K prior to NMR measurement. Animal handling and experiments are fully conformed with European regulations and were approved by the French Ministry of Higher Education and Research (APAFIS #4420-2016030813053199 to Christophe Moreau). Authorization of the animal facility has been delivered by the Prefect of Isere (Authorization # D3818510001).

NMR Spectroscopy

NMR experiments were performed on Bruker spectrometers interfaced to cryoprobes operating at ¹H frequencies of 600, 700, 850 and 950 MHz). Temperature calibration was carried out as described,²³ all samples were equilibrated for 20 minutes before tuning/matching and shimming. Under crowding conditions, ¹⁵N spin relaxation rates were measured at 298 (600 and 850MHz) and 288K (700 MHz) for N_{TAIL} and 273 (600 MHz) and 278K (600 or 700 MHz) for MKK4. Spin relaxation rates at all four magnetic field strengths were also measured at 273, 278, 283 and 288 K for MKK4 in buffer. Concentration of N_{TAIL} was approximately 300 μM and 1.3 mM for MKK4.

¹⁵N $R_{1\rho}$ (using a spin lock of 1.5 kHz), R_1 and $\{^1\text{H}\}$ -¹⁵N heteronuclear nOe were measured as described.^{23,68} ¹⁵N R_2 was determined from $R_{1\rho}$ and R_1 .⁶⁹ ¹⁵N-¹H CSA/DD transverse cross-correlated (η_{xy}) and longitudinal (η_z) relaxation rates were measured as described.^{70,71} Recycle delays of 1-2.5s and

64-128 dummy scans were used for R_1 , $R_{1\rho}$, η_{xy} and η_z . Typical set of relaxation delays included points measured at [1, 30, 50, 70, 90, 130, 170, 210 and 250] ms for $R_{1\rho}$ and [0, 0.08, 0.2, 0.4, 0.6, 0.8, 1.04, 1.6 and 1.9] s for R_1 (including repetition of one delay) Cross-relaxation delays were set to 50 ms (N_{TAIL}) and 60 ms (MKK4) for η_{xy} and 100 ms for η_z . All *in vitro* measurements were performed in 3 mm NMR tubes containing 10% D₂O.

All spectra were processed in NMRPipe⁷² and analyzed in NMRFAM-Sparky.⁷³ Relaxation rates were estimated using in-house software, and errors estimated on the basis of noise-based Monte-Carlo simulation.

In-cell NMR

In vivo measurements were performed in 5 mm Shigemi tubes at 600 and 700 MHz ¹H frequency using a TCI and QCI Bruker cryoprobe respectively. Cell viability was monitored by comparing ¹⁵N edited ¹H spectra and of the supernatant immediately before and directly after acquisition. Prior to each experiment, supernatant from the suspension of injected oocytes was collected and single-increment ¹⁵N HSQC and $R_{1,water}$ were measured. For in-cell NMR, MKK4-injected oocytes were gently transferred into a Shigemi tube within a minimal volume of oocyte buffer containing approximately 10% D₂O final volume. Static field homogeneity was optimized on buffer in the same Shigemi tube. Typically NMR studies started 4-5 hours post-injection. All oocyte NMR was carried out at 292K to ensure optimal conditions for live oocytes. Typical sets of experiment included single increment ¹⁵N HSQC for detection of overall protein signal (2 mins), water proton-saturation recovery (20 mins) and ¹⁵N-¹H CSA/DD transverse cross-correlated η_{xy} (16 scans, 1.5 recycle delay, 128 increments per plane, acquisition time 9 hours) and {¹H}-¹⁵N heteronuclear nOe (8 scans, 8 s saturation and recovery delays, 128 increments per plane, acquisition time 9 hours). The length of dephasing gradients was shortened from 1 ms to 500 μ s to decrease losses due to static field inhomogeneity. Experiments were only analysed from the time window where the cells were viable (normally within the first 50 hours post-injection). MKK4 assignments were transferred from η_{xy} spectra from *in vitro* samples recorded at 288K. Upon calculation of relaxation rates, intensities from interleaved experiments were boot-strap averaged between three replica experiments to yield η_{xy} of single biological replica. η_{xy} values were averaged over two biological replicas, giving the final in cell relaxation rates.

Longitudinal relaxation of water

$R_{1,water}$ was measured using a saturation recovery sequence using 20 delays in the range 1 ms to 12 seconds. Data were analysed using NMRPipe and integrals fitted to exponential recovery using in-house software. In cell $R_{1,water}$ data were fitted to a bi-exponential function assuming that the observed water peaks comprised two components corresponding to intra and extracellular water. One of the components was assumed to have the same relaxation characteristics as buffer measured *a posteriori* from the same preparation.

Modelling of ¹⁵N relaxation rates

^{15}N relaxation rates are given by the known functions:

$$R_1 = \frac{1}{10} \left(\frac{\mu_0 \hbar \gamma_H \gamma_N}{4\pi r_{NH}^3} \right)^2 \left(J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N) \right) + \frac{2}{15} \omega_N^2 (\sigma_{\parallel} - \sigma_{\perp})^2 J(\omega_N) \quad (7)$$

$$R_2 = \frac{1}{20} \left(\frac{\mu_0 \hbar \gamma_H \gamma_N}{4\pi r_{NH}^3} \right)^2 \left(4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N) + 6J(\omega_H) \right) + \frac{1}{45} \omega_N^2 (\sigma_{\parallel} - \sigma_{\perp})^2 (4J(0) + 3J(\omega_N)) \quad (8)$$

$$\sigma_{NH} = \frac{1}{10} \left(\frac{\mu_0 \hbar \gamma_H \gamma_N}{4\pi r_{NH}^3} \right)^2 \left(6J(\omega_H + \omega_N) - J(\omega_H - \omega_N) \right) \quad (9)$$

$$\eta_{xy} = \frac{1}{15} P_2(\cos\theta) \left(\frac{\mu_0 \hbar \gamma_H \gamma_N}{4\pi r_{NH}^3} \right) (\sigma_{\parallel} - \sigma_{\perp}) \omega_N (4J(0) + 3J(\omega_N)) \quad (10)$$

$$\eta_z = \frac{1}{15} P_2(\cos\theta) \left(\frac{\mu_0 \hbar \gamma_H \gamma_N}{4\pi r_{NH}^3} \right) (\sigma_{\parallel} - \sigma_{\perp}) \omega_N (6J(\omega_N)) \quad (11)$$

where μ_0 the permittivity of free space, \hbar Planck's constant and θ is the angle between the principal axis of the chemical shift anisotropy (CSA) tensor (assumed axially symmetric with anisotropy $\sigma_{\parallel} - \sigma_{\perp} = -172\text{ppm}$), $J(\omega)$ is the angular spectral density function at frequency ω . r_{NH} is the N-H internuclear distance (assumed to be 1.015\AA), γ_N and γ_H are the gyromagnetic ratios of ^{15}N and ^1H respectively.

A minimization algorithm was used for the fit data to the expressions in equations (7-11) using the definition of the spectral density function in equation 3, by minimizing the following function:

$$\chi_i^2 = \sum_{n=1}^5 \sum_{m=1}^N \left\{ \frac{(R_{n,exp}^m - R_{n,calc}^m)}{\sigma_{n,exp}^m} \right\}^2 \quad (12)$$

for each residue i where n identifies the different rates and m identifies the different conditions used in the fit (temperature and/or crowding). Errors in fitted parameters were estimated using noise-based Monte-Carlo simulations. Overall experimental errors were scaled by a factor of approximately 2 to respect 95% confidence limits. Following results of earlier studies²³ θ was treated as a residue specific parameter that neither evolves with respect to temperature or viscosity. Data-fitting programs are available on request.

Acknowledgements

This work was supported by the ANR (ANR-18-CE29-003) and the Labex GRAL (ANR-10-LABX-49-01). The work used the platforms of the Grenoble Instruct-ERIC centre (ISBG; UMS 3518 CNRS-CEA-UJA-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB) and the Swiss National Science Foundation (Advanced Postdoc Mobility Fellowship P300P2_167742 to N.S.). This work was also supported by grants from the French Agency ANSES (EST-16-66) to C.J.M, from the National Institutes of Health (Grant N°EB007047). We thank Hervé Pointu, Soumalamaya Bama Toupet and Charlène Caloud for the management and the maintenance of Xenopus and acknowledge the platform supported by GRAL,

financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA).

Supporting Information

Additional figures showing NMR spectra of MKK4 and N_{TAIL} in the presence of dextran and PEG, additional analysis, cross validation of data removed from the analysis and water relaxation in MKK4 samples, as well as estimation of extinction coefficient of dextran40 and 1D ¹H spectra *in cellulo*.

REFERENCES

- (1) Tompa, P.; Davey, N. E.; Gibson, T. J.; Babu, M. M. A Million Peptide Motifs for the Molecular Biologist. *Mol. Cell* **2014**, *55* (2), 161–169.
- (2) Dyson, H. J.; Wright, P. E. Intrinsically Unstructured Proteins and Their Functions. *Nat. Rev. Mol. Cell Biol.* **2005**, *6* (3), 197–208.
- (3) Uversky, V. N.; Dunker, A. K. Understanding Protein Non-Folding. *Biochim. Biophys. Acta* **2010**, *1804* (6), 1231–1264.
- (4) Csermely, P.; Palotai, R.; Nussinov, R. Induced Fit, Conformational Selection and Independent Dynamic Segments: An Extended View of Binding Events. *Trends Biochem. Sci.* **2010**, *35* (10), 539–546.
- (5) Jensen, M. R.; Zweckstetter, M.; Huang, J.; Blackledge, M. Exploring Free-Energy Landscapes of Intrinsically Disordered Proteins at Atomic Resolution Using NMR Spectroscopy. *Chem. Rev.* **2014**, *114* (13), 6632–6660.
- (6) Schneider, R.; Blackledge, M.; Jensen, M. R. Elucidating Binding Mechanisms and Dynamics of Intrinsically Disordered Protein Complexes Using NMR Spectroscopy. *Curr. Opin. Struct. Biol.* **2019**, *54*, 10–18.
- (7) Bernado, P.; Blanchard, L.; Timmins, P.; Marion, D.; Ruigrok, R. W. H.; Blackledge, M. A Structural Model for Unfolded Proteins from Residual Dipolar Couplings and Small-Angle x-Ray Scattering. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (47), 17002–17007.
- (8) Dedmon, M.; Lindorff-Larsen, K.; Christodoulou, J.; Vendruscolo, M.; Dobson, C. Mapping Long-Range Interactions in Alpha-Synuclein Using Spin-Label NMR and Ensemble Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2005**, *127* (2), 476–477.
- (9) Mittag, T.; Forman-Kay, J. D. Atomic-Level Characterization of Disordered Protein Ensembles. *Curr. Opin. Struct. Biol.* **2007**, *17* (1), 3–14.
- (10) Tamiola, K.; Acar, B.; Mulder, F. A. A. Sequence-Specific Random Coil Chemical Shifts of Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* **2010**, *132* (51), 18000–18003.
- (11) Salmon, L.; Nodet, G.; Ozenne, V.; Yin, G.; Jensen, M. R.; Zweckstetter, M.; Blackledge, M. NMR Characterization of Long-Range Order in Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* **2010**, *132* (24), 8407–8418.
- (12) Ozenne, V.; Schneider, R.; Yao, M.; Huang, J.-R.; Salmon, L.; Zweckstetter, M.; Jensen, M. R.; Blackledge, M. Mapping the Potential Energy Landscape of Intrinsically Disordered Proteins at Amino Acid Resolution. *J. Am. Chem. Soc.* **2012**, *134* (36), 15138–15148.
- (13) Hofmann, H.; Soranno, A.; Borgia, A.; Gast, K.; Nettels, D.; Schuler, B. Polymer Scaling Laws of Unfolded and Intrinsically Disordered Proteins Quantified with Single-Molecule Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (40), 16155–16160.
- (14) Nath, A.; Sammalkorpi, M.; DeWitt, D. C.; Trexler, A. J.; Elbaum-Garfinkle, S.; O’Hern, C. S.; Rhoades, E. The Conformational Ensembles of α -Synuclein and Tau: Combining Single-Molecule FRET and Simulations. *Biophys. J.* **2012**, *103* (9), 1940–1949.
- (15) Milles, S.; Lemke, E. A. Single Molecule Study of the Intrinsically Disordered FG-Repeat

Nucleoporin 153. *Biophys. J.* **2011**, *101* (7), 1710–1719.

(16) Bracken, C.; Carr, P. A.; Cavanagh, J.; Palmer, A. G. Temperature Dependence of Intramolecular Dynamics of the Basic Leucine Zipper of GCN4: Implications for the Entropy of Association with DNA11 Edited by P. E. Wright. *J. Mol. Biol.* **1999**, *285* (5), 2133–2146.

(17) Modig, K.; Poulsen, F. M. Model-Independent Interpretation of NMR Relaxation Data for Unfolded Proteins: The Acid-Denatured State of ACBP. *J. Biomol. Nmr* **2008**, *42* (3), 163–177.

(18) Xue, Y.; Skrynnikov, N. R. Motion of a Disordered Polypeptide Chain as Studied by Paramagnetic Relaxation Enhancements, N-15 Relaxation, and Molecular Dynamics Simulations: How Fast Is Segmental Diffusion in Denatured Ubiquitin? *J. Am. Chem. Soc.* **2011**, *133* (37), 14614–14628.

(19) Kaderavek, P.; Zapletal, V.; Rabatinova, A.; Krasny, L.; Sklenar, V.; Zidek, L. Spectral Density Mapping Protocols for Analysis of Molecular Motions in Disordered Proteins. *J. Biomol. Nmr* **2014**, *58* (3), 193–207.

(20) Gill, M. L.; Byrd, R. A.; Arthur G. Palmer, I. I. Dynamics of GCN4 Facilitate DNA Interaction: A Model-Free Analysis of an Intrinsically Disordered Region. *Phys. Chem. Chem. Phys.* **2015**, *18* (8):5839–49

(21) Khan, S. N.; Charlier, C.; Augustyniak, R.; Salvi, N.; Dejean, V.; Bodenhausen, G.; Lequin, O.; Pelupessy, P.; Ferrage, F. Distribution of Pico- and Nanosecond Motions in Disordered Proteins from Nuclear Spin Relaxation. *Biophys. J.* **2015**, *109* (5), 988–999.

(22) Salvi, N.; Abyzov, A.; Blackledge, M. Atomic Resolution Conformational Dynamics of Intrinsically Disordered Proteins from NMR Spin Relaxation. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *102–103*, 43–60.

(23) Abyzov, A.; Salvi, N.; Schneider, R.; Maurin, D.; Ruigrok, R. W. H.; Jensen, M. R.; Blackledge, M. Identification of Dynamic Modes in an Intrinsically Disordered Protein Using Temperature-Dependent NMR Relaxation. *J. Am. Chem. Soc.* **2016**, *138* (19), 6240–6251.

(24) Salvi, N.; Abyzov, A.; Blackledge, M. Multi-Timescale Dynamics in Intrinsically Disordered Proteins from NMR Relaxation and Molecular Simulation. *J. Phys. Chem. Lett.* **2016**, *7* (13), 2483–2489.

(25) Salvi, N.; Abyzov, A.; Blackledge, M. Analytical Description of NMR Relaxation Highlights Correlated Dynamics in Intrinsically Disordered Proteins. *Angew. Chem. Int. Ed Engl.* **2017**, *56* (45), 14020–14024.

(26) Theillet, F.-X.; Binolfi, A.; Frembgen-Kesner, T.; Hingorani, K.; Sarkar, M.; Kyne, C.; Li, C.; Crowley, P. B.; Gierasch, L.; Pielak, G. J.; et al. Physicochemical Properties of Cells and Their Effects on Intrinsically Disordered Proteins (IDPs). *Chem. Rev.* **2014**, *114* (13), 6661–6714.

(27) Zimmerman, S. B.; Trach, S. O. Estimation of Macromolecule Concentrations and Excluded Volume Effects for the Cytoplasm of Escherichia Coli. *J. Mol. Biol.* **1991**, *222* (3), 599–620.

(28) Ellis, R. J. Macromolecular Crowding: Obvious but Underappreciated. *Trends Biochem. Sci.* **2001**, *26* (10), 597–604.

(29) Zeskind, B. J.; Jordan, C. D.; Timp, W.; Trapani, L.; Waller, G.; Horodincu, V.; Ehrlich, D. J.; Matsudaira, P. Nucleic Acid and Protein Mass Mapping by Live-Cell Deep-Ultraviolet

Microscopy. *Nat. Methods* **2007**, *4* (7), 567–569.

(30) Cohen, R. D.; Pielak, G. J. A Cell Is More than the Sum of Its (Dilute) Parts: A Brief History of Quinary Structure. *Protein Sci.* **2017**, *26* (3), 403–413.

(31) Davey, N. E. The Functional Importance of Structure in Unstructured Protein Regions. *Curr. Opin. Struct. Biol.* **2019**, *56*, 155–163.

(32) Bülow, S. von; Siggel, M.; Linke, M.; Hummer, G. Dynamic Cluster Formation Determines Viscosity and Diffusion in Dense Protein Solutions. *Proc. Natl. Acad. Sci.* **2019**, *116* (20), 9843–9852.

(33) Schuler, B.; Eaton, W. A. Protein Folding Studied by Single-Molecule FRET. *Curr. Opin. Struct. Biol.* **2008**, *18* (1), 16–26.

(34) Soranno, A.; Buchli, B.; Nettels, D.; Cheng, R. R.; Müller-Späh, S.; Pfeil, S. H.; Hoffmann, A.; Lipman, E. A.; Makarov, D. E.; Schuler, B. Quantifying Internal Friction in Unfolded and Intrinsically Disordered Proteins with Single-Molecule Spectroscopy. *Proc. Natl. Acad. Sci.* **2012**, *109* (44), 17800–17806.

(35) Echeverria, I.; Makarov, D. E.; Papoian, G. A. Concerted Dihedral Rotations Give Rise to Internal Friction in Unfolded Proteins. *J. Am. Chem. Soc.* **2014**, *136* (24), 8708–8713.

(36) König, I.; Zarrine-Afsar, A.; Aznauryan, M.; Soranno, A.; Wunderlich, B.; Dingfelder, F.; Stüber, J. C.; Plückthun, A.; Nettels, D.; Schuler, B. Single-Molecule Spectroscopy of Protein Conformational Dynamics in Live Eukaryotic Cells. *Nat. Methods* **2015**, *12* (8), 773–779.

(37) Paudel, B. P.; Fiorini, E.; Börner, R.; Sigel, R. K. O.; Rueda, D. S. Optimal Molecular Crowding Accelerates Group II Intron Folding and Maximizes Catalysis. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (47), 11917–11922.

(38) Selenko, P.; Serber, Z.; Gadea, B.; Ruderman, J.; Wagner, G. Quantitative NMR Analysis of the Protein G B1 Domain in *Xenopus Laevis* Egg Extracts and Intact Oocytes. *Proc. Natl. Acad. Sci.* **2006**, *103* (32), 11904–11909.

(39) Li, C.; Charlton, L. M.; Lakkavaram, A.; Seagle, C.; Wang, G.; Young, G. B.; Macdonald, J. M.; Pielak, G. J. Differential Dynamical Effects of Macromolecular Crowding on an Intrinsically Disordered Protein and a Globular Protein: Implications for in-Cell NMR Spectroscopy. *J. Am. Chem. Soc.* **2008**, *130* (20), 6310–6311.

(40) Ito, Y.; Selenko, P. Cellular Structural Biology. *Curr. Opin. Struct. Biol.* **2010**, *20* (5), 640–648.

(41) Smith, A. E.; Zhang, Z.; Pielak, G. J.; Li, C. NMR Studies of Protein Folding and Binding in Cells and Cell-like Environments. *Curr. Opin. Struct. Biol.* **2015**, *30*, 7–16.

(42) Theillet, F.-X.; Binolfi, A.; Bekei, B.; Martorana, A.; Rose, H. M.; Stuiver, M.; Verzini, S.; Lorenz, D.; van Rossum, M.; Goldfarb, D.; et al. Structural Disorder of Monomeric α -Synuclein Persists in Mammalian Cells. *Nature* **2016**, *530* (7588), 45–50.

(43) Luchinat, E.; Banci, L. In-Cell NMR in Human Cells: Direct Protein Expression Allows Structural Studies of Protein Folding and Maturation. *Acc. Chem. Res.* **2018**, *51* (6), 1550–1557.

(44) Guseman, A. J.; Perez Goncalves, G. M.; Speer, S. L.; Young, G. B.; Pielak, G. J. Protein

Shape Modulates Crowding Effects. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (43), 10965–10970.

(45) Sekhar, A.; Latham, M. P.; Vallurupalli, P.; Kay, L. E. Viscosity-Dependent Kinetics of Protein Conformational Exchange: Microviscosity Effects and the Need for a Small Viscogen. *J. Phys. Chem. B* **2014**, *118* (17), 4546–4551.

(46) Roos, M.; Ott, M.; Hofmann, M.; Link, S.; Rössler, E.; Balbach, J.; Krushelnitsky, A.; Saalwächter, K. Coupling and Decoupling of Rotational and Translational Diffusion of Proteins under Crowding Conditions. *J. Am. Chem. Soc.* **2016**, *138* (32), 10365–10372.

(47) Jensen, M. R.; Houben, K.; Lescop, E.; Blanchard, L.; Ruigrok, R. W. H.; Blackledge, M. Quantitative Conformational Analysis of Partially Folded Proteins from Residual Dipolar Couplings: Application to the Molecular Recognition Element of Sendai Virus Nucleoprotein. *J. Am. Chem. Soc.* **2008**, *130* (25), 8055–8061.

(48) Abragam, A. *The Principles of Nuclear Magnetism*, Reprint.; Clarendon Press, 1994.

(49) Persson, E.; Halle, B. Cell Water Dynamics on Multiple Time Scales. *Proc. Natl. Acad. Sci.* **2008**, *105* (17), 6266–6271.

(50) Kimmich, R.; Fatkullin, N. Self-Diffusion Studies by Intra- and Inter-Molecular Spin-Lattice Relaxometry Using Field-Cycling: Liquids, Plastic Crystals, Porous Media, and Polymer Segments. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *101*, 18–50.

(51) Korb, J.-P. Multiscale Nuclear Magnetic Relaxation Dispersion of Complex Liquids in Bulk and Confinement. *Prog. Nucl. Magn. Reson. Spectrosc.* **2018**, *104*, 12–55.

(52) Lipari, G.; Szabo, A. Model-Free Approach To The Interpretation Of Nuclear Magnetic-Resonance Relaxation In Macromolecules .1. Theory And Range Of Validity. *J. Am. Chem. Soc.* **1982**, *104* (17), 4546–4559.

(53) Cukier, R. I. Diffusion of Brownian Spheres in Semidilute Polymer Solutions. *Macromolecules* **1984**, *17* (2), 252–255.

(54) Barshtein, G.; Almagor, A.; Yedgar, S.; Gavish, B. Inhomogeneity of Viscous Aqueous-Solutions. *Phys. Rev. E* **1995**, *52* (1), 555–557.

(55) Lavalette, D.; Tétreau, C.; Tourbez, M.; Blouquit, Y. Microscopic Viscosity and Rotational Diffusion of Proteins in a Macromolecular Environment. *Biophys. J.* **1999**, *76* (5), 2744–2751.

(56) Szymański, J.; Patkowski, A.; Wilk, A.; Garstecki, P.; Holyst, R. Diffusion and Viscosity in a Crowded Environment: From Nano- to Macroscale. *J. Phys. Chem. B* **2006**, *110* (51), 25593–25597.

(57) Sekhar, A.; Vallurupalli, P.; Kay, L. E. Defining a Length Scale for Millisecond-Timescale Protein Conformational Exchange. *Proc. Natl. Acad. Sci.* **2013**, *110* (28), 11391–11396.

(58) Kalwarczyk, T.; Sozanski, K.; Ochab-Marcinek, A.; Szymanski, J.; Tabaka, M.; Hou, S.; Holyst, R. Motion of Nanoprobes in Complex Liquids within the Framework of the Length-Scale Dependent Viscosity Model. *Adv. Colloid Interface Sci.* **2015**, *223*, 55–63.

(59) Wisniewska, A.; Sozanski, K.; Kalwarczyk, T.; Kedra-Krolik, K.; Holyst, R. Scaling Equation for Viscosity of Polymer Mixtures in Solutions with Application to Diffusion of Molecular Probes. *Macromolecules* **2017**, *50* (11), 4555–4561. <https://doi.org/10.1021/acs.macromol.7b00545>.

(60) Einstein, A. Eine Neue Bestimmung Der Moleküldimensionen. *Ann. Phys.* **1906**, *324* (2),

289–306.

- (61) Delaforge, E.; Kragelj, J.; Tengo, L.; Palencia, A.; Milles, S.; Bouvignies, G.; Salvi, N.; Blackledge, M.; Jensen, M. R. Deciphering the Dynamic Interaction Profile of an Intrinsically Disordered Protein by NMR Exchange Spectroscopy. *J. Am. Chem. Soc.* **2018**, *140* (3), 1148–1158.
- (62) Selenko, P.; Wagner, G. Looking into Live Cells with In-Cell NMR Spectroscopy. *J. Struct. Biol.* **2007**, *158* (2), 244–253.
- (63) Salvi, N.; Abyzov, A.; Blackledge, M. Solvent-Dependent Segmental Dynamics in Intrinsically Disordered Proteins. *Sci. Adv.* **2019**, *5* (6), eaax2348.
- (64) Huang, J.; Ozenne, V.; Jensen, M. R.; Blackledge, M. Direct Prediction of NMR Residual Dipolar Couplings from the Primary Sequence of Unfolded Proteins. *Angew. Chem. Int. Ed Engl.* **2013**, *52* (2), 687–690.
- (65) Frauenfelder, H.; Fenimore, P. W.; Chen, G.; McMahon, B. H. Protein Folding Is Slaved to Solvent Motions. *Proc. Natl. Acad. Sci.* **2006**, *103* (42), 15469–15472.
- (66) Houben, K.; Blanchard, L.; Blackledge, M.; Marion, D. Intrinsic Dynamics of the Partly Unstructured PX Domain from the Sendai Virus RNA Polymerase Cofactor P. *Biophys. J.* **2007**, *93* (8), 2830–2844.
- (67) Moreau, C. J.; Niescierowicz, K.; Caro, L. N.; Revilloud, J.; Vivaudou, M. Ion Channel Reporter for Monitoring the Activity of Engineered GPCRs. *Methods Enzymol.* **2015**, *556*, 425–454.
- (68) Lakomek, N.-A.; Ying, J.; Bax, A. Measurement of ^{15}N Relaxation Rates in Perdeuterated Proteins by TROSY-Based Methods. *J. Biomol. NMR* **2012**, *53* (3), 209–221.
- (69) Akke, M.; Palmer, A. G. Monitoring Macromolecular Motions on Microsecond to Millisecond Time Scales by R1 ρ –R1 Constant Relaxation Time NMR Spectroscopy. *J. Am. Chem. Soc.* **1996**, *118* (4), 911–912.
- (70) Pelupessy, P.; Espallargas, G. M.; Bodenhausen, G. Symmetrical Reconversion: Measuring Cross-Correlation Rates with Enhanced Accuracy. *J. Magn. Reson. San Diego Calif 1997* **2003**, *161* (2), 258–264.
- (71) Pelupessy, P.; Ferrage, F.; Bodenhausen, G. Accurate Measurement of Longitudinal Cross-Relaxation Rates in Nuclear Magnetic Resonance. *J. Chem. Phys.* **2007**, *126* (13), 134508.
- (72) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J. Biomol. NMR* **1995**, *6* (3), 277–293.
- (73) Lee, W.; Tonelli, M.; Markley, J. L. NMRFAM-SPARKY: Enhanced Software for Biomolecular NMR Spectroscopy. *Bioinformatics.* **2015**, *31* (8), 1325–1327.

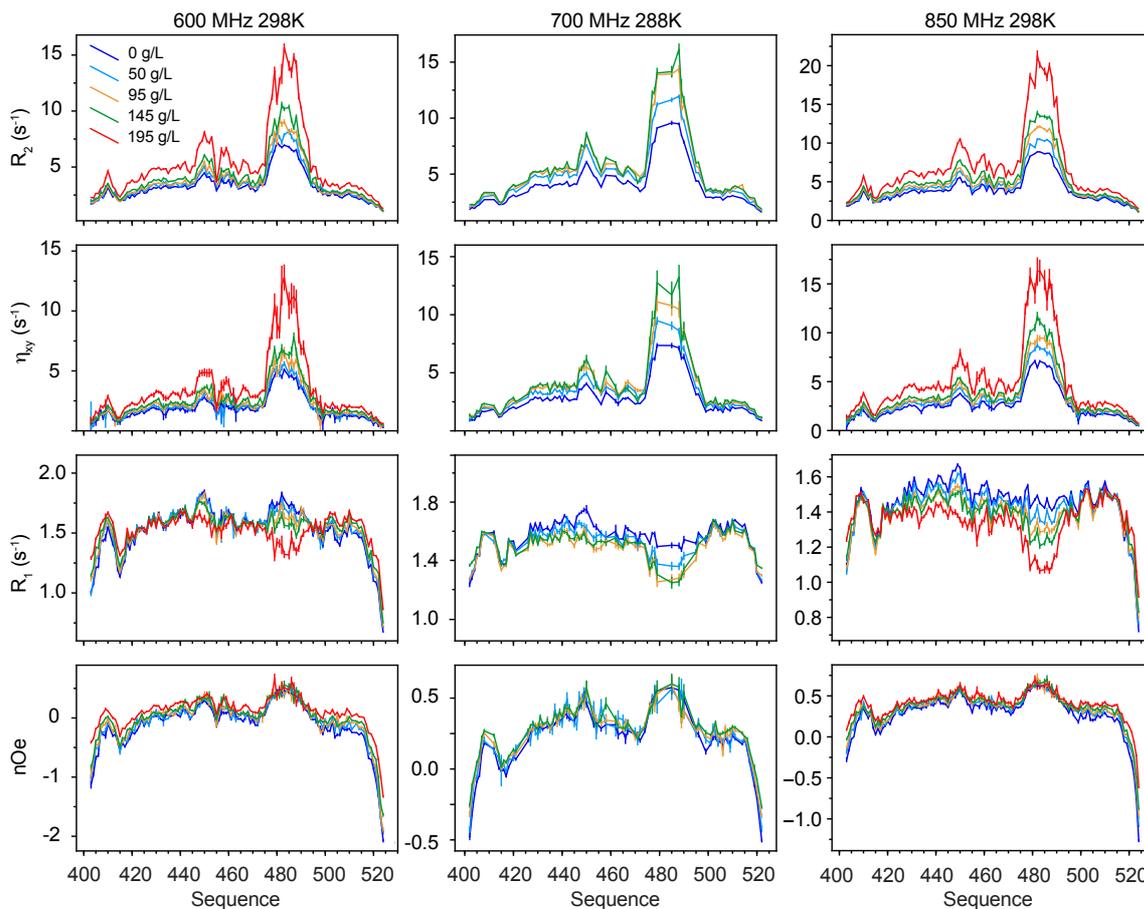


Figure 1. Experimental ^{15}N relaxation rates measured on N_{TAIL} as a function of molecular crowding.

Longitudinal (R_1) and transverse (R_2) autocorrelated relaxation, transverse cross-correlated DD/CSA (η_{xy}) and heteronuclear $\{^1\text{H}\}$ - ^{15}N nuclear Overhauser enhancement (nOe) measured at three magnetic field strengths (600, 700 and 850 MHz) as a function of viscogen concentration (dextran40). Additional longitudinal cross-correlated DD/CSA (η_z) were measured at 600 and 850 MHz and are shown in supplementary figure S2.

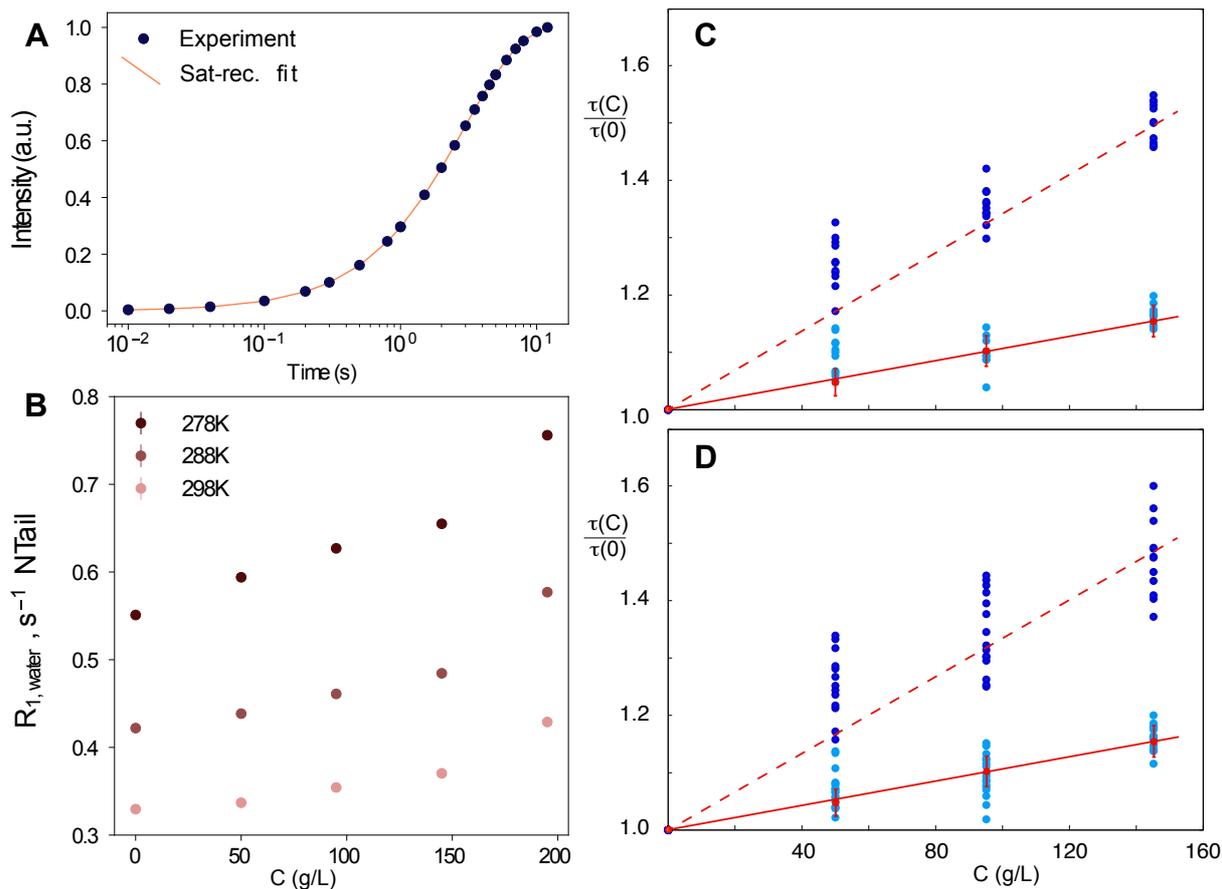


Figure 2. Comparison of the influence of crowding on dynamics of water protons and ^{15}N backbone amide correlation times.

A – Example of evolution of water proton intensity during saturation recovery and fit to $I(t) = I_0 \left(1 - \exp(-R_{1,\text{water}}t) \right)$. B - longitudinal relaxation of water protons for admixtures of N_{TAIL} with different concentrations (C) of dextran40 recorded at three temperatures at 600 MHz. C and D show relative changes of backbone amide dynamic timescales for N_{TAIL} (dark blue: τ_3 , light blue: τ_2) estimated from an 11 parameter model-free of 45 relaxation rates recorded at two magnetic fields (600 and 850 MHz) at 298K. τ_3 and τ_2 were optimized separately for each concentration with τ_1 kept constant at 50 ps outside and 0 ps inside the transiently folded helical region, keeping amplitudes constant at all concentrations but allowing them to vary in the range of 0 to 1 for each site. Examples of two regions (panel C: 20-40; panel D: 100-120) are shown. Red line shows relative changes in dynamics of water measured by relaxation of water protons (red solid line: $\rho(C) + 1$, red points and error bars show average and standard deviation over three temperatures (278, 288 and 298K) respectively; red dotted line: $(3\rho(C) + 1)$). This figure demonstrates that intermediate and slow motional timescales exhibit distinct responses to viscosity, with the intermediate motion mirroring the friction coefficient of water.

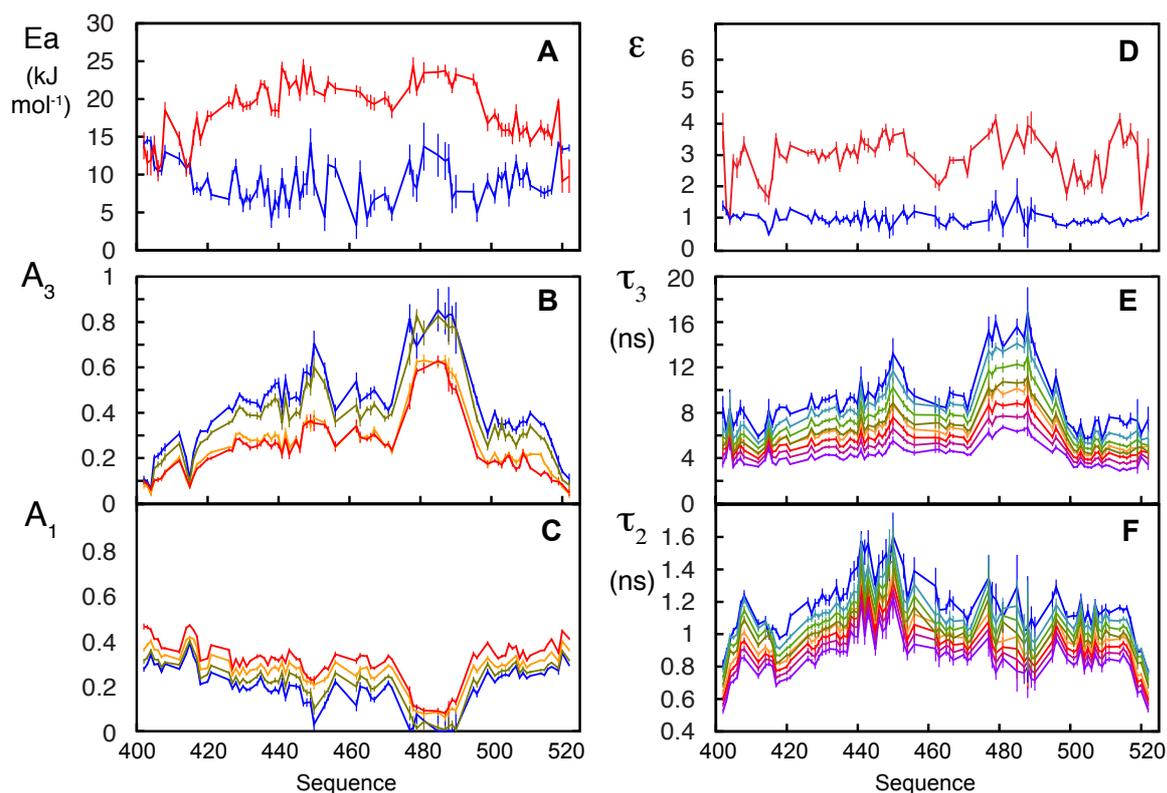


Figure 3. Combined analysis of dynamic behaviour of N_{TAIL} as a function of temperature and viscosity.

105 relaxation rates were analysed and fitted to the general expression relating rotational correlation times to viscosity and temperature (equation 5). 15 parameters were optimized for each ^{15}N - ^1H spin pair ($\epsilon_3, \epsilon_2, \tau_{3,\infty}, \tau_{2,\infty}, E_{a,3}, E_{a,2}, \theta$ and A_2 and A_3 at each T). A – activation energies for backbone (blue) and segmental (red) contributions. B – amplitudes of segmental motions at 274 (blue), 278 (green), 288 (orange) and 298K (red). C - amplitudes of fast timescale motions (see B). D – Friction coefficients (ϵ_k) for segmental (red) and backbone (blue) motions. E – Timescales of segmental motions: 0 g/L dextran40, 298K (purple), 50 g/L, 298K (dark red), 95 g/L, 298K (red), 145 g/L, 298K (orange), 195 g/L, 298K (blue), 50 g/L, 288K (bottle green), 95 g/L, 288K (green), 145 g/L, 288K (blue-green). F – Timescales of intermediate, backbone motions (legend as E).

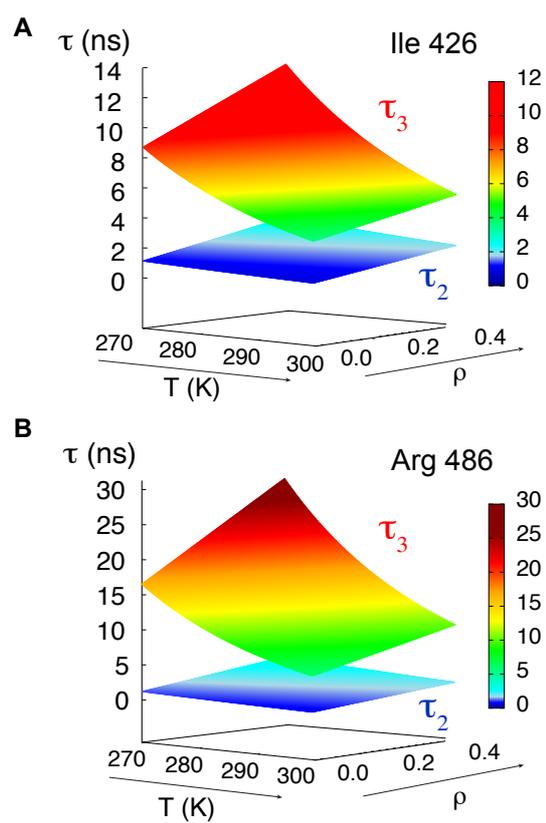


Figure 4. Mapping the dynamic behaviour of N_{TAIL} at amino acid specific resolution as a function of temperature and viscosity.

Experimental data fitted to equation 5 allow us to map the intermediate and slow correlation times as a function of temperature and solvent friction for each amino acid in the chain. Shown are two examples (Ile 426 and Arg 486 from the disordered and partially helical regions respectively).

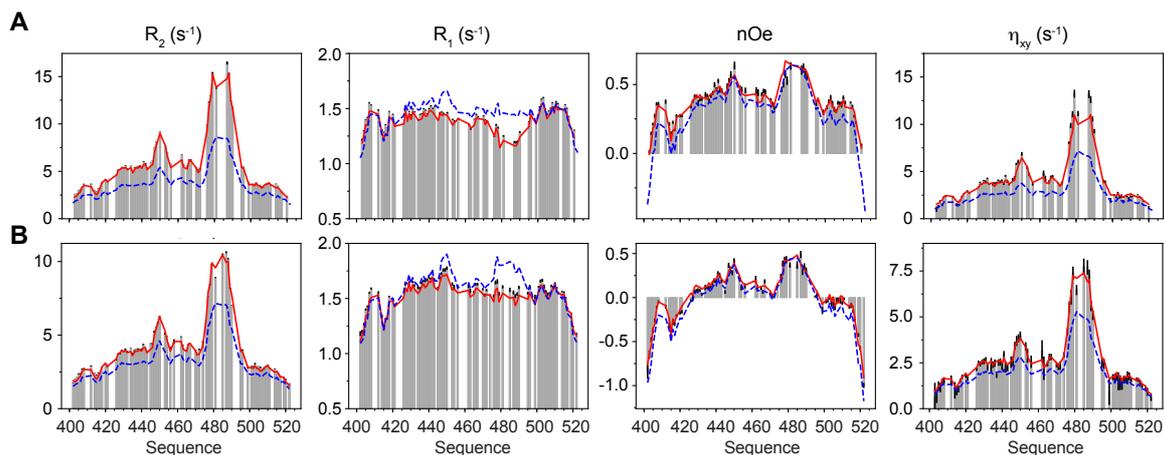


Figure 5 Comparison of predicted and experimental relaxation from N_{TAIL} measured in the presence of different concentration of PEG crowding agent

Comparison of ^{15}N relaxation rates recorded on N_{TAIL} in the presence of PEG (grey bars) with rates calculated using equation 5 (red lines) on the basis of experimentally determined ε_2 and ε_3 and energies of activation and $\rho(C)$ determined from the experimental measurement of $R_{1,water}$. Relaxation rates predicted under dilute conditions ($\rho(C) = 0.0$) at the same temperature are shown as blue dashed lines for comparison. (A) Relaxation rates measured at 850 MHz ^1H frequency at a concentration of 135 g/L PEG at 298K ($\rho(C) = 0.25$). (B) Relaxation rates measured at 600 MHz ^1H frequency at a concentration of 90 g/L PEG at 298K ($\rho(C) = 0.16$).

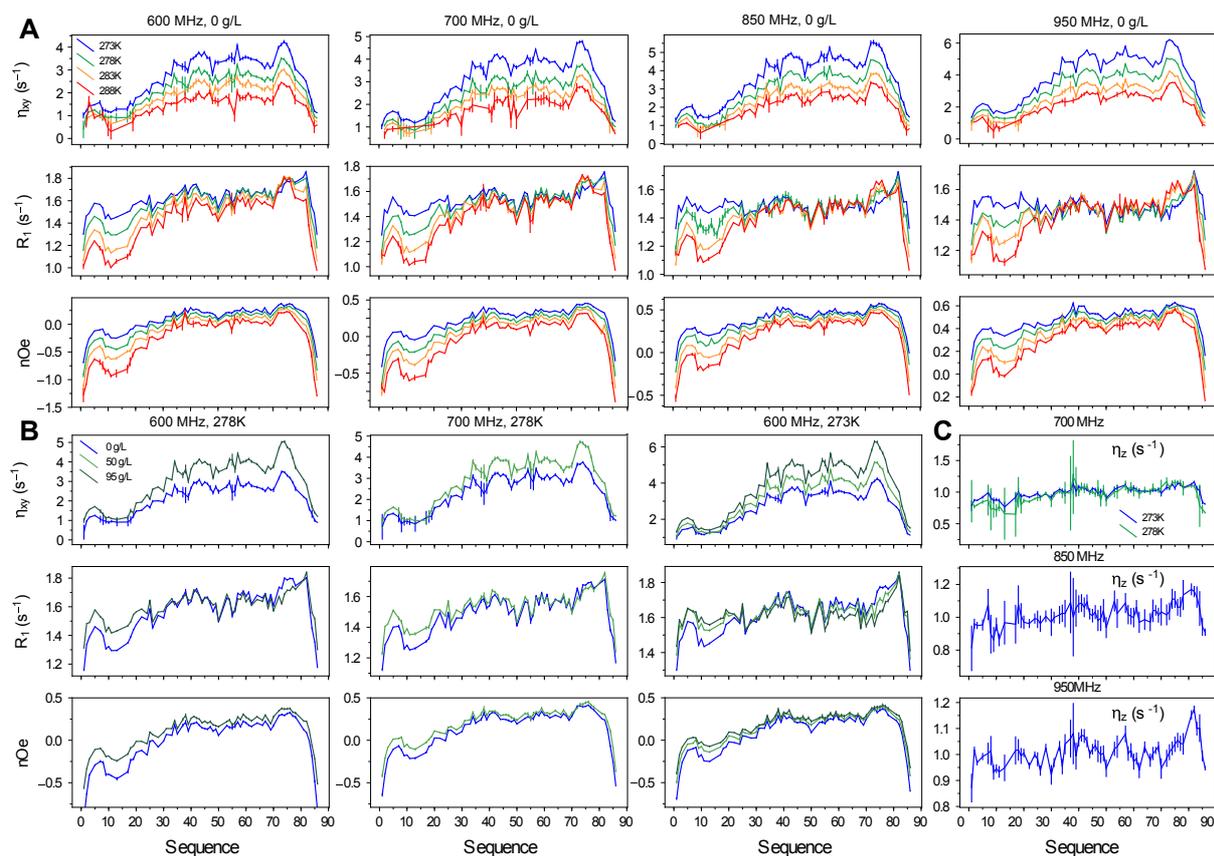


Figure 6. Relaxation data measured from MKK4 as a function of temperature and viscosity

Longitudinal (R_1) autocorrelated relaxation, transverse cross-correlated DD/CSA (η_{xy}) and heteronuclear $\{^1\text{H}\}$ - ^{15}N nuclear Overhauser enhancement (nOe) measured at four magnetic field strengths (600, 700, 850 and 950 MHz) as a function of temperature (A and C) and viscogen concentration (dextran40) (B). Legends identify the colour code in (A, B and C).

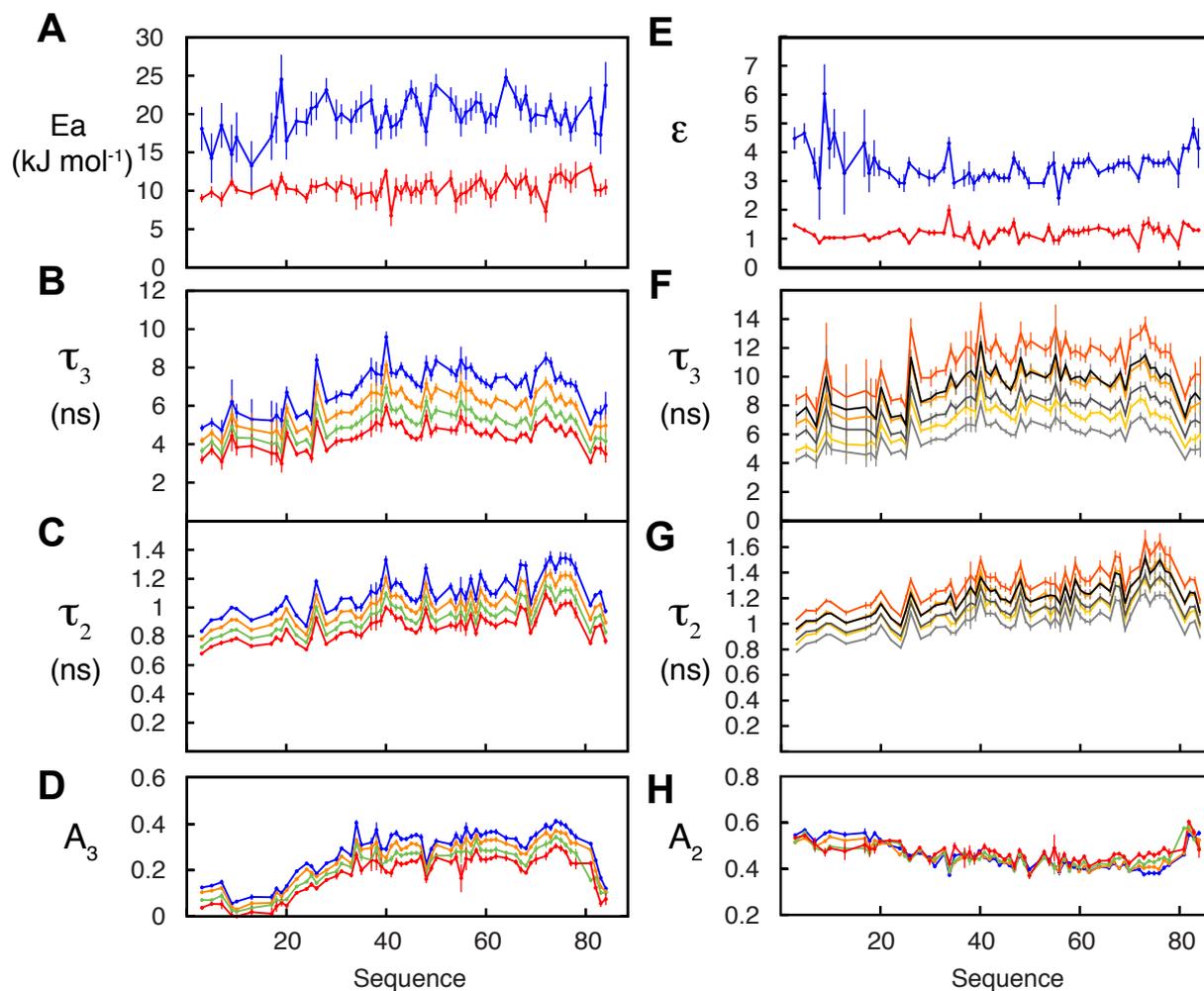


Figure 7. Combined analysis of dynamics of MKK4 as a function of temperature and viscosity.

68 relaxation rates were simultaneously analysed and fitted to equation 5 relating correlation times to viscosity and temperature. 15 parameters were again optimized for each ^{15}N - ^1H spin pair ($\epsilon_3, \epsilon_2, \tau_{3,\infty}, \tau_{2,\infty}, E_{a,3}, E_{a,2}, \theta$ and A_2 and A_3 at each temperature). Selected values are shown in this figure.

A - activation energies for backbone (blue) and segmental (red) contributions. B - Timescales of segmental motions in the absence of crowder at 273 (blue), 278 (orange), 283 (green) and 288K (red). C - Timescales of intermediate, backbone motions (legend as B). D - amplitudes of segmental motions (for legend see B). E - Friction coefficients (ϵ_k) for segmental (red) and backbone (blue) motions. F - Timescales of segmental motions: 0 g/L, 278K (light grey), 50 g/L, 278K (grey), 95 g/L, 278K (dark grey), 0 g/L, 273K (yellow), 50 g/L, 273K (orange), 95 g/L, 273K (red). G - Timescales of intermediate backbone motions (legend as in F). H - amplitudes of intermediate timescale motions (for legend see B).

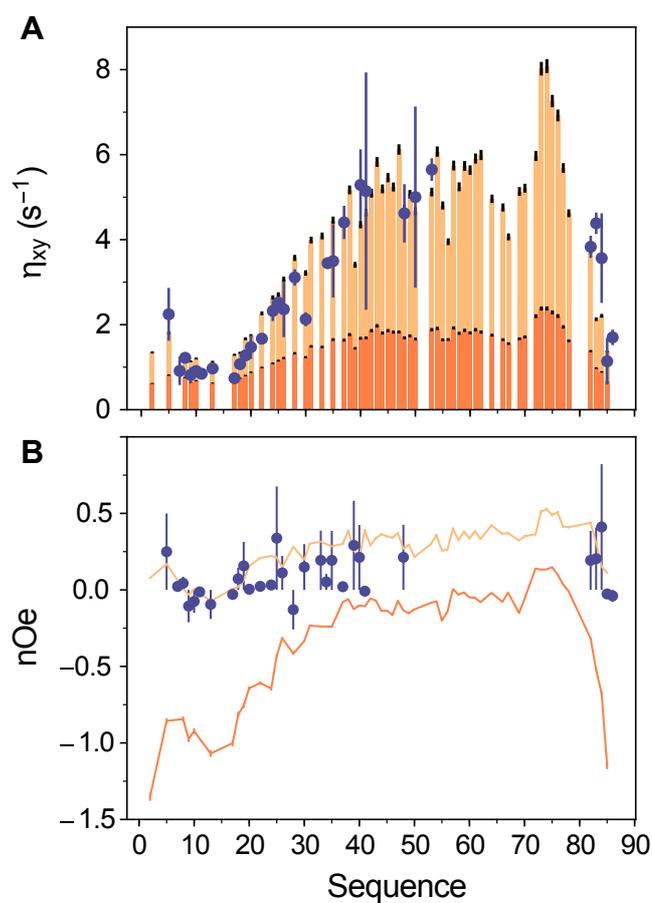


Figure 8. Comparison of predicted and experimental relaxation measured in vivo from MKK4.

Comparison of ¹⁵N relaxation rates recorded in-cell (blue points) with rates calculated on the basis of dynamic parameters (pale orange bars and lines): timescales and amplitudes determined *in vitro* were extrapolated to 292K using energies of activation and linear temperature dependence respectively; experimentally determined ε_2 and ε_3 values determined *in vitro* were incorporated into equation 5 and $\rho(C)$ determined from the experimental measurement of $R_{1,water}$ in *cellulo*. For comparison – dark orange shows the predicted relaxation rates at 292K in non-crowded conditions. A – Cross correlated DD/CSA relaxation, B – heteronuclear nOe.

Graphical abstract (TOC)

