Biophysical Journal
Article

Convergent views on disordered protein dynamics from NMR and computational approaches

Nicola Salvi,[1] Vojtěch Zapletal,[2] Zuzana Jaseňáková,[2] Milan Zachrdla,[3] Petr Padrta,[4] Subhash Narasimhan,[2] Thorsten Marquardsen,[5] Jean-Max Tyburn,[6] Lukáš Žídek,[2] Martin Blackledge,[1,*] Fabien Ferrage,[3,*] and Pavel Kadeřávek[4,*]

[1]Institut de Biologie Structurale (IBS), CEA, CNRS, University Grenoble Alpes, Grenoble, France; [2]National Centre for Biomolecular Research, Faculty of Science and Central European Institute of Technology, Masaryk University, Brno, Czech Republic; [3]Laboratoire des Biomolécules, LBM, Département de chimie, École normale supérieure, PSL University, Sorbonne Université, CNRS, Paris, France; [4]Central European Institute of Technology, Masaryk University, Brno, Czech Republic; [5]Bruker BioSpin GmbH, Rheinstetten, Germany; and [6]BioSpin, Wissembourg Cedex, France

ABSTRACT Intrinsically disordered proteins (IDPs) or intrinsically disordered regions (IDRs) is a class of biologically important proteins exhibiting specific biophysical characteristics. They lack a hydrophobic core, and their conformational behavior is strongly influenced by electrostatic interactions. IDPs and IDRs are highly dynamic, and a characterization of the motions of IDPs and IDRs is essential for their physically correct description. NMR together with molecular dynamics simulations are the methods best suited to such a task because they provide information about dynamics of proteins with atomistic resolution. Here, we present a study of motions of a disordered C-terminal domain of the delta subunit of RNA polymerase from *Bacillus subtilis*. Positively and negatively charged residues in the studied domain form transient electrostatic contacts critical for the biological function. Our study is focused on investigation of ps-ns dynamics of backbone of the delta subunit based on analysis of amide ^{15}N NMR relaxation data and molecular dynamics simulations. In order to extend an informational content of NMR data to lower frequencies, which are more sensitive to slower motions, we combined standard (high-field) NMR relaxation experiments with high-resolution relaxometry. Altogether, we collected data reporting the relaxation at 12 different magnetic fields, resulting in an unprecedented data set. Our results document that the analysis of such data provides a consistent description of dynamics and confirms the validity of so far used protocols of the analysis of dynamics of IDPs also for a partially folded protein. In addition, the potential to access detailed description of motions at the timescale of tens of ns with the help of relaxometry data is discussed. Interestingly, in our case, it appears to be mostly relevant for a region involved in the formation of temporary contacts within the disordered region, which was previously proven to be biologically important.

SIGNIFICANCE Dynamics of proteins is essential for their function in biological systems. The importance of the dynamics is even more pronounced in the case of intrinsically disordered proteins (IDPs), which lack a stable three-dimensional structure and are highly flexible. A combination of analyses of molecular dynamic simulations and NMR relaxation data provides detailed information about the motions at the ps-ns timescale, but sampling of slower motions is limited. We are presenting an approach overcoming this limitation by employing high-resolution relaxometry as demonstrated for a case study of the delta subunit of RNA polymerase from *Bacillus subtilis,* where it is used to describe biologically relevant dynamics of its C-terminal domain involved in regulation of transcription.

INTRODUCTION

Proteins are dynamic biomolecules that feature motions occurring on timescales from ps to days. It is widely accepted that the problem of describing the functional mechanisms of a protein is essentially equivalent to the problem of describing its functional dynamics with sufficient accuracy (1). This statement is particularly relevant in the case of intrinsically disordered regions (IDRs) and intrinsically disordered proteins (IDPs), which lack a

---

Submitted June 11, 2022, and accepted for publication September 15, 2022.

*Correspondence: martin.blackledge@ibs.fr or fabien.ferrage@ens.psl.eu or pavel.kaderavek@mail.muni.cz

Nicola Salvi's present address is Bio Structure and Biophysics (BSB), Sanofi R&D, 94400 Vitry-sur-Seine, France.

Thorsten Marquardsen is deceased.

Editor: Scott Showalter.

https://doi.org/10.1016/j.bpj.2022.09.016

well-defined three-dimensional fold and are best described by an ensemble of diverse conformations sampled in solution (2,3). A direct yet often unclear link between the structural ensemble of an IDP, the dynamics of interconversion between conformers within such ensemble, and protein function exists. The development of approaches to describe this link has been the subject of considerable research effort in recent years (4–11).

Among the experimental techniques applied to IDPs, NMR spectroscopy is unique in providing multiple probes of dynamics, effectively covering most timescales from ps to days with atomic resolution. In particular, in the present work, we focus on motions occurring on tens of ps to tens of ns. Dynamics on these timescales is conveniently probed by nuclear spin relaxation. High magnetic fields are, in general, required to obtain sufficient spectral resolution and sensitivity in complex biomolecules, so relaxation is conventionally measured at high (>9.4 T) magnetic field (high-field [HF]) strengths in the case of proteins.

More precisely, spin relaxation probes the spectral density function J(ω) which is the Fourier transform of the correlation function, describing the power dissipated by the fluctuations as a function of frequency ω. Measured spin relaxation rates depend on the spectral density function evaluated at specific ω values dictated by spin physics, usually the interaction of nuclear spins with the strong magnetic field of the NMR magnet. For example, in the case of the relaxation of ^{15}N spins in the backbone of a protein used extensively in the present work, measured relaxation rates depend on linear combinations of the spectral density function evaluated at the following frequencies: 0, the Larmor frequency of ^{15}N spins (ω_N) and that of ^{1}H spins (ω_H). As both ω_N and ω_H are proportional to the magnetic field strength B_0, it is evident that a strategy to maximize the amount of information that can be extracted from relaxation studies is to measure relaxation at as many B_0 values as possible, as availability of instruments, resolution, and sensitivity requirements allow. It has been shown in a number of systems and experimental conditions (4,5,7,12,13) that HF relaxation data measured on disordered proteins at multiple fields can be convincingly interpreted in the framework of (extended) model-free (MF) analysis (14), whereby the spectral density function is decomposed into the sum of three Lorentzian components, characterized by their amplitudes (A_i) and timescales (τ_i). These three components are associated with fast bond librations on tens of ps, intermediate sampling of backbone dihedral angles on hundreds of ps, and slower chain-like motions occurring on timescales of several ns (9,12).

The timescales of these chain-like motions have been proposed to follow a continuous distribution (15,16). An analysis based on the interpretation of motions by projection onto an array of correlation times (IMPACT), which uses a discretized distribution of correlation times, could not identify slower motional modes or a possible tail of the distribution of correlation times for chain-like motions from the analysis of relaxation rates measured at high magnetic fields (5). It is still unclear whether the presence of additional motions is masked by the particular range of timescales HF spin relaxation is intrinsically most sensitive to, even when multiple HFs are combined. Because the sensitivity of NMR relaxation depends on the available transition frequencies of the spin system of interest, the range of magnetic fields at which relaxation is measured defines the range of timescales of the motions that can be probed by a particular set of relaxation rates (17–22). Measurements of ^{15}N relaxation at magnetic fields from 9.4 to 28 T can, in principle, probe motions in the range of tens of ps to tens of ns, with limited sensitivity above a few ns. The accurate and sensitive determination of motions slower than a few ns timescales requires measurement at magnetic fields lower than those allowed by high magnetic fields. Relaxation measurements have already been performed over a much broader range of magnetic fields down to 0.5 mT by sacrificing high-resolution and thus the ability to assign motions to any particular region of a protein (23).

Recently, high-resolution relaxometry, which combines relaxation at low magnetic fields and detection at high magnetic fields, emerged as a promising tool to expand the sampling of the spectral density function to otherwise inaccessible ω values while preserving the resolution and sensitivity of high-field NMR. Applications of high-resolution relaxometry to small folded proteins have demonstrated enhanced sensitivity to motions with correlations times of hundreds of ps to a few ns, both on the backbone (24,25) and side chains (26,27).

Insight into the nature of dynamic fluctuations that is complementary to experimental NMR studies can be provided by molecular dynamics (MD) simulations. In fact, MD trajectories are used to calculate rotational autocorrelation functions of NH bond vectors (and consequently J(ω) values) describing the reorientational properties of relaxation-active interactions such as dipole-dipole couplings and chemical shift anisotropy (CSA), thereby providing a direct and quantitative way of comparing simulated dynamics with experimental spin relaxation rates. In the case of unfolded proteins, MD simulations have been shown to be largely consistent with the three-timescale model derived from MF studies (9,28,29).

In this work, we perform high-resolution relaxometry, HF spin relaxation measurements, and MD simulations on a particularly challenging half structured and half disordered δ subunit of RNA polymerase. The δ subunit of RNA polymerase is a subunit unique for Gram-positive bacteria that was shown to be essential for virulency of some pathogens like *Streptococcus agalactiae* and *Staphylococcus aureus* (30), and it was demonstrated that the δ subunit causes a sensitivity of RNA polymerase activity to the concentration of initiating nucleoside triphosphates, making the δ subunit essential for rapid changes of gene expression (31). Here, the δ subunit of a model Gram-positive bacteria *Bacillus subtilis* is investigated. The δ subunit is composed from two domains: the N-terminal domain has a well-defined structure formed by a core of three α-helices, while the C-terminal domain is disordered (32), and it does not show any propensity to form any secondary structure element within its sequence. The C-terminal domain is highly negatively charged except a lysine-rich motif ^{96}KAKKKKAKK^{104} involved in the formation of transient electrostatic contacts with negatively charged amino acids in the remaining parts of the C-terminal domain (32).

We combine high-resolution relaxometry measurements with HF spin relaxation data to probe as broad a range of correlation times as possible in the disordered domain of the δ subunit of RNA polymerase. Although the high-field dataset analyzed here is relatively sparse, this combination also provides us with the opportunity to assess whether low-field measurements can identify additional dynamic processes to those derived from uniquely HF spin relaxation or indeed whether this information can be refined by probing the spectral density function at lower frequencies. We find that all measurements are consistent and mostly report on the same three motional processes described above. This observation is confirmed by a detailed analysis of motions sampled by multi-μs MD simulation. Our results indicate that small inconsistencies between the analyses of relaxation data measured at low and high field would be the hallmark of IDP dynamics more complex than the state-of-the-art MF model.

MATERIALS AND METHODS

Sample preparation

The δ subunit of RNA polymerase was prepared as uniformly ^{15}N and ^{13}C labeled recombinant protein expressed in *Escherichia coli* BL21(DE3) strain. The purification protocol is described elsewhere (33). NMR samples contain 1.2 mM of the protein in 20 mM phosphate buffer (pH = 6.6; uncorrected reading) and 10 mM NaCl and 10% D$_2$O.

Prepared NMR samples were degassed upon mild vacuum and sealed in a special tube (25) for the high-resolution relaxometry experiments. The design of the tube was modified compared with the reported prototype (25) to increase the active sample volume to 120 μL. HF NMR experiments were measured in standard 5 mm NMR tubes.

NMR measurements

All NMR experiments were performed at (300.0 ± 0.2) K, which was calibrated using temperature standards before every measurement.

High-resolution relaxometry relaxation rates were acquired with a 600 MHz spectrometer equipped with a special 3.2 mm room-temperature TXI probe and a shuttling device (25). The shuttling device allows sample movement along the bore of the magnet in order to reach a position corresponding to any desired magnetic field between 0.01 and 14.1 T, at which relaxation rates are to be measured. The pulse sequence for the measurement of high-resolution relaxometry rates (Fig. S1) is derived from the previously published pulse sequence for the determination of relaxation rates at 0.33 T using a two-field NMR spectrometer (34–37). ^{15}N longitudinal relaxation rates R$_1$ were measured at nine magnetic fields: 0.10, 0.33, 0.67, 1.00, 1.46, 2.00, 2.50, 4.00, and 6.00 T. Three-dimensional HNCO nonuniformly sampled (for ^{15}N and ^{13}C dimensions) spectra with varying relaxation delays (summarized in Table S1) were acquired for each magnetic field in an interleaved manner as a pseudo-four-dimensional experiment (experimental relaxometry relaxation rates are available in Tables S2–S4). An in-house program was used to generate the NUS schedule using Poisson discs sampling (38) with the Gauss-function used to weight the density of sampled points. A complete set of real and imaginary components for quadrature detection in indirect dimensions were acquired, each of them measured with four scans to accumulate the signal. Additional details about the experimental setup of the high-resolution relaxometry measurements can be found in Table S1.

Experiments for the determination of proton exchange rates with the solvent were carried out according to the RELAX-EXSY protocol (39) on samples of the delta subunit with identical composition but different contents of D$_2$O (1%, 10%, 30%, and 50% contents of D$_2$O was used). These experiments were carried out on a 600 MHz NMR spectrometer equipped with a cryo-cooled 5 mm TCI probe. Relaxation delays of 44.8, 179.2, 380.8, 627.2, and 1,030.4 ms were used.

Steady-state nuclear Overhauser effect (NOE) experiment and measurement of longitudinal relaxation rates were carried out on an 850 MHz NMR spectrometer equipped with a cryo-cooled 5 mm TCI probe (BMRB: 27245) using the published pulse sequence for uniformly ^{15}N and ^{13}C labeled disordered proteins (40). The longitudinal relaxation rates were measured with relaxation delays 0.0448, 0.0672, 0.112, 0.1792*, 0.2464, 0.3808, 0.784, and 1.232 s (the asterisk denotes the spectra repeated twice). The saturation in the steady-state NOE experiment was achieved with 5 s irradiation composed from inversion proton pulses separated by 22.22 ms, and the reference experiment was measured with a 15 s interscan delay (41).

NMR processing and data analysis

All NMR spectra were processed using NMRpipe (42) software v.9.9 and SMILE 2.0beta (43) for the nonuniformly sampled spectra.

No extrapolation (NUS zero-filling) was used in the processing of the nonuniformly sampled data, and the same signal downscaling factor was used for the independently processed spectra of various relaxation delays. The spectra were analyzed using the program NMRFAM-Sparky 1.413 (44). The extracted peak intensities of the high-resolution relaxometry spectra were fitted to mono-exponential decay functions in the Octave 3.8.2 program (45) using the function leasqr from the package optim.Errors of the fitted relaxation rates were estimated based on the smooth Bootstrap method (46,47). 300 Monte Carlo simulations for each Bootstrap sample were generated based on the estimated error of the peak intensities from the noise in the spectra. The steady-state NOE was determined from the ratio of the peak intensities in the spectra acquired with saturation and reference spectra. The error of signal intensities was estimated from the noise. The longitudinal relaxation rates were fitted in the Octave 4.0.3 program (45), and the error was obtained by Bootstrap method.

Analysis of relaxation rates

The high-resolution relaxometry rates were combined with previously published longitudinal relaxation rates and steady-state NOE acquired at 500 MHz spectrometer; longitudinal relaxation rates, steady-state NOE, and transverse and longitudinal cross-correlated cross-relaxation rates measured at 600 MHz (using a uniformly ^{15}N labeled sample (4)); longitudinal relaxation rates, steady-state NOE, and longitudinal and transverse cross-correlated cross-relaxation rates acquired with ^{13}C, ^{15}N labeled sample at 600 MHz (40); and longitudinal relaxation rates and steady-state NOE measured with the identical ^{13}C, ^{15}N labeled sample at 850 MHz. Transverse auto-relaxation rates were not used in this analysis due to possible contributions of chemical exchange (4), resulting in one single high-resolution relaxation rate that reports on J(0). Errors of the relaxation rates were estimated to be at least 2% of the relaxation rate.

The measured relaxometry relaxation rates and high-field relaxation rates were analyzed using the IMPACT approach (5). The relaxometry relaxation rates were corrected via the ICARUS protocol (25,48) to consider the effects of cross-relaxation during the sample shuttle transfers, stabilization delays, and relaxation delays during the measurements. A 1 ms time step was used in the ICARUS correction simulation. The accurate low-field (0.33 T) relaxation rates measured in a two-field NMR spectrometer (37) were not used in the analysis, so they could serve as an independent verification of the ICARUS correction procedure.

The applied IMPACT method follows the originally outlined methodology (5), but the analysis protocol was extended, and a grid search for the optimal ratio (ranging between 200 and 10,000 with steps of 200) between the smallest and the largest correlation time (ranging between 10 and 120 ns with steps of 10 ns) was included. The analysis was performed using inhouse script written in Mathematica 10.1.0 (49). The error of the optimized parameters of the selected variant of $J(\omega)$ was estimated based on 4,000 Monte Carlo simulations. The IMPACT analysis was also performed for individual residues separately using high-field data and the corrected relaxometry relaxation rates obtained after the last iteration of the ICARUS-IMPACT analysis.

MF analysis was carried out as previously described (12). Briefly, experimental relaxation rates were modeled using a spectral density function of the form $J(\omega) = 0.4\ \Sigma_i\ A_i\ \tau_i/(1 + (\omega\ \tau_i)^2)$, in which i = 1–3, the sum of the three amplitudes A_i is constrained to 1, and τ_1 is fixed to 45 ps. Three Lorentzian terms have been used because this number provided the best fit for all residues in previous studies of disordered proteins (7,12). Thus, the model effectively contains only four parameters to be determined (two amplitudes and two timescales). Monte Carlo simulations were used to quantify the error bars on fitted parameters. Chemical shielding tensor of amide ^{15}N was defined following published analysis (50) with the angle between the main tensor component and ^{15}N-^{1}H bond set to 21° (4).

MD simulations

MD simulations were performed with the software Gromacs 5.0 (51) using the force-field Amber99SB_ILDN (52). The simulations were run with explicit solvent water molecules TIP4P-D (6). The choice of the setup reflected previous analysis of the effect of the force-field and water molecule models (53). Thirteen MD trajectories were calculated with various initial conformations of the C-terminal domain. The starting structures are representative of the structural flexibility of the C-terminal domain, and they covered both extended conformations of the C-terminal domain and a more compact state that features transient electrostatic interactions between the IDR and the folded N-terminal domain. The charge of the system was neutralized using sodium and chloride ions, and additional Na^+ and Cl^- ions were added to match the experimental NaCl concentration (10 mM). The appropriate protonation of side chains in the MD simulations was checked experimentally by pH titration in the range 8 to 5.4 and analysis of the chemical shifts of side-chain carbons (54).

Each trajectory was divided into 22 nonoverlapping segments of variable length (100–120 ns). For each segment, the rotational correlation function of each NH bond vector was calculated and fitted to a predefined grid of timescales as previously described (8).

RESULTS AND DISCUSSION

Remarkably, in spite of very limited signal dispersion (32) and technical challenges associated with the high-resolution relaxometry setup, resolution and sensitivity of the high-resolution relaxometry experiments were sufficient to probe dynamics of 68 out of 90 nonproline residues in the disordered region of the delta subunit (in order to achieve spectra with well-resolved peaks suitable for a quantitative analysis, we performed relaxation measurements as pseudo-four-dimensional experiments using the HNCO correlation; the benefit of a better resolution is documented in Fig. S2).

The ICARUS-IMPACT analysis of the experimental data results in a description of the spectral density function in the form of $J(\omega) = 0.4\ \Sigma_i\ B_i\ T_i/(1 + (\omega\ T_i)^2)$ with i = 1–7. The seven T_i values span the range between 11 ps and 70 ns and are equidistant on a logarithmic scale (correlation times approximately 0.01, 0.05, 0.2, 0.9, 4, 16, and 70 ns). Experimental data and back-calculated values are in excellent agreement (Fig. S3–S5). The correction factors applied to consider the cross-relaxation pathways during the relaxometry experiment range from 3.0% to 11.5% depending on the experimental setup (the corrected relaxometry relaxation rates are in Tables S5–S7). These values are consistent with those previously reported for ubiquitin (25). To further verify the accuracy of the correction factors derived from the ICARUS procedure, we compared the relaxometry relaxation rates measured at 0.33 T with longitudinal relaxation rates measured at the same field (37) with a two-field NMR spectrometer (34,35), in which the effect of cross-relaxation pathways can be removed by means of ^{13}C and proton inversion pulses during the relaxation delay (36). While uncorrected high-resolution relaxometry rates are systematically lower than the values acquired suppressing cross-relaxation pathways (Fig. 1 *a*), the correction factors derived from the ICARUS protocol effectively restore a quantitative agreement with the data measured at a two-field NMR spectrometer (Fig. 1 *b*). Therefore, these correction factors are used throughout this work.

The result of IMPACT analysis provides insight into the distribution of timescales of motions. Of the seven fixed timescales, only four appear to contribute significantly to the spectral density function. The fastest motions on the order of several tens of ps B_1 and B_2 have a cumulated amplitude of about 0.3, with little if any sequence dependence (Fig. 2 *a* shows the joint amplitudes B_1+B_2; separated plots for B_1 and B_2 can be found in Fig. S6). B_4 amplitudes, associated with a timescale close to 1 ns (Fig. 2 *c*), are the largest contribution to relaxation for most residues and appear to monotonically increase from the end of the N-terminal folded domain to the C-terminus, indicating a gradual reduction of the effect of the folded domain on the dynamics of the IDR, despite the presence of long-range contacts of electrostatic nature (55). Finally, the fifth component associated with a timescale of the order of several ns (Fig. 2 *d*) is an important driver of relaxation for the residues close to the folded domain and those including and surrounding the lysine-rich stretch (K-tract, K96–K104), which is involved in charge-mediated interactions (55).

Of the remaining three timescales, values of $B_3 = 0.1$ are found for all residues outside the K-tract (Fig. 2 *b*), indicating that relatively fast dynamics might be quenched by long-range interactions in the K-tract. Similarly, nonzero values of B_6 are found in the proximity of the folded domain (Fig. 2 *e*), probably to compensate for the fact that

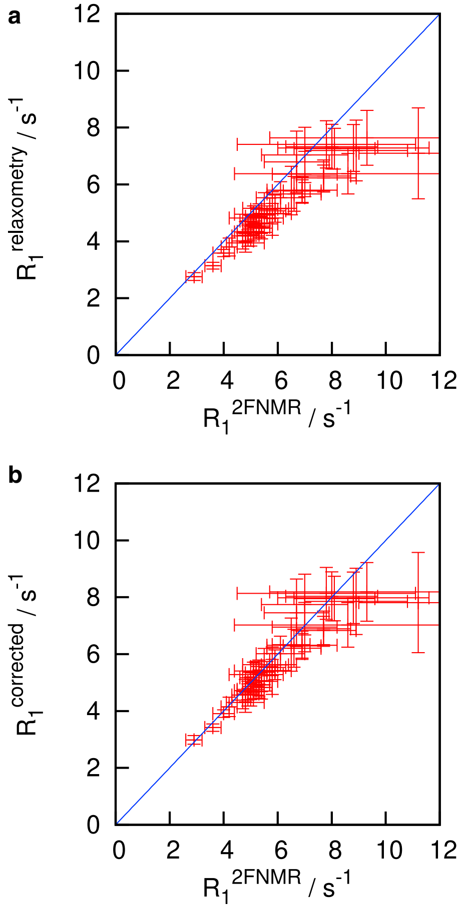

FIGURE 1 Correlation of the relaxation rates measured at 0.33 T (experimental uncertainties are indicated by error bars). Correlation of the longitudinal relaxation rates R_1^{2FNMR} measured with two-field NMR spectrometer and the relaxometry decay rates $R_1^{relaxometry}$ (*a*) and their corrected values $R_1^{corrected}$ (*b*).

$T_5 = 3.81$ ns is smaller than the actual rotational correlation time of the folded N-terminal domain $\tau_c = 5$ ns based on hydrodynamical calculations (56,57) (estimated for N-terminal domain of the δ subunit alone; the comparison of the relaxation data acquired for the separated N-terminal domain (58) and the full length δ subunit (4) shows a negligible effect of the C-terminal domain on the rotational diffusion of the N-terminal domain). This effect is similar to the one observed in the IMPACT analysis of HF relaxation data measured on the partially disordered Engrailed2 (5). Finally, the parameter B_7 (Fig. 2 *f*), associated with a correlation time on the orders of several tens of ns features very small values throughout the sequence. While these values are significantly different from zero for residues 110 to 135, which might suggest a tail of the distribution of correlation times with a nonzero density for correlation times $\tau > T_6 = 16.3$ ns, they never represent more than 1% of the angular correlation function/spectral density function. Similar conclusions can be obtained from the analysis done on a per-residue basis (the results are shown in Fig. S7). The results deviate from the monotonous dependence of the amplitudes on the correlation times of the Zimm model for polyelectrolytes (59). It documents a strong effect of electrostatic contacts on the conformation of the C-terminal domain of the delta subunit studied previously by small angle X-ray scattering and paramagnetic relaxation enhancement (32,55).

In order to characterize the extent of slow ns dynamics and quantify their potential contributions to NMR relaxation from a theoretical standpoint, we performed a pool of MD simulations as described in the materials and methods section and calculated relaxation rates averaged across the pool (Fig. S8 and S9). While the simulation is in good agreement with experimental relaxometry data, longitudinal relaxation rates measured at HF are systematically overestimated by the simulation, whereas the only rate that depends on J(0) in our dataset—the CSA/dipole-dipole cross-correlated cross-relaxation rates η_{xy} measured at 600 MHz—is significantly higher in the experiment than in the simulation. Taken together, these two observations indicate that protein dynamics are excessively fast in the simulation compared with in the experiment. We attribute this quantitative disagreement between experiment and simulation to our choice of water model (TIP4P-D), which has been shown to promote excessive flexibility in both folded and unfolded proteins (10).

We applied the average block selection using relaxation data (ABSURD) method to select, among the segments of trajectories in our pool, a subset of trajectories minimizing the root-mean-square deviation between experimental and simulated data (8). This approach accounts for conformational variability across conformers interconverting on timescales faster than the coalescence limit while filtering for poor sampling of slower motional modes responsible for unstable averaging of the autocorrelation function. All rates measured at 500, 600, and 850 MHz except heteronuclear NOEs were combined in the ABSURD target function for selection. Somewhat disappointingly, we obtain a subensemble of four out of a total of 44 segments. In addition, simulated relaxation rates evaluated from this subensemble (Figs. S10 and S11) are very close to those averaged over the entire pool, indicating that protein dynamics are accelerated in the entire pool of trajectories to a similar degree, rendering the ABSURD selection process very inefficient. Although the experimental sequence dependence of relaxation rates presented in Figs. S8–S11 is qualitatively reproduced by the ABSURD ensemble of trajectories, there are systematic differences that appear to report on systematic errors in the depiction of both ns motions (only one high-field longitudinal rate is correctly reproduced, the 500 and 600 MHz R_1 values are clearly not) and, most severely, the unique probe of slower motions (η_{xy}). Under these conditions, it would be hazardous to over-interpret the fact that relaxometry data are, apparently serendipitously, actually quite well reproduced by the simulation.

In an attempt to shed further light on the shortcomings of the simulation, we compare the dynamic fingerprint of the

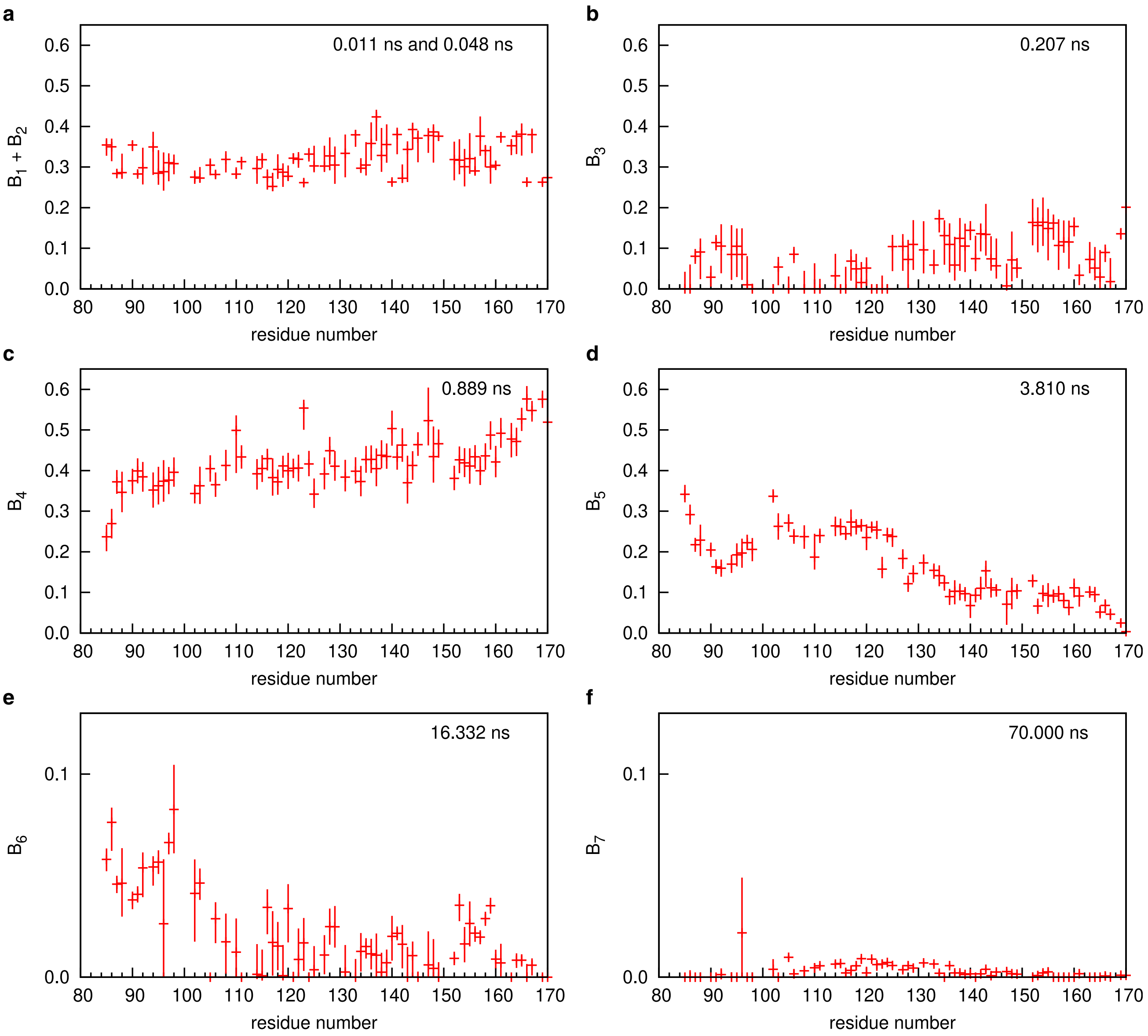

FIGURE 2 Fitted coefficients B_i (with error bars) used to model the distribution of the timescales. The coefficients are related to the terms of the spectral density function associated with correlation times 11 and 48 ps taken together (*a*), 0.2 ns (*b*), 0.9 ns (*c*), 3.8 ns (*d*), 16.3 ns (*e*), and 70 ns (*f*).

protein derived from the IMPACT analysis (Fig. 2) with an analog fingerprint extracted from the simulation.

To do so, for each NH bond vector, we average the rotational correlation functions calculated for each segment and fit the averaged autocorrelation function to a predefined grid of 128 timescales as previously described (8). The results of the fit are shown in Fig. 3. Besides a minor component at ∼5 ns, most ns dynamics in the folded domain (residues 1–90) occur at ∼3 ns, which is shorter than the expected rotational correlation time of the protein, highlighting yet again that our choice of water model does not reproduce protein hydrodynamics with sufficient accuracy. In both the folded domain and IDR, the two fastest motional processes occur around 10 and 100 ps, suggesting, assuming the simulation to be accurate, that the derived timescale of ∼50 ps often found in IDPs irrespective of their sequence and length (12,13) is in fact a value representative of a distribution of fast dynamics. In the IDR, we find two additional motional processes, one around 1 ns and a slower one, around 4 ns, that identify with the fourth and fifth component of the IMPACT analysis, respectively. We note that both these processes appear to occur on timescales that are close to, but well distinct from, the reorientation of the folded domain. This indicates that slow dynamics in IDRs is of segmental nature and decoupled from the rotational diffusion of the folded domain beyond the persistence length of several residues, implying that the use of “order parameters” is not the most informative for understanding

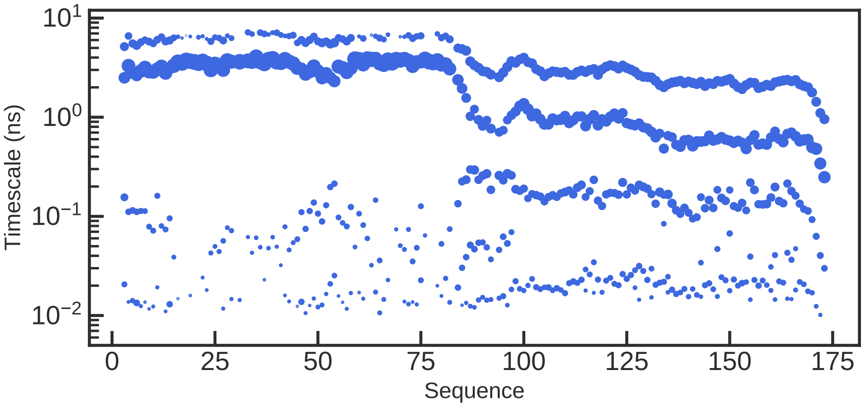

FIGURE 3 Motional timescales derived from fits of average autocorrelation functions to a predefined grid of 128 timescales, from 1 ps to 50 ns. The size of each dot is proportional to the amplitude of the associated timescale.

the dynamic richness of long IDRs. We suggest that a dynamic description in terms of multiple timescales and their associated amplitudes is a more natural framework to understand the functional mechanisms of disordered proteins.

We note that, despite a lack of quantitative agreement with the high-field data due to excessively fast multi-ns motions, the simulation, particularly with the use of ABSURD, reproduces relaxometry rates well. This demonstrates that relaxometry data can be described with a model that does not include >10 ns motions, as an alternative to the use of the correlation times T_6 and T_7 of the IMPACT analysis.

MF analysis of HF data and high-resolution relaxometry measurements simultaneously reproduces relaxation rates measured at low (Fig. S12) and high (Fig. S13) fields with good accuracy. The sequence dependence of fitted parameters (Fig. 4, we remind the reader that τ_1 is fixed to 45 ps) displays many features identified above. τ_2 is close to 900–1,000 ps for the entire sequence, with the exception of the region broadly centered around the K-tract. This observation is consistent with our interpretation of the third and fourth components of the IMPACT analysis. τ_3, which appears to be poorly defined because of the scarcity of information on J(0) in our dataset, does not show any evident sequence dependence, suggesting again that slow motions are segmental chain-like dynamics largely independent of the rotation of the N-terminal domain. We find that the product of A_3 and τ_3 is much more robustly determined in our analysis, in line with decades of applications of MF analysis to folded proteins. Consistently with B_1-B_2 features in IMPACT, the amplitude of fast motions (Fig. 4 *f*) increases only slightly going from the N- to the C-terminus. The amplitudes of intermediate (Fig. 4 *e*)

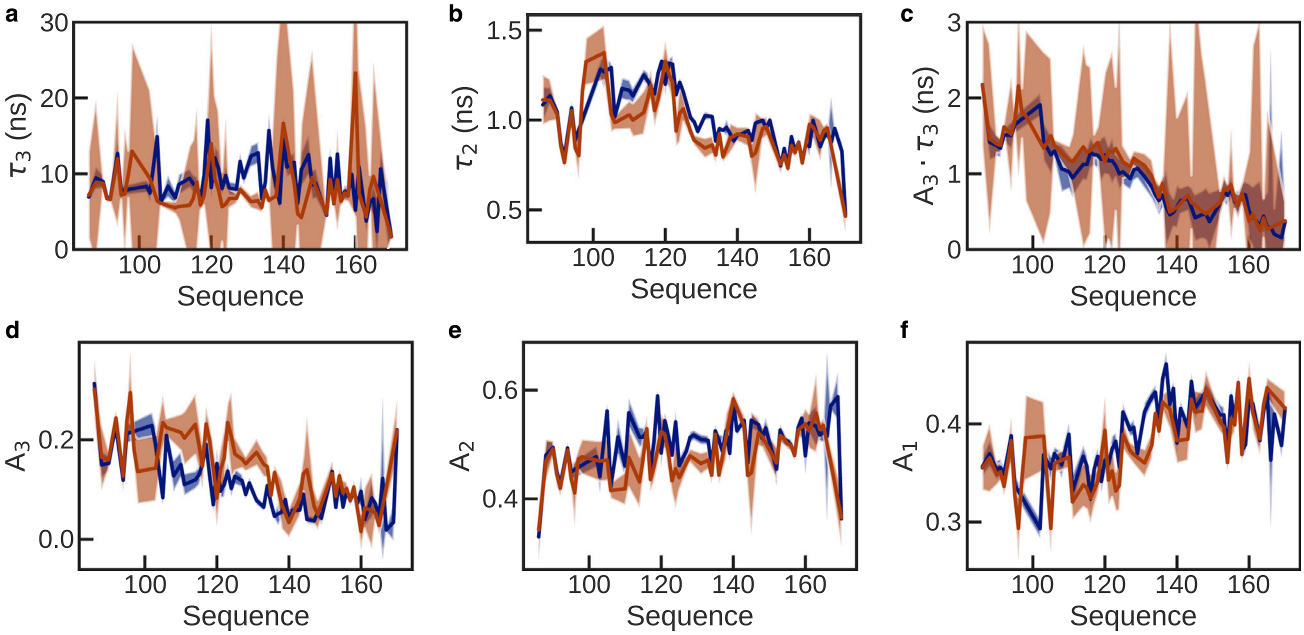

FIGURE 4 Motional timescales (*a* and *b*) and amplitudes (*d*–*f*) resulting from the model-free analysis of either the entire dataset comprising both high-resolution relaxometry and HF spin relaxation data (blue) or HF measurements only (orange). The product of A_3 and τ_3 is better defined than the two individual parameters (*c*). Shaded areas represent uncertainties on fitted parameters, estimated by Monte Carlo simulations.

and slow (Fig. 4 *d*) motions match B_4 and B_5 in the IMPACT analysis, respectively. Overall, the results are consistent with what was expected for an IDR of similar length in the absence of significant partially formed secondary structure elements.

Recent studies of the dynamics of IDPs have exploited full sets of R_1, R_2, ^{1}H-^{15}N NOE and cross- correlated CSA-dipole cross-relaxation η_{xy} at three or more magnetic fields. The relative sparsity of experimental data available at high field in this study precludes a systematic comparison of the information content present in high- and low-field relaxometry data. We have nevertheless repeated the MF analysis using only the spin relaxation data measured at 500, 600, and 850 MHz. Rates back calculated from the results of this MF analysis closely match the experimental values (Fig. S14). Interestingly, cross-validation of the MF results using the relaxometry data, which are not actively used in the fit, reveals that the overall trend of longitudinal relaxation at low field can be predicted with good accuracy from the information about dynamics encoded in the high-field data in most cases (Fig. S15). Such overall agreement is expected since high-field relaxation probes the spectral density function, a monotonous, decreasing function of the frequency, down to 50 MHz and at zero frequency. In addition, the precision of low-field relaxation rates is much lower than that of high-field rates due to the limited sensitivity of the shuttle and probe apparatus (about 20% of the sensitivity of a room-temperature probe or ~7% of the 600 MHz spectrometer used for high-field relaxation measurements).

The spectral density at zero frequency J(0) obtained from high-field measurements probes the slowest motions in a way where amplitude and correlation times are convoluted. The magnetic-field dependence of low-field relaxation rates should therefore provide novel information concerning small, but statistically significant, differences between experimental relaxometry relaxation rates and those predicted by the analysis of high-field relaxation alone. The presence of additional information contained in relaxometry relaxation rates is suggested by a direct comparison of the MF parameters obtained by an analysis of high-field relaxation alone and with the addition of relaxometry data (Fig. 4), which have different profiles in the region between 105 and 137 (results are listed in Tables S8 and S9, respectively). Both analyses provide similar estimates of the product of A_3 and τ_3, which is just slightly lower in the analysis of the full dataset (Fig. 4 *c*). Not surprisingly, this product is defined with more precision because of the information about the spectral density values close to 0, due to the higher number of relaxometry data at the lowest fields reporting on slower motions. However, overall A_3 parameters are lower and τ_3 higher in this region when the relaxometry data are included (Fig. 4 *a* and *d*), which could be numerically similar to the appearance of very small contribution of the longest correlation time T_7 in the ICARUS-IMPACT analysis of the full set of high- and low-field relaxation rates. We note, however, that it is again difficult to make any meaningful comparison because of the limited size of the high-field data set (only 6 rates are available for the determination of 4 independent parameters).

The region spanning residues 105–140 was shown (55) to electrostatically interact with the K-tract and with the structured part of the δ subunit. The electrostatic interactions with the K-tract are essential for the role of the δ subunit in regulation of RNA polymerase activity by a stabilization of initiation complexes of RNA polymerase-DNA interaction (55). The accurate description of the dynamics in this region is therefore of a particular interest. If confirmed by further investigations, motions with an effective timescale of 10 ns or more, such as the contribution for T_7 in the ICARUS-IMPACT analysis, could be related to the formation of a temporarily compacted conformation of the part of the C-terminal domain involving the positively charged K-tract and the part of the acidic regions. Yet, the contribution of such motions is very small in the ICARUS-IMPACT analysis and therefore not reflected by MF analysis and not captured by MD simulations.

CONCLUSIONS

We presented a detailed comparison of complementary NMR- and computational-based approaches to probe and describe ps-ns dynamics in disordered systems using a well-characterized model system. Overall, our results indicate that three timescales and their associated amplitudes are in general sufficient to model most observations in vitro and in silico. While this has been shown already in a number of IDPs, here we add an example of a disordered region in a partially folded protein, and we show that the presence of the folded domain does not increase the complexity of the dynamic features, which can be captured by the tools and models developed specifically for IDPs.

Our study combines ^{15}N relaxation measured at atomic resolution of an IDP over an unprecedented range of magnetic field strengths and therefore a broader spectrum of timescales of reorientational modes than has been available until now. We observe that the model assuming three major contributions to the spectral density function can predict the relaxometry data measured down to 4.26 MHz with good accuracy in most cases, confirming the overall robustness of this commonly used framework. The potential for refining more complex motional models by including relaxometry data into such an analysis is, however, self-evident. Although our study suffers from suboptimal sampling of high-field data and force-field inaccuracies, which have been amply discussed in the literature, our analysis nevertheless provide tantalizing indications of the potential of combining high-field relaxation and high-resolution relaxometry to enhance sampling of the spectral density function at frequencies not accessible to HF relaxation measurements, in particular with respect to the details of slower, possibly functionally important, motions in the δ subunit of RNA polymerase. We expect that further studies applying high-resolution relaxometry with enhanced sensitivity or other techniques will help to further refine motions in this important family of proteins beyond the model with three ps-ns effective processes.

SUPPORTING MATERIAL

Supporting material can be found online at https://doi.org/10.1016/j.bpj.2022.09.016.

AUTHOR CONTRIBUTIONS

N.S., V.Z., M.B., and P.K. analyzed data; Z.J. and S.N. prepared NMR samples; V.Z., M.Z., J.-M.T., T.M., L.Z., F.F., and P.K performed NMR experiments; V.Z., Z.J., P.P., and P.K. processed and analyzed NMR spectra; N.S., L.Z., M.B., F.F., and P.K. wrote the manuscript; F.F., M.B., and P.K. obtained funding for the project.

ACKNOWLEDGMENTS

This work was supported by Czech Science Foundation grant no. GJ18-04197Y (to P.K., V.Z., and Z.J.), from European Regional Development Fund-Project MSCAfellow2@MUNI, No. CZ.02.2.69/0.0/0.0/18_070/0009846 (to P.P. and P.K.), and grant no. ANR-18-CE29-0003 (NANO-DISPRO), provided by Agence Nationale de la Recherche (to F.F. and M.B.). Short scientific mission of P.K. to perform measurements at the NMR spectrometer allowing high-resolution relaxometry experiments was supported by a STSM grant from the EURELAX COST Action CA15209. CIISB, Instruct-CZ Centre of Instruct-ERIC EU consortium, funded by MEYS CR infrastructure project LM2018127, is gratefully acknowledged for the financial support of the measurements at the Josef Dadok National NMR Centre.

DECLARATION OF INTERESTS

T.M. and J.-M.T. were employees of the Bruker BioSpin. The other authors declare no other conflict of interest.

REFERENCES

1. Henzler-Wildman, K., and D. Kern. 2007. Dynamic personalities of proteins. *Nature*. 450:964–972.
2. Dyson, H. J., and P. E. Wright. 2005. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* 6:197–208.
3. Jensen, M. R., M. Zweckstetter, …, M. Blackledge. 2014. Exploring free-energy landscapes of intrinsically disordered proteins at atomic resolution using NMR spectroscopy. *Chem. Rev.* 114:6632–6660.
4. Kadeřávek, P., V. Zapletal, …, L. Žídek. 2014. Spectral density mapping protocols for analysis of molecular motions in disordered proteins. *J. Biomol. NMR*. 58:193–207.
5. Khan, S. N., C. Charlier, …, F. Ferrage. 2015. Distribution of pico- and nanosecond motions in disordered proteins from nuclear spin relaxation. *Biophys. J.* 109:988–999.
6. Piana, S., A. G. Donchev, …, D. E. Shaw. 2015. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *J. Phys. Chem. B*. 119:5113–5123.
7. Gill, M. L., R. A. Byrd, and A. G. Palmer, III. 2016. Dynamics of GCN4 facilitate DNA interaction: a model-free analysis of an intrinsically disordered region. *Phys. Chem. Chem. Phys.* 18:5839–5849.
8. Salvi, N., A. Abyzov, and M. Blackledge. 2016. Multi-timescale dynamics in intrinsically disordered proteins from NMR relaxation and molecular simulation. *J. Phys. Chem. Lett.* 7:2483–2489.
9. Salvi, N., A. Abyzov, and M. Blackledge. 2017. Analytical description of NMR relaxation highlights correlated dynamics in intrinsically disordered proteins. *Angew. Chem., Int. Ed. Engl.* 56:14020–14024.
10. Robustelli, P., S. Piana, and D. E. Shaw. 2018. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc. Natl. Acad. Sci. USA*. 115:E4758–E4766.
11. Thomasen, F. E., F. Pesce, …, K. Lindorff-Larsen. 2022. Improving martini 3 for disordered and multidomain proteins. *J. Chem. Theor. Comput.* 18:2033–2041.
12. Abyzov, A., N. Salvi, …, M. Blackledge. 2016. Identification of dynamic modes in an intrinsically disordered protein using temperature-dependent NMR relaxation. *J. Am. Chem. Soc.* 138:6240–6251.
13. Adamski, W., N. Salvi, …, M. Blackledge. 2019. A unified description of intrinsically disordered protein dynamics under physiological conditions using NMR spectroscopy. *J. Am. Chem. Soc.* 141:17817–17829.
14. Clore, G. M., A. Szabo, …, A. M. Gronenborn. 1990. Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J. Am. Chem. Soc.* 112:4989–4991.
15. Buevich, A. V., and J. Baum. 1999. Dynamics of unfolded proteins: incorporation of distributions of correlation times in the model free analysis of NMR relaxation data. *J. Am. Chem. Soc.* 121:8671–8672.
16. Hsu, A., F. Ferrage, and A. G. Palmer. 2018. Analysis of NMR spin-relaxation data using an inverse Gaussian distribution function. *Biophys. J.* 115:2301–2309.
17. Bloembergen, N., E. M. Purcell, and R. V. Pound. 1948. Relaxation effects in nuclear magnetic resonance absorption. *Phys. Rev.* 73:679–712.
18. Wangsness, R. K., and F. Bloch. 1953. The dynamical theory of nuclear induction. *Phys. Rev.* 89:728–739.
19. Solomon, I. 1955. Relaxation processes in a system of two spins. *Phys. Rev.* 99:559–565.
20. Redfield, A. G. 1957. On the theory of relaxation processes. *IBM J. Res. Dev.* 1:19–31.
21. Abragam, A. 1961. The Principles of Nuclear Magnetism. Clarendon Press.
22. Smith, A. A., M. Ernst, and B. H. Meier. 2017. Because the light is better here: correlation-time analysis by NMR spectroscopy. *Angew. Chem., Int. Ed. Engl.* 56:13590–13595.
23. Parigi, G., N. Rezaei-Ghaleh, …, C. Luchinat. 2014. Long-range correlated dynamics in intrinsically disordered proteins. *J. Am. Chem. Soc.* 136:16201–16209.
24. Clarkson, M. W., M. Lei, …, D. Kern. 2009. Mesodynamics in the SARS nucleocapsid measured by NMR field cycling. *J. Biomol. NMR*. 45:217–225.
25. Charlier, C., S. N. Khan, …, F. Ferrage. 2013. Nanosecond time scale motions in proteins revealed by high-resolution NMR relaxometry. *J. Am. Chem. Soc.* 135:18665–18672.
26. Cousin, S. F., P. Kadeřávek, …, F. Ferrage. 2018. Time-resolved protein side-chain motions unraveled by high-resolution relaxometry and molecular dynamics simulations. *J. Am. Chem. Soc.* 140:13456–13465.
27. Smith, A. A., N. Bolik-Coulon, …, F. Ferrage. 2021. How wide is the window opened by high-resolution relaxometry on the internal dynamics of proteins in solution? *J. Biomol. NMR*. 75:119–131.
28. Prompers, J. J., and R. Brüschweiler. 2002. General framework for studying the dynamics of folded and nonfolded proteins by NMR relaxation spectroscopy and MD simulation. *J. Am. Chem. Soc.* 124:4522–4534.

29. Salvi, N., A. Abyzov, and M. Blackledge. 2019. Solvent-dependent segmental dynamics in intrinsically disordered proteins. *Sci. Adv.* 5:eaax2348.

30. Seepersaud, R., R. H. V. Needham, …, A. L. Jones. 2006. Abundance of the δ subunit of RNA polymerase is linked to the virulence of Streptococcus agalactiae. *J. Bacteriol.* 188:2096–2105.

31. Rabatinová, A., H. Šanderová, …, L. Krásný. 2013. The δ subunit of RNA polymerase is required for rapid changes in gene expression and competitive fitness of the cell. *J. Bacteriol.* 195:2603–2611.

32. Papoušková, V., P. Kadeřávek, …, L. Žídek. 2013. Structural study of the partially disordered full-length δ subunit of RNA polymerase from Bacillus subtilis. *Chembiochem.* 14:1772–1779.

33. López de Saro, F. J., A. Y. Woody, and J. D. Helmann. 1995. Structural analysis of the Bacillus subtilis delta factor: a protein polyanion which displaces RNA from RNA polymerase. *J. Mol. Biol.* 252:189–202.

34. Cousin, S. F., P. Kadeřávek, …, F. Ferrage. 2016. Recovering invisible signals by two-field NMR spectroscopy. *Angew. Chem., Int. Ed. Engl.* 55:9886–9889.

35. Cousin, S. F., C. Charlier, …, F. Ferrage. 2016. High-resolution two-field nuclear magnetic resonance spectroscopy. *Phys. Chem. Chem. Phys.* 18:33187–33194.

36. Kadeřávek, P., N. Bolik-Coulon, …, F. Ferrage. 2019. Protein dynamics from accurate low-field site-specific longitudinal and transverse nuclear spin relaxation. *J. Phys. Chem. Lett.* 10:5917–5922.

37. Jaseňáková, Z., V. Zapletal, …, P. Kadeřávek. 2020. Boosting the resolution of low-field $${15}\hbox {N}$$ relaxation experiments on intrinsically disordered proteins with triple-resonance NMR. *J. Biomol. NMR.* 74:139–145.

38. Kazimierczuk, K., A. Zawadzka, and W. Koźmiński. 2008. Optimization of random time domain sampling in multidimensional NMR. *J. Magn. Reson.* 192:123–130.

39. Lopez, J., P. Ahuja, …, G. Lippens. 2014. H/D exchange of a 15N labelled Tau fragment as measured by a simple Relax-EXSY experiment. *J. Magn. Reson.* 249:32–37.

40. Srb, P., J. Nováček, …, L. Žídek. 2017. Triple resonance ^{15}N NMR relaxation experiments for studies of intrinsically disordered proteins. *J. Biomol. NMR.* 69:133–146.

41. Ferrage, F., D. Cowburn, and R. Ghose. 2009. Accurate sampling of high-frequency motions in proteins by steady-state 15N−{1H} nuclear overhauser effect measurements in the presence of cross-correlated relaxation. *J. Am. Chem. Soc.* 131:6048–6049.

42. Delaglio, F., S. Grzesiek, …, A. Bax. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR.* 6:277–293.

43. Ying, J., F. Delaglio, …, A. Bax. 2017. Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data. *J. Biomol. NMR.* 68:101–118.

44. Lee, W., M. Tonelli, and J. L. Markley. 2015. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics.* 31:1325–1327.

45. Eaton, J. W., D. Bateman, and S. Hauberg. 2009. GNU Octave Version 3.0.1 Manual: A High-Level Interactive Language for Numerical Computations. CreateSpace Independent Publishing Platform.

46. Tibshirani, R. 1988. Variance stabilization and the Bootstrap. *Biometrika.* 75:433–444.

47. Hall, P. 1988. Theoretical comparison of Bootstrap confidence intervals. *Ann. Stat.* 16:927–953.

48. Bolik-Coulon, N., P. Kadeřávek, …, S. F. Cousin. 2020. Theoretical and computational framework for the analysis of the relaxation properties of arbitrary spin systems. Application to high-resolution relaxometry. *J. Magn. Reson.* 313:106718.

49. Wolfram Research, Inc. 2015. Mathematica. 10.1. Champaign Illinois.

50. Loth, K., P. Pelupessy, and G. Bodenhausen. 2005. Chemical shift anisotropy tensors of carbonyl, nitrogen, and amide proton nuclei in proteins through cross-correlated relaxation in NMR spectroscopy. *J. Am. Chem. Soc.* 127:6062–6068.

51. Abraham, M. J., T. Murtola, …, E. Lindahl. 2015. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 1–2:19–25.

52. Lindorff-Larsen, K., S. Piana, …, D. E. Shaw. 2010. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins.* 78:1950–1958.

53. Zapletal, V., A. Mládek, …, J. Hritz. 2020. Choice of force field for proteins containing structured and intrinsically disordered regions. *Biophys. J.* 118:1621–1633.

54. Tollinger, M., J. D. Forman-Kay, and L. E. Kay. 2002. Measurement of side-chain carboxyl pK(a) values of glutamate and aspartate residues in an unfolded protein by multinuclear NMR spectroscopy. *J. Am. Chem. Soc.* 124:5714–5717.

55. Kubáň, V., P. Srb, …, L. Žídek. 2019. Quantitative conformational analysis of functionally important electrostatic interactions in the intrinsically disordered region of delta subunit of bacterial RNA polymerase. *J. Am. Chem. Soc.* 141:16817–16828.

56. García de la Torre, J., M. L. Huertas, and B. Carrasco. 2000. HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson.* 147:138–146.

57. García De La Torre, J., M. L. Huertas, and B. Carrasco. 2000. Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J.* 78:719–730.

58. Kadeřávek, P., C. Diehl, …, M. Akke. 2011. Complementation of 3D structure of delta subunit of RNA polymerase from Bacillus subtilis with description of internal motions in terms of reduced spectral density mapping. *Mater. Struct.* 18:3–5.

59. Zimm, B. H. 1956. Dynamics of polymer molecules in dilute solution: viscoelasticity, flow birefringence and dielectric loss. *J. Chem. Phys.* 24:269–278.