Investigating the relationship between intrinsic protein disorder and biological function : mapping interaction trajectories at atomic resolution using nuclear magnetic resonance spectroscopy and molecular simulation

Wiktor Adamski

To cite this version:

Wiktor Adamski. Investigating the relationship between intrinsic protein disorder and biological function : mapping interaction trajectories at atomic resolution using nuclear magnetic resonance spectroscopy and molecular simulation. Biological Physics [physics.bio-ph]. Université Grenoble Alpes [2020-..], 2021. English. NNT : 2021GRALY057 . tel-03577038

HAL Id: tel-03577038
https://theses.hal.science/tel-03577038
Submitted on 16 Feb 2022

Université
Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : Physique pour les Sciences du Vivant
Arrêté ministériel : 25 mai 2016

Présentée par

Wiktor ADAMSKI

Thèse dirigée par **Martin BLACKLEDGE**

préparée au sein du **Laboratoire Institut de Biologie Structurale**
dans **l'École Doctorale Physique**

Détermination de la relation entre désordre intrinsèque des protéines et fonction biologique: description à l'échelle atomique des trajectoires d'interactions moléculaires par résonance magnétique nucléaire (RMN) et la simulation moléculaire

Investigating the relationship between intrinsic protein disorder and biological function: Mapping interaction trajectories at atomic resolution using nuclear magnetic resonance spectroscopy and molecular simulation

Thèse soutenue publiquement le **8 octobre 2021**,
devant le jury composé de :

Madame Cecile BREYTON
DIRECTEUR DE RECHERCHE, CNRS, Examinatrice
Monsieur Fabien FERRAGE
DIRECTEUR DE RECHERCHE, CNRS, Rapporteur
Monsieur Arthur G. PALMER III
PROFESSEUR, Columbia University, Rapporteur
Madame THérèse MALLIAVIN
DIRECTEUR DE RECHERCHE, CNRS, Examinatrice
Monsieur Christian GRIESINGER
PROFESSEUR, NMR based Structural Biology, Max Planck Institute for biophysical Chemistry, Examinateur
Monsieur Marc Jamin
PROFESSEUR, UGA, Président

Résumé

La thèse décrit la complexité de la dynamique locale des protéines dépliées sur des échelles de temps ps-ns sondées par relaxation RMN ^{15}N dans des environnements macromoléculaires complexes dans une gamme de conditions. L'application de l'analyse sans modèle dépendant de la température aux taux de relaxation du squelette ^{15}N révèle l'existence de trois modes dynamiques dominants représentant les distributions des échelles de temps des mouvements. La sensibilité expérimentale fixe la limite de la précision de l'estimation des modes dynamiques à partir des taux de relaxation RMN suite à l'application de l'analyse sans modèle dans les limites de l'approximation adiabatique. Les profils dynamiques hétérogènes et distincts des distributions des échelles de temps des mouvements présentés par les unités peptidiques des protéines dépliées à l'étude, des régions de protéines présentant une dominance de la dynamique ps-ns rapide pour les résidus situés dans la queue ultra-flexible de Map kinase kinase 4 à dynamique ps-ns lente en vedette élément déplié impliqué dans l'échange rapide avec les intermédiaires repliés du fragment C-terminal de la nucléoprotéine du virus Sendai, réaffirme l'affectation précédente des modes dynamiques aux mouvements de libration dans un puits de potentiel plat, échantillonnage dièdre local de l'épine dorsale et corrélé mouvements de plusieurs unités peptidiques (mouvements segmentaires). La modification systématique de la composition du milieu par addition de viscogène permet d'évaluer l'influence de l'encombrement sur la dynamique locale du squelette de trois protéines dépliées, révélant à son tour un couplage étroit des mouvements locaux du squelette avec la dynamique du solvant sur différentes durées et échelles de temps. Les constantes de proportionnalité associées à la décélération relative des fluctuations du squelette corrélées à la perturbation rotationnelle du solvant (coefficients de frottement) résultant de l'encombrement révèlent une dépendance de séquence marquée, qui est influencée par l'architecture moléculaire de la protéine dépliée. L'analyse est entièrement cohérente envers l'analyse sans modèle dépendant de la température développée précédemment et valide dans les limites de l'approximation adiabatique. L'orthogonalité relative des deux analyses permet de combiner l'expression reliant la décélération relative des unités peptidiques en fonction de la température et de la viscosité à une seule équation d'Arrhenius-Einstein, permettant son application facile à l'analyse sans modèle de ^{15}N taux de relaxation dans une gamme de conditions et de paramètres macromoléculaires. Relier les changements relatifs de viscosité à la perturbation relative de la dynamique du solvant, plutôt qu'à la concentration du viscogène permet de prédire la dynamique locale des protéines dépliées dans des environnements macromoléculaires complexes de nature physico-chimique différente, au-delà de la gamme des conditions étudiées, avec application dans la prédiction des fluctuations du squelette NH dans les phases cellulaires et condensées. Le travail présenté met en lumière la dynamique complexe présentée par les protéines intrinsèquement désordonnées dans des conditions environnementales complexes sur des échelles de temps de la pico à la nanoseconde.

ABSTRACT

The Thesis describes the intricacy of local dynamics of unfolded proteins on ps-ns timescales probed by ^{15}N NMR relaxation in complex macromolecular environments across a range of conditions. Application of temperature-dependent Model-Free analysis to ^{15}N backbone relaxation rates reveals existence of three dominant dynamic modes representing distributions of timescales of motions. The experimental sensitivity sets the limit on the accuracy of estimation of dynamic modes from NMR relaxation rates following application of Model-Free analysis within the limits of adiabatic approximation. The heterogenous and distinct dynamic profiles of the distributions of timescales of motions featured by peptide units of unfolded proteins under investigation, from regions of protein featuring dominance of fast ps-ns dynamics for residues located in the ultra-flexible tail of Map kinase kinase 4 to slow ps-ns dynamics featured unfolded element involved in the fast exchange with folded intermediates of C-terminal fragment of Sendai Virus Nucleoprotein, reaffirm previous assignment of the dynamic modes to librational motions in a flat potential well, local dihedral sampling of the backbone and correlated motions of several peptide units (segmental motions). Systematic modification of the composition of the medium by addition of viscogen allows to assess the influence of crowding on local backbone dynamics of three unfolded proteins, in turn revealing tight coupling of local backbone motions to the solvent dynamics on different length and timescales. The proportionality constants associated with relative deceleration of correlated backbone fluctuations to the rotational perturbation of the solvent (friction coefficients) arising due to crowding reveal marked sequence-dependence, which are influenced by the molecular architecture of the unfolded protein. The analysis is fully consistent with the previously developed temperature-dependent Model-Free analysis and valid within the bounds of adiabatic approximation. The relative orthogonality of the two analyses allows to combine the expression relating relative deceleration of peptide units as a function of temperature and viscosity to a single Arrhenius-Einstein equation, allowing its facile application to the Model-Free analysis of ^{15}N relaxation rates across a range of conditions and macromolecular settings. Relating the relative changes in viscosity to the relative perturbation of the solvent dynamics, rather than to the concentration of the viscogen allows to predict the local dynamics of unfolded proteins in complex macromolecular environments of different physicochcemical nature, beyond the range of conditions studied, with application in prediction of NH-backbone fluctuations in cell and in condensed phases. The presented work sheds light on the complex dynamics exhibited by intrinsically disordered proteins in complex environmental conditions on pico to nanosecond timescales.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

Martin and Malene

Being able to pursue research goals and merge the scientific ideas and goals encountered in my scientific development, from the first year of Bachelors to the end of Masters degree, has allowed me to find my life goal and contribute to the Scientific community in a way that exceeded my undergraduate expectations. Coming from different and diverse areas of research, the scientific path that You have invited me on has allowed me to form a unique set of perspectives on the research in the field. The familial atmosphere in the Lab has allowed me to relax into the research project and pursue calm scientific development, while being able to develop further understanding in the analytical methods guiding the interpretation of experimentally accessible parameters, which development remains at the boundaries of different domains of science. Being able to develop new points of understanding of the methods applied, introduce a set of ideas associated with interpretation of the physical phenomena to the scientific community based on them, and defend the perspectives representing fundamental aspects of methods with which we can develop a general framework of understanding of mechanisms underlying functioning of the biomolecular machineries. The experience You have allowed me to gather throughout my Doctoral experience has shaped me on multiple levels allowing to view the Research in more complete set of perspectives and immerse in the international Scientific community. Seeing You as Group Leaders, able to merge different aspects of Science, from scientific subjects through collaborative effort of community members both within and between Research fields, fostering Scientific development of members of the community has inspired me for pursuing challenges in my Academic path.

Nicola, Sima, Sigrid, Damien, Laura Luiza, Anas, Frank, Loïc, Aldo, Thibault and Carmen

Appreciation of every minute of work and pursuit of the ideas would not be possible without frequent, day-to-day discussions on the subject relating to the development and application of the methods further understanding of the subject. The day to day discussions that we had have given me an additional set of perspectives on the methods applied and functioning of the scientific community. Expanding the levels of understanding of the methods applied, would not have been possible without initial guidance into the subject and gradual development of fundamental understanding of principles underlying biochemistry, biophysics, NMR spectroscopy and MD simulations. The discussions we had allowed me to view the subject of my studies from different perspectives and develiop complementary understanding of the dynamics of the underlying biological processes and recapitulate my knowledge on the subject.

Christian, Arthur, Fabien and Rafael

The schools, conferences and meetings featuring presentations and poster sessions we had accelerated my progress, understanding and were verifying my knowledge in the intermediate stages of my Doctoral experience.

Florent, Nina, Lukas, Mizar, Joanna, Julliette and Mark

The discussions on scientific perspectives gathered during occasional local conferences and meetings, fostering formation of local collaborative network, allowing to present and share scientific ideas to the community prior to communication of the ideas to the broader audience.

THESIS OUTLINE

The intricacy of the intrinisically disordered (unfolded) proteins featuring highly heterogenous dynamics involved in functional binding equilibria requires understanding of the complexity of motions exhibited by unfolded protein fragments and underlying distributions of timescales of motions representing reoerientation rates of peptide units and their modifications associated with intra and intermolecular interactions. Nuclear Magnetic Resonance relaxation and Molecular Dynamics simulations allow to access their conformational equilibria on multiple length and timescales with atomic resolution. The reorientation of peptide planes may be represented by the dominant dynamic modes (timescales and amplitudes of the modes of distributions of timescales of motions with respect to experimental sensitivity) estimated from Model-Free analysis (featuring no assumptions on the complexity of underlying motions) of NMR relaxation rates originating from spin transtions arising due to the stochastically modulated motions. Further modifications of NH-backbone dynamics associated with changes in the macromolecular environement featuring crowding induced modification of viscosity rendering modulation to size and shape of the unfolded proteins and associated with it potential for phase separation, set the context for studying unfolded protein dynamics in highly complex, increasingly physiologically relevant environments.

Chapter 1 introduces the reader to the general context associated with complex dynamics featured by unfolded proteins in increasingly relevant complex macromolecular environments and methods used for their estimation.

Chapter 2 presents the theoretical basis of NMR relaxation used to access the information on ps-ns dynamics exhibited by reorienting peptide units - from spin and spin density operator and associated with it evolution, through second-order perturbation description of stochastically modulated motions-induced spin relaxation via one of the mechanisms accessible for the spin half-integer nuclei.

Chapter 3 presents fundamental aspects of Model-Free analysis and alternative methods associated with analysis of NMR relaxation rates, as well as, their limitations. The following part presents limitations associated with interpretation of NMR relaxation rates and statistical independence (adiabatic approximation) of stochastically modulated motions implicit in the analysis within the framework of Model-Free analysis and exemplifies averaging of timescales of motions with respect to the sensitivity of the NMR relaxation rates to different decades of timescales on the model systems used in the study.

Chapter 4 describes the temperature modulation of the ps-ns unfolded protein dynamics estimated from Model-Free analysis of NMR relaxation rates within the limits of adiabatic approximation and interpretation of dynamic modes in terms of distributions of timescales of motions. The temperature dependence of dynamic modes reveals distinct distributions of timescales of motions associated with reorientation of peptide units for structurally heterogenous unfolded proteins Map kinase kinase 4 (MKK4) and C-terminal part of Sendai Virus Nucleoprotein (SeV NTail).

Chapter 5 features viscosity-associated modulations to dynamic modes of SeV NTail estimated from viscosity and temperature-dependent (Einstein-Arrhenius) Model-Free analysis, limits associated with its application and the tight coupling between crowding-induced deceleration of peptide units and the slow-down of the reorientation rate of the solvent and potential origin of the proportionality constants describing relative slow-down of the NH-backbone motions with resepct to the solvent (friction coefficients).

Chapter 6 demonstrates application of the Einsten-Arrhenius Model-Free analysis to prediction of unfolded protein dynamics microinjected to *Xenopus oocytes* (MKK4, *in vivo*) and in concentrated phases originating from phase separated proteins (Measles Virus NTail). The exemplification of friction coefficients on three distinct unfolded proteins leads to estimation of their physical origin associated with internal and external limitations imposed on reorientation of peptide units both in dilute solutions and in complex macromolecular environment.

CHAPTER 1

Introduction

Proteins play a critical role in biology, as they carry out a vast array of functions, including for instance sustaining the cellular structure, synthesis, processing, transport and storage of biomolecules, regulation of biochemical reactions and immune responses. In the spirit of the structure-function paradigm which laid the foundation of structural biology, the unstructured parts of protein were deemed unimportant for function for many years. However, in the last 20 years a wealth of studies demonstrated that unfolded or Intrinsically Disordered Proteins and Regions (IDPs and IDRs respectively) represent above 40 percent of the proteome and their function or malfunction are associated with a variety of diseases, such as cancer, neurodegenerative disorders and viral infections. Unfolded proteins do not adopt a single well-defined three-dimensional structure, but rather an ensemble of rapidly inter-converting conformations defined by a relatively flat free-energy landscape. The conformational heterogeneity allows them to engage in a myriad of interactions, often without formation of secondary structure, thereby putting a spotlight on the role of dynamics in protein function. In the presented chapter, the systems under study, their biophysical properties and the influence of physical parameters, modify the operating conditions of proteins in cells (e.g. viscosity, temperature) that affect the dynamics and, consequently the function of unfolded proteins.

1.1 Unfolded proteins

1.1.1 From static to dynamic picture

Escaping classical structure-function paradigm, Intrinsically Disordered Proteins or, unfolded proteins, instead of exhibiting a single well-defined structure, cover a wide range of conformations, which in turn allows them to exhibit unique characteristics and engage in a myriad, new hitherto undiscovered functions. Given the dynamical diversity of the stochastically modulated by the local (amino acid residue) and global (overall shape and size) environment ensemble necessitates multi-faceted approach in characterization of their functional state (1–6). Studies of unfolded proteins cover a wide range of experimental measurements, starting from circular dichroism, through analytical ultracentrifugation to tryptophan-bleaching experiments, allowed to qualitatively probe the physocochemical properties of unfolded or denatured state. Further improvements in understanding of these biomolecules was supported by the evidence of binding between unfolded proteins with isothermal calorimetry, providing clues towards affinity of the formed complexes. Functional complexes formed by the proteins feature much weaker binding affinities than folded proteins with K_d on the order of μM puzzling biophysicists for many years. The possibility of exchanging buffers, adding binding agents, opened up avenues in terms of understanding of these highly mysterious class of biomacromolecules. Additional information provided small-angle scattering of X-ray and neutron (SAXS and SANS), methods allowing for estimation of the overall size and shape of the ensemble (7–9).

A more detailed view on the function of unfolded proteins has been supplemented only when the dynamic nature of unfolded proteins was taken into account. An array of studies following single molecule fluorescence approaches featuring specifically designed biopolymer construct with fluorescent dyes exhibiting Förster Resonant Energy Transfer (FRET) phenomenon, one can monitor potential self-association effects, binding between the biopolymer and gain initial insight on its overall size (fragment-to-fragment distance and radii of gyration) distribution. More advanced studies on unfolded proteins and their functions were allowed when super-resolution microscopes with single photon detection schemes allowed to obtain ns resolution on global reorientation dynamics as accessed by fluorescence correlation spectroscopy and two-photon correlation spectroscopy (10–20).

By far, the most thorough experimental investigation of unfolded proteins in solution offers Nuclear Magnetic Resonance spectroscopy with unique spatio-temporal (from Å to nm and ps to s) resolution. NMR spectroscopy allows to monitor coherent evolution of nuclear magnetic moments (which direction is represented by spins) and their concomitant interactions in the presence of external magnetic field, in turn allowing to view the structure and dynamics of proteins with single residue or chemical bond resolution. More specifically, spin relaxation of spin half-integer nuclei (i.e. ^{15}N R$_1$, ^{15}N R$_1\rho$ and ^{1}H-^{15}N heteronuclear nuclear Overhauser effect - nOe) offers access to the local dynamics (e.g. reorientation of the rank 2 spherical harmonics representing spatial part of spin-lattice interaction tensor coupled to the NH-backbone vector) occurring on ps-ns timescales, while the long-range dynamics may be probed by techniques such as paramagnetic relaxation enhancement. Chemical shifts supplement information on the time and ensemble averaged local electronic density surrounding the nuclei, which provide an overall idea on the conformational landscape (population-weighted average information on local molecular architecture of) macromolecule fragments averaged over timescales ranging from ps to ms. To gain further insight into longer timescale fluctuations one can apply relaxation dispersion methods, which are sensitive to modifications of nuclear magnetic moments evolution occurring on a range of timescales from μs to ms. NMR is, however, limited by size (or rather the tumbling rate) of protein (or its fragment) under evaluation, with small macromolecule fragments being far better resolved than large macromolecular complexes (with weight exceeding 50 kDa, typically possessing reorientation rates exceeding 20-30 ns). Additionally, NMR features a time-window (from tens of ns to tens to μs) for which estimation of detailed information on dynamics is essentially impossible due to either physical or instrumental limitations (21, 22).

On a par with NMR, the spatio-temporal resolution is very often supplemented from the point of view of molecular dynamics (MD) simulation, which relies on implementation of Newtonian physics to describe the time-dependent fluctuations of atomic positions of a biopolymer model. The information on forces, affinities and energetic levels of (macro)molecule fragments of interest may be estimated from the MD trajectories allowing to obtain a unique, atomistic view on the functional behavior of proteins in the medium of choice. To date there exist a myriad of approaches to effectively sample the potential energetic landscape of disordered proteins, which in turn reveal fundamentals on relationship between conformational heterogeneity and multi-timescale dynamics of the highly abundant class of proteins. Depending on the specificity of information sought from the simulation, the spatio-temporal resolution of the simulation ranges from Å to nm and from ps to μs. More specifically, the computational power and costs limit MD simulations to μs-long observation of highly detailed structures, with up to ms observation of coarse grained models. Still, current force fields suffer from inaccuracies due to the application of classical principles to describe are almost quantum in nature interactions. Concurrent developments in the area necessitate benchmarking of the methodology with the aid of experimental methods reaching atomistic resolution such as NMR spectroscopy. Combination of NMR and MD simulations, should therefore allow for detailed studies of biological function of these highly abundant, yet still enigmatic class of proteins (23–27).

1.1.2 Ensemble based descriptions of unfolded proteins

One of the features of unfolded proteins when observed by NMR spectroscopy is a relatively narrow spread of chemical shifts in the proton dimension when compared to the folded protein. This occurs when the stochastically modulated large scale polypeptide fluctuations occur faster than the chemical shift evolution (up to ms). In other words, exchange between members of the conformational ensemble featuring distinct local electronic density and nuclear environment occurs faster than the evolution of the average magnetic moments of nuclei under evaluation. As a result, all of the amino acid residues instead of yielding different chemical shifts (each for a single conformer) collapse into one single, well-resolved peak with chemical shift (CS) value equal to the population-weighted average of CS originating from the amino acid residue of each of the conformer. In order to obtain quantitative information from the chemical shift data on the local molecular architecture of the macromolecule fragment containing nuclei under evaluation, one may compare the experimental values with so-called random-coil CS, which values for different residues as a function of amino acid sequence are generally accessible. Deviations from the random-coil values presented in the form of Secondary Chemical Shifts allow for initial, coarse-grained structural assessment of the unfolded protein conformational ensemble (28, 29).

Valuable piece of information on local backbone sampling of unfolded proteins may be obtained from residue resolved Residual Dipolar Couplings, which inform on the long timescale (up to the ms) average perturbation to the nuclear magnetic moments evolution as a result of internuclear dipole-dipole interaction, which does not average to zero in the presence of alignment medium. Practically, estimation of RDC values relies on calculation of principal components of the alignment tensor and normalized spherical coordinates of the internuclear vector connecting interacting nuclei relative to the tensor for each of the conformer (30–36). An information on the long-range contacts may be additionally obtained from paramagnetic relaxation enhancement studies, which rely on incorporation of a tag containing paramagnetic nuclei (i.e. DOTA-Gd^{3+}) or stable radical (i.e. TEMPO) in the structure of the protein (37, 38). The interaction between the paramagnetic moiety and spin-1/2 nuclei of interest leads to a distance-dependent attenuation of the intensity in a manner depending on the ensemble-averaged inverse sixth power of the electron-nuclei distance ($\mathrm{I}_{para}/\mathrm{I}_{diam} \propto \langle r^{-6} \rangle$). The additional information on an overall shape and dimensions of the unfolded chain is accessible from X-ray and neutron scattering, as well as single-molecule FRET experiments (8, 9). To quantitatively assess the conformational space spanned by an unfolded protein of interest one typically resorts to sample and select approaches, first by building statistical coil ensemble and selecting the conformers satisfying the experimental data. For instance, one can build a library of all possible conformers starting from one of the termini with successive addition of peptide planes limited by steric repulsion of neighboring residue side-chains (*flexible-meccano*) on the basis of the conformational energy landscape of a statistical coil. In the following step, using a genetic algorithm (e.g. ASTEROIDS, A Selection Tool for Ensemble Representations Of Intrinsically Disordered States) and ensemble-averaged data, selection is being made to narrow down the pool of conformers from range of tens of thousands to hundreds of conformers that best represent the experimental data. In order to asses the quality of representation of experimentally accessible information on the conformational space of an unfolded protein, one can remove part of experimental data used for selection and estimate the experimental values representing population-weighted ensemble average not used in the selection (cross-validation). Obtained this way conformational ensemble may be used as a starting point for MD simulation studies of protein dynamics and their binding with other partner proteins allowing to estimate single-residue resolved binding kinetics - association and dissociation rates, shedding light on the overall binding equilibrium and protein function (39–42).

1.1.3 Complex binding equilibria of unfolded proteins

The intricacy of functional behavior of unfolded proteins in solution may be derived from the spin relaxation or relaxation dispersion studies. Binding events between proteins or their fragments slow down the overall reorientation of the protein on ps-ns timescales resulting in perturbation of spin states giving rise to the spin relaxation phenomena. The formed binding equilibrium, depending on the system of interest, modifies evolution of nuclear magnetic moments on average represented by bulk magnetization potentially leading to differences in chemical shift values of the bound and unbound forms. The most fundamental experiments used to derive information on dynamics from NMR spectroscopy measurements are experiments which monitor magnetization evolution of prepared spin state (coherence) of interest. The coherent evolution of nuclear magnetic moments represented by bulk magnetization is encapsulated within longitudinal (R_1), and transverse (R_2) relaxation rates informing on rates at which magnetization of given coherence comes back to equilibrium with bulk magnetization vector aligned along the lines of the external magnetic field (R_1) and loss of magnetization in the transverse plane (perpendicular to the external magnetic field, R_2) (43, 44). The overall line shape of peaks and the resultant spectral resolution inversely depends on the transverse relaxation rate - the larger R_2, the broader the peak. Binding and unbinding between protein fragments cause more efficient loss of magnetization in the transverse plane allowing to obtain first insight on the interaction dynamics between binding partners (45–49). A more detailed description of the binding equlibria is accessed via a suite of relaxation dispersion experiments, which allow to monitor the chemical exchange between the major (ground, G) and minor (excited, E) state of the resolved site. On timescales ranging from approximately 10 ms to 1 s (slower than the chemical shift evolution) one is expecting two peaks with distinct chemical shift values. On faster timescales (up to 100 ns - 1 μs) the observed chemical shift represents population-weighted average of the two states. If the exchange rate is comparable with the timescale of chemical shift evolution (around tens of μs to tens of ms), one may have difficulties in resolving peaks representing combination of the states (intermediate regime) (50–54).

One of the first successful attempts to probe the exchange effects on μs to ms timescales was Carr-Purcell-Meiboom-Gill (CPMG) experiment during which coherent nuclear magnetic moments evolution in the transverse plane is achieved using a series of refocusing pulses. Application of the pulses at varying frequency allows to quench the transverse relaxation contribution from exchange between major (G) and minor (E) state, while providing information on the interconversion rate between the states. CPMG is one of the simplest experiments that can be applied to study backbone or side-chain dynamics of folded, as well as, unfolded proteins (55–57). Concurrent development in NMR methodology led to a myriad of experiments based on application of additional radio-frequency field which limit the nuclear magnetic moments evolution of the nuclei of interest to oscillations under an angle θ allowing to study the coherent evolution of magnetization in so-called rotating frame - the magnetization is spin-locked between the longitudinal and transverse planes. As the evolution of magnetization in the rotating frame depends on the rate of exchange between the states, its chemical shifts and relaxation rates, as well as strength of the additional radio-frequency field and offset from the resonance, fine tuning of the spin-lock strength allows to map binding equilibria with interconversion rates ranging from tens to tens of thousands per second (58–62).

The application of additional radio-frequency field causes partial saturation of magnetization of the nuclei within the molecule fragment leading to loss of magnetization associated with its evolution in the rotating frame. Monitoring of the signal intensity attenuation associated with transfer of magnetization saturation occurring as a result of exchange between major and minor states allows to quantify the populations, rates of interconversion and inherent relaxation rates, giving an overall idea on the ground and excited state dynamics. More specifically, Chemical Exchange Saturation Transfer (CEST) finds an application whenever the chemical shift difference between excited and ground state ($\Delta\omega_{GE}$) is significant, the transverse relaxation rate difference between those states ($\Delta \mathrm{R}_{2,GE}$) is small and the population of the excited state is significant (e.g. in the case of weak binding events between similarly sized macromolecules) (49, 63). In the signal attenuation profile one then expects two dips, on and off-resonance originating from the ground and excited state respectively. Dark-state Exchange Saturation Transfer (DEST) on the other hand, is not limited to the significant chemical shift difference and relies on large differences between transverse relaxation rates (more than 3 orders of magnitude) of the bound and unbound states (e.g. in the case of binding of unfolded protein to a large macromolecule) (64, 65). For this type of exchange, the intensity profile features a single, but much broader dip than for the protein in the free form. Conveniently, all of the experiments may be described by the same Bloch-McConell formalism allowing to combine information from multiple sources in order to describe the complex binding equilibria featured by unfolded proteins (66, 67).

The kinetic exchange studies have allowed to observe transient effects associated with formation of weak, multivalent or transiently formed complexes often without formation of a stable secondary structure. One of the interesting examples is the binding equilibrium of C-terminal fragment of Sendai Virus Nucleoprotein (NTail) and its Phosphoprotein

(PX). In the free form, NTail undergoes rapid exchange (10^5 s^{-1}) between unfolded and transiently formed helical intermediates in the so-called Molecular Recognition Element (MoRE). Introduction of binding partner changes the folding-unfolding equilibrium allowing to sample the surface of PX forming so-called encounter complex (k_{ex} 850 s^{-1}). In the subsequent step, the complex is transiently stabilized forming weakly bound intermediate state, which is indispensable in the viral assembly process (68). The information on the binding equilibrium could not be provided without a suite of multinuclear relaxation dispersion experiments. As such, ^{1}H and ^{15}N are sensitive to the local electronic and nuclear environment (directly reflecting on the molecular architecture and its temporal modifications) and respond to the binding-unbinding events in much more pronounced way, while the ^{13}C' relaxation dispersion (more sensitive to changes in the local intra-chain hydrogen bonding pattern) informs more on conformational equilibrium between the transiently folded encounter and functional form of the complex. Similar free-state equilibrium between folded and unfolded states reveals mitogen-activated protein kinase kinase 7 (MKK7), which binding to c-Jun Kinase 2 stabilizes one of the conformations as revealed by additional X-ray crystallography studies (69).

On the other hand, one may find protein binding equilibria, which occur without formation of a stable folded architecture. An example of such phenomenon is supplemented by Delaforge et al., 2018 where N-regulatory domain of mitogen-activated protein kinase kinase 4 (MKK4) binds to its partner p38α to activate the kinase cascade. Application of combination of CPMG and CEST experiments allows to estimate the information on the binding equilibrium properties (k_{ex} around 270 s^{-1}) and dissect the contributions from the bound (excited) state of MKK4 to p38α, which within the framework of Stokes-Einstein description of isotropic tumbling corresponds to an astounding reorientation rate of approximately 47 ns, when compared to tumbling rate of MKK4 free form fragments oscillating within the range of 5-10 ns (70). Very similar binding equilibrium characteristics features disordered C-terminal region of Artemis and DNA binding domain of Ligase IV, which has been revealed from combination of NMR relaxation dispersion studies (CPMG and CEST) with ITC and molecular simulation docking (71). An analogous example is ultra-weak binding of importins to the FG-rich nucleoporins with dissociation constants on the order of single mM, which occurs with minor modifications to the conformational sampling of the protein during the course of fast transport through the nuclear pore membrane (72). More complex binding equilibria features Influenza A (PB2) polymerase machinery, as the temperature-dependent open-close form equilibrium requires large scale reorientation of the entire 627-NLS domain (NLS denotes nuclear localization sequence). The closed form, which structure has been originally solved on the basis of X-ray crystal diffraction patterns is essential in the viral replication machinery, while opening of the 627-NLS enables otherwise unattainable due to steric clashes binding of importins to the NLS allowing for nuclear import where the replication occurs. Application of temperature-dependent CEST, smFRET and SAXS allowed to map the equilibrium and concurrent interaction trajectory between 627-NLS and importinα (73, 74).

1.2 From structrure and dynamics to function

1.2.1 NMR-based description of protein dynamics

If the structure and topology represented by local and global protein backbone and side-chain sampling is not the only factor influencing functional state of protein, the spotlight in pursuit of an accurate description of conformational behavior of proteins lands on fast (ps to ns) and slow (μs to ms) dynamics. Investigations in this direction require more general understanding of restrictions imposed on peptide units influencing the stochastically modulated motions exhibited by polypeptide units in solution on multiple timescales. More than four decades of studies on biomolecular dynamics by NMR provided initial insights on rigidity and domain dynamics of folded proteins (36, 75, 76). The information on ps to ns dynamics of proteins estimated from NMR relaxation encapsulates the ensemble of spin interactions of the backbone or side-chain spin pairs arising due to stochastically modulated motions acting on reorienting peptide unit and entire macromolecule via dipolar interactions of both direct (between neighboring nuclear magnetic moments) and indirect (associated with fluctuations of electronic density surrounding the nuclei) origin, as wells as interaction between the quadrupole nuclear magnetic moment with electric field gradients, or cross-correlation between the dipolar interaction mechanisms. The ensemble of these processes may be described by the semi-classical second-order perturbation theory (Bloch-Wangness-Redfield) linking dipolar or quadrupolar interaction-associated modulation of nuclear magnetic moments to the coherent evolution of its average, represented by spin interactions and bulk magnetization respectively. The separation of spin and spatial part of spin-lattice interaction tensor allows to describe reorientation process of macromolecule in a form of correlation function of the latter represented by rank 2 spherical harmonics, encapsulating connection between stochastically modulated motions and allowed spin interactions, which following series of spatio-temporal averaging processes are represented by coherent evolution of bulk magnetization observed experimentally in the form of relaxation rates (44, 77). Within the classical treatment, reorientation of a macromolecule (i.e. folded protein) is described by rotational diffusion of the entire biological entity with internal degrees of freedom allowing for localized fluctuations of its fragments (bond vectors, peptide planes and larger polypeptide fragments). Both of the processes are described by an auto-correlation function, which represents cumulative fluctuations of spatial functions of spin-lattice interaction tensor coupled to the reorientation of the interacting spin pairs representing backbone or side-chain vector of interest (for description of internal degrees of freedom) and the entire macromolecule (for global reorientation rate) via successive application of transformations from the interaction frame, through the internal mobility tensor to the global reorientation tensor. The correlation function (C(t)) describes the orientation of the spatial part of spin-lattice interaction tensor at **t** equal to zero and its orientation after a time interval (at **t** equal to τ), leading to a maximum value of **1** at **t** equal to zero, with successive decay of the C(t) value towards *plateau* value associated with long timescale equilibrium fluctuations as stochastic motions lose memory of the initial state of the system (78).

In the standard Model-Free treatment of relaxation rates, one assumes statistic independence of processes originating from slow global reorientation and fast internal motions effectively contributing to the rotational diffusion of protein represented by an angular correlation function. In order to interpret NMR relaxation, one has to consider Fourier Transform of the auto-correlation function, i.e. SDF, which describes probability of observing contributions to relaxation from (dipolar or quadrupolar) interaction associated modulation to nuclear magnetic moments evolution arising represented by allowed spin transitions arising concurrently to the stochastically modulated motions at timescales proportional to an inverse frequency of the spin interactions. The linear combination of SDFs at different frequencies associated with the interactions constitutes the experimentally accessible relaxation rates (44, 77). Separation of processes contributing to the reorientational motion of a protein allows for more straightforward treatment of the datasets and estimation of reorientation rates (timescales of motion), as well as, so-called order parameter (S^2), which in the classical view of Model-Free analysis is proportional to the degree of rotational freedom of a bond vector (i.e. NH-backbone or CH_3-side-chain) attached to a body reorienting on much slower timescale (79). From the perspective of a correlation function, S^2 is equal to the *plateau* value of a correlation function component (i.e. internal motion) and represents the relative amount of memory consumed by a reorientation process (of the entire macromolecule or larger macromolecular fragment) leading to the observed spin transitions. The remaining part of dynamics represented by $1\text{-}S^2$ is associated with relative amount of information on reorientation process of peptide units including their internal and external degrees of freedom encapsulated within experimentally accessible spin relaxation rates. The greater the order parameter, the smaller relative amount of faster dynamics occurring on the relaxation active timescales. For backbone or side-chain vectors located within folded proteins of similar size (featuring global reorientation of similar rate), S^2 informs on the relative amount of degrees of freedom available to the peptide unit, i.e. local rigidity of the protein. The spin pairs situated within peptide plane in the folded regions of the proteins typically reach S^2 values ranging from 0.8 to 1, while smaller values are typically observed for unfolded parts of the proteins, i.e. flexible loops and the termini (80–89). The description of internal motions occurring on ps-ns timescales of folded protein have been approximated by so-called wobbling in a cone model in one of energetic potentials describing relative degrees of freedom available to the peptide unit (90–93). The changes in order parameters upon binding of both amide backbone and methyl side-chain have been linked to changes in conformational entropy and used in interpretation of conformational preferences for intra and inter-molecular interactions between folded and unfolded proteins with potential applications in drug target design (94–101).

A complementary point of view on the structure and dynamics of folded proteins, occurring on ps to ms timescales may be obtained from Scalar and Dipolar Couplings representing interactions between spin pairs arising either due to coupling between nuclear magnetic moments and electronic environment of the nuclei associated with recombination of the valence shell spin orbitals (through bond) or the direct dipolar interaction of spin pairs (through space). Both of the interactions reveal dependence on the conformational sampling of the polypeptide units. The incomplete averaging of the dipolar interactions introduced by perturbation to isotropic magnetic properties of medium in which the protein is immersed in via application of so-called alignment medium allows to quantify the Residual Dipolar Couplings (RDCs), which are estimated from difference between sum of scalar and dipolar couplings in the aligned and isotropic medium for different spin pairs. The Model-Free treatment of RDCs allows to estimate information on relative orientation of alignment tensor with respect to the molecular reference frame, leading to estimation of modulatory potential of reorientational motions limited by internal and external degrees of freedom of peptide planes averaged on timescales of up to the ms with timescale limit set by the chemical shift evolution (36, 102, 103). Estimation of alignment tensor components represented by its magnitude, anisotropy and its average orientation with respect to the external magnetic field provides information on the exact atomic positions in the peptide unit, allowing for accurate structure determination and estimation of relative degrees of freedom, represented by long timescale average rank 2 spherical harmonics (order parameter) (32, 34, 35, 85, 104–106). Comparison of order parameters derived from relaxation rates and RDCs allows to quantify polypeptide unit dynamics occurring on timescales limited by the global reorientation of the macromolecule (S^2, relaxation) with respect to the degrees of freedom sampled by the reorienting peptide unit(s) averaged over long-timescale window (S^2, RDC) respectively (107–112). A more quantitative view on the reorientation process of a macromolecule fragment may be achieved by approximation of polypeptide unit orientational modulation by Gaussian Axial Fluctuations allowing to quantify their relative degrees of freedom in three orthogonal dimensions (113, 114). Since, the measurement of RDCs necessitates estimation of scalar couplings of all interacting spin pairs, the information on the peptide plane geometry and its long-timescale average modulation may be supplemented by analysis of three-bond scalar couplings acquired in the same set of experiments. For instance, $J_{NC'}$ three-bond scalar couplings depend on relative orientation of vectors connecting the interacting nuclei and allow to estimate the (ensemble-weighted) average orientation of peptide units in the Ramachandran space represented by the ϕ-ψ angles encapsulated within Karplus relationship and modulations to geometry of peptide units associated with hydrogen bonding or transient interactions between polypeptide fragments (115–117).

Application of the NMR relaxation and RDCs to study of local protein dynamics and their modification due to the inter and intramolecular contacts between neighboring polypeptide fragments. The latter has allowed application of the methods in interpretation of modulatory potential of stochastically modulated motions of peptide planes and their binding induced modifications allowing to describe modulations to internal and external degrees of freedom of reorienting polypeptide units (98–101). For instance, binding of SH2 to pY1021 causes increase in methyl (axis) order parameters in apolar region of the protein (L35-I55) exhibiting greater order parameters in free form with respect to average value observed for other polypeptide fragments. Minor modifications to its dynamics in flexible region (featuring smaller order parameters) suggest limited modification to the ps-ns dynamics of methyl side-chains in the pTyr binding region. The limitations imposed on the degrees of freedom of side chains have been revealed from perturbation to the three-bond scalar couplings, which are subject to modulations as a function of distance and angle in-between the coupled nuclei, overall suggesting changes in the degree of mobility represented by variations in torsional angles (118). Similarly, limited modifications to NH-backbone order parameters, associated with binding have been observed for the SH2 PLCγ1 in-between complexed and uncomplexed with pY1021 protein. Minor modifications to the order parameters of residues in the apolar patch (residues 70-90, hydrophobic binding pocket), further away from phosphotyrosine binding pocket suggest increase in the degree of mobility of NH-backbone allowing for greater fluctuations on hundreds of ps timescales in small β-sheet (βE) and AA and EF loops, originating presumably due to large scale conformational rearrangement associated with formation of a complex with pY1021 peptide (119). In another example, estimation of structure and dynamics of Ubiquitin from experimentally accessible Scalar and Residual Dipolar Couplings, which structure and long-timescale dynamics has been validated against experimental data from tens of independent alignment media, coupled with reproduction of cross-correlated relaxation rates arising due to spin interactions between neighboring backbone amide nitrogen-proton pairs, as well as, with α carbon-proton pairs, sensitive to motions occurring on longer timescales reveal correlated motions within β-strand of reorienting protein. Perturbations to hydrogen bonding network of reorienting peptide units engaged in formation of β-strand suggest correlated motions of peptide units, which to similar extent have been observed in MD simulations. Variations in number of degrees of freedom available to peptide units within the β-strand arising due to intramolecular hydrogen bonding allow for greater collectivity of reorientation process of the polypeptide fragment (120, 121). Similar observations on correlated nature of motions in the folded and unfolded segments of peptide units have been revealed by independent, site-resolved analysis of RDCs originating from different backbone pairs of GB3 revealing restrictions imposed on reorienting peptide unit represented by order parameters decreasing in the order of NH, CαHα and NC'. The degree of correlation between the NH and CαHα order parameters suggests greater interdependence of dihedral angle ($\phi\psi$) fluctuations between subsequent peptide units positioned in-between α-helical element in the central part of the protein and surrounding it elements of the β-strand (107).

More accurate determination of the extent of peptide unit fluctuations is supplemented from the 3D GAF model to the analysis of RDCs of protein G, allowing to decompose the ns-ms oscillations of the peptide planes in three orthogonal directions: rotation around the long peptide plane axis defined by the CαCα vector (γ), in-plane (β) and out-of-plane oscillations (α). For folded elements, the γ fluctuations are dominant with standard deviation of underlying distribution on the order of 20-25°, while the latter two motions reveal smaller degree of fluctuations with oscillations about the average position remaining within 10°, and dropping to zero in the α-helical element. For flexible elements the magnitude of α oscillations is greater, reaching up to 20-25°, suggesting that the large scale fluctuations in the direction orthogonal to the rotation axis are allowed. The extent of fluctuations is associated with relative number of degrees of freedom for peptide plane reorientation along the orthogonal axis. The oscillatory pattern of the α

motions in flexible elements suggests greater number of degrees of freedom for residues 14, 41 and 44, than for the neighboring ones. The latter is presumably associated with positioning of residues within polypeptide loops with respect to distance from β-strand, suggesting that neighboring residues might posses greater restrictions on their mobility due to conformational strain associated with small amount of peptide units separating the unfolded fragment from other folded parts. The greater extent of γ fluctuations of residues in the β-strand suggests correlated motions of peptide planes influencing hydrogen bonding pattern of the polypeptide units. The comparison of $J_{NC'}$ scalar couplings for static structure and the one exhibiting correlated motions implies correlated motions between residues contributing to formation of the β-sheet. Comparison of order parameters estimated from experimental RDCs with the ones estimated from NMR relaxation suggests pronounced α and γ motions within and in-between the β-strands, which occur on timescales up to the ns and beyond (112, 122).

1.2.2 MD-simulation guided interpretation of unfolded protein dynamics

The complexity of unfolded protein dynamics precludes straightforward description of their heterogeneous conformational fluctuations solely on the basis of NMR relaxation or RDCs - one is typically challenged with interpretation of experimentally accessible values on the basis of global and local motions. More complete understanding of the motions featured by unfolded proteins necessitates development of more complex formalism. Drawing from the polymer nature of this highly abundant class of proteins, one can distinguish multiple dynamic modes occurring on different timescales - from very fast oscillatory movements of bonds leading to fluctuations of local degrees of freedom within peptide planes, through correlated motions of several peptide units (segmental motions) to large-scale reorientation of domains, as well as, rotational and translational motions of entire protein (123, 124). With an aid in interpretation of experimentally observed parameters come ensemble based Molecular Dynamics simulations, which allow to better visualize the complex processes occurring on multiple timescales. A series of examples is supplied by Propmers and Bruschweiler, who by an application of principal component analysis of covariance matrices of spin-lattice interaction tensors of ^{15}N-backbone nuclei estimated motional eigenmodes from MD trajectory (Reorientational Eigenmode Dynamics analysis, RED). The parameters estimated from the RED analysis represent mode collectivity and amplitude of fluctuations of the mode representing degree of correlation of reorientational motions and their relative contribution to the reorientation process which drive the spin relaxation observed experimentally respectively (125, 126). The associated information on timescales of motion, may be estimated following a fit to the experimental NMR relaxation rates, while diagonal elements of the covariance matrix represent order parameter S^2. Application of the RED analysis to Ubiquitin in native form and in A-state (partially unfolded) convincingly shows correlated motions featuring high mode collectivity of unfolded parts of the protein. Convincingly, in a follow-up study, by varying the simulated block length from 1 ns to 500 ns on cold shock protein (Csp) of *E. coli*, RED analysis shows convergence of calculated order parameters to the ones estimated from extended Model-Free analysis (88, 127).

The information encoded in the experimental data is both ensemble and time-averaged necessitating multiple averaging steps to fully represent the conformational fluctuations of protein fragments in solution from molecular simulations. In order to perform the MD simulations, one typically begins by designing a pool of conformations best representing ensemble averaged parameters averaged on timescales of up to the ms (e.g. chemical shifts and RDCs). Choice of starting conformers from the pool able to reproduce the experimental data should be non-redundant in terms of local conformations and topology of conformers, i.e. the chosen conformers should be separated on potential free energy landscape to as greatest extent as possible. In order to satisfy convergence criteria of an estimated parameter from which information on stochastically modulated motions exhibited by protein fragments may be estimated, length of the trajectories should exceed several correlation times of the longest reorientation rate of the polypeptide unit. The resultant information, however, especially in the case of disordered protein dynamics occurring on ps to ns timescales rarely matches the experimentally observed values prior to reweighing or rescaling of the resultant trajectories (128–130). The latter is associated with parametrization of force fields involving contributions to energetic potential originating from bonded and non-bonded terms, i.e. van der Waals and electrostatic interactions (131–134). One is typically challenged with tuning of the force-field parameters or water model with an aim to better represent the experimental data and the underlying protein dynamics (135–137). An advancement in interpretation of MD simulations of unfolded proteins is based on the concept of Average Block Selection Using Relaxation Data (ABSURD). In principle, improvement in reproduction of the experimentally accessible relaxation rates on three decades of timescales is achieved by selection of blocks of trajectories on the basis of comparison of relaxation rates representing dynamics derived from the auto-correlation functions with the experimental relaxation rates encapsulating information on motional process on one of the decades of timescales (138).

Further efforts in description of the complex disordered protein backbone dynamics from MD simulations has allowed to decompose the angular correlation function into the product of three correlation functions informing on modulation of internal and external degrees of freedom of peptide planes, represented by local librational motions, oscillations within and in-between the Ramachandran basins, as well as, reorientation of entire segments composed of multiple peptide units. The order parameter describing averaging process of rank 2 spherical harmonics due to the fastest processes is almost sequence independent and oscillates around value of 0.9 manifesting almost negligible contribution of the processes to memory loss of NH-backbone vector represented by a correlation function of spatial part of spin-lattice tensor. The fast intermediate motions occurring on hundreds of ps exhibit little correlation with regard to molecular architecture of side-chains belonging to the reorienting peptide units, which suggests that the motions are dominated by fluctuations of degrees of freedom of peptide planes not exceeding the Ramachandran basins. The intermediate and slower motions with values oscillating around single and several ns reveal marked sequence dependence with significantly higher values for more bulky and highly polar residues possessing greater limitations to the external degrees of freedom, presumably due to steric hindrance with neighboring side-chains, electrostatic interactions and hydrogen bonding between polar groups of neighboring residues. Further, significantly greater values are observed for residues involved in fast exchange between folded and unfolded states, potentially due to greater moments of inertia of reorienting (folded) polypeptide fragment. The slowest of the NH-backbone reorientational motions reveal timescales within the same range as $C\alpha C\alpha$ tumbling representing correlated motions of several residues in a row (segments). Observations of correlated dynamics of consecutive peptide planes as measured by standard deviation of angles formed by subsequent peptide planes ($\sigma(C_{\alpha,i}$-$C_{\alpha,i+1}))$ spanning three (unfolded part) to seven residues (α-helical MoRE) provide evidence towards correlated dynamics and support polymer-like nature of unfolded proteins (139). The observed by MD dynamics has been validated against an ensemble of experimental ^{15}N relaxation rates. A follow-up study shows that the quality of reproduction of the experimental data strongly depends on water model and applied force-field in the simulation(s) (140). Subsequent analysis of the segmental motions allows to separate contributions from reorientation of entire segments with respect to laboratory frame (inter-segmental dynamics) and those originating from modulation of internal degrees of freedom of the segment. Additional temperature-dependent analysis reveals that intra but not inter-segmental chain dynamics is directly affected by lifetime of hydrogen bonds between polar atoms of side-chains of neighboring peptide units and surrounding water molecules. This type of tight coupling between the solvent and the protein dynamics is not unexpected, as slow-down of water molecules in vicinity of the protein surface, clefts and in binding pockets has been observed in a number of studies (141–145).

Comparison of MD-derived NH-backbone and methyl side chain order parameters of folded proteins with up to μs long trajectories with experimentally accessible RDCs or NMR relaxation rates reveals long timescale processes contributing to the reorientation of peptide fragment of interest. The μs long MD simulations of Ubiquitin reveal $\phi\psi$ anti-correlated dynamics in β1-β2 subdomains (residues 12-16), and correlated motions between β3 and β5 subunits (146, 147). The correlated nature of these motions readily observed in 20 ns simulations substantiates point of view of highly correlated motions featured by polypeptide loops connecting β subunits. The RED analysis confirms correlated reorientation of peptide planes in the fragment of the protein with the highest mode collectivity for residues 10 and 11. The obtained on the basis of joint RED and relaxation rate analysis timescales of motions for the estimated eigenmodes oscillate within hundreds of ps for the modes of highest collectivity. The degree of correlation between motions exhibited by the residues is concurrent with greater number of degrees of freedom available for the residues with $\phi\psi$ mobility within or between the Ramachandran basins reaching up to 25° (126). The RED analysis allows to obtain insight into sparse reorientation events of high collectivity, which in ensemble based description of MD simulation or NMR relaxation disappear in the average. For instance, correlated motions exhibited by the entire β4-β5 loop (encompassing more than 20 residues) or residues in the α-helical fragment feature amplitude, which is an order of magnitude lower than the one representing correlated reorientational motion of β1-β2 loop (representing local $\phi\psi$ backbone flips). Of specific interest are correlated transitions of several peptide planes in a row, which degree of collectivity and amplitude varies across the sequence which may inform on long-range correlations of peptide units fluctuations, i.e. concerted dihedral hops (148). Similarly, μs-long MD simulations of Ubiquitin allow to estimate amount of time required for averaging of methyl side-chains degrees of freedom encapsulated within the experimentally accessed, RDC-derived order parameters. The improvement in reproduction of the RDC-derived order parameter is observed when the timescale window of MD simulation from which the experimental variable or analytical parameter is estimated is longer than tens of ns (149, 150). The long timescale dynamics presumably involves exploration of additional degrees of freedom allowed by reorientation of surrounding peptide units and their side-chains, which otherwise remain in transient intramolecularly bound states limiting the range of motion accessible for the methyl side-chains (151, 152). Both in NMR and MD simulation, the access to long-range dynamics of unfolded proteins is limited either by the lifetime of a correlation function (tens of ns) or accuracy of the force-field to accurately extrapolate the local ns dynamics to timescales longer than the reorientation rate of polypeptide units (represented by correlation time of spatial part of the spin-lattice tensor) without sufficient experimental validation. Additionally, the greater spatial extent of motion, the longer expected time for a biomolecule to execute it. Consequently, the duration of the MD simulation and concurrent computation cost become a hurdle. By sacrificing the spatial resolution and coarse-graining the biopolymer model, one may extend the simulation time allowing to visualize complex processes occurring on tens of ns to single μs timescales (153, 154).

1.2.3 Long-range dynamics

Experimental access to long-range dynamics from NMR may be obtained via experiments based on dipolar interaction either between (paramagnetic or diamagnetic) nuclear spins or (diamagnetic) nuclear spin-electron spin pairs separated by a distance of up to 10Å. In both cases, one is able to measure time and distance dependent fluctuations of reorientation vector connecting the two interacting moieties. In the first case, known from static structure determination of folded proteins the nuclear Overhauser effect between protons is utilized, which together with ^{15}N spectral editing allows to quantify the time-dependent signals originating solely from backbone amide protons. The information encoded in the nOe intensity build-ups is both time, distance and conformation averaged, posing a challenge to accurately estimate the dynamic information from cross-peak time evolution. Additional difficulties are associated with experimental design, as the nOes measured by ^{15}N edited HMQCs or HSQCs might contain contributions from spin diffusion, which necessitate estimation of relaxation matrix covering all of the spins in the vicinity of nuclei of interest. eNOEs have been used in qualitative assessment of loop dynamics of folded proteins such as GB3 and Ubiquitin (155, 156). A quantitative framework facilitating analysis of time and distance-averaged dipolar interactions observed from nOe build-ups has been summarized recently in a so-called Kinetic Ensemble approach which involves estimation of transition rate matrix describing probability of transition between states in the ensemble with different distance between the interacting nuclei and distance dependent strength of the dipolar interaction between them. Diagonalization of probability matrix allows to obtain eigenvalues corresponding to rates of decay of equivalent correlation function, which can be in turn used to evaluate SDF for estimation of NOE. For exactly three dynamic processes with fixed distance the approach simplifies to the extended Model-Free formalism (157).

Alternative way of estimating long-range dynamics may be achieved from studies of bio-macromolecules equipped with paramagnetic moiety equipped with an unpaired electron (e.g. located on nitroxide radical). In that case, one focuses on observation of enhancement of relaxation rates originating from nuclear-electron spin interactions. The paramagnetic relaxation enhancement which is dominated by the hyperfine interaction of nuclei of interest with an unpaired electron may be observed as an additional contribution to the relaxation rates. The PRE is usually estimated from peak intensity ratio or relaxation rates of spin half-integer nuclei of macromolecules to which spin label is attached in its paramagnetic and diamagnetic form, which for nitroxide radical correspond to oxidized and reduced spin-label. As the interaction between the unpaired electron and the nuclei of interest depends on average inverse sixth power of distance, PREs have been used as a measure of long-range distance restraints in structure calculation of folded proteins (38). Further improvements in understanding of the PRE have allowed application of the methodology to incorporate long-range effects for better estimation of structure and dynamics of unfolded proteins represented by RDCs (37, 158, 159). Following Solomon-Blombergen theory the paramagnetic relaxation contribution to the relaxation rates is directly proportional to SDFs characterizing probability of probing spin interactions, which occur concomitantly with the motion. Presence of the label, however, modifies the moments of inertia of the polypeptide fragments, which leads to deceleration of the local dynamics of unfolded chain, while the PRE additionally broadens the observed resonances leading to lower intensities and potential peak overlap, which directly influence the degree of accuracy with which experimental parameter is estimated i.e. relaxation rates or kinetic parameters from exchange transfer experiments. Additionally, the PRE effect is remarkably sensitive to the diamagnetic impurities of the sample - even 3 to 5% of diamagnetic species may lead to underestimation of Γ value by up to 30% calculated from peak intensity ratio (160). Measurement of Γ PRE should allow to obtain insight on time and ensemble-averaged long-range dynamics in folded and unfolded proteins within the range of 10 to 35Å for paramagnetic moieties featuring spin 1/2 and spin 7/2 respectively. In contrast to Γ_2 typically used for structure refinement and evaluation of intramolecular contacts, the PRE contribution to longitudinal relaxation depends on internal dynamics occurring on timescales comparable with inverse frequency of the Larmor frequency of the nuclei which relaxation is evaluated, and, consequently requires more parameters for its successful estimation. Therefore, in practical application the use of PRE, as an additional contributor to transverse relaxation dominates due to its relative independence on strength of external magnetic field. The incorporation of PRE restraints in calculation of SRY/DNA complex allowed to minimize number of conformers representing experimental data and more precisely define contours of conformational space (38). Inclusion of a paramagnetic label to phosphocarrier HPr protein allowed to visualize complex with N-terminal domain of enzyme I (EIN) with bound state population of 10% remaining in fast exchange solely on the basis of intermolecular PREs (161). PREs have also taken its part in elucidation of intermolecular translocation mechanism of HOXD9 homeodomain protein on surface of DNA, which is functionally important in target search process during transcription (162). Observation of ^{13}C PREs in addition to ^{1}H PREs additionally allows for more accurate estimation of the very fast kinetics of exchange between Ubiquitin and unilamellar vesicles. The extracted ratio between the ^{13}C and ^{1}H Γ R_2, which is independent on apparent association rate allowed for more accurate estimation of the k_{off} (161). PREs have also assisted in description of oligomerization process of DNAJB6b chaperoned by Hsp70 (163). Additionally, immersion of stabilized radicals in protein solution allowed for estimation of solvent accessibility of folded and unfolded proteins (164, 165).

1.3 Towards biologically relevant conditions

As the complexity of macromolecular environment increases, description of protein dynamics becomes more challenging. In order to answer, how does the complex macromolecular environment affect the size, shape and dynamics of proteins, and more specifically disordered proteins, requires understanding of behavior of a protein under crowded, native-like conditions. In this section, I begin my considerations with simplified derivation of Stokes-Einstein equation, which is followed by more specific solution for macromolecule fragments across different length-scales. Further, general thermodynamic arguments allowing to describe crowded environment and phase separating systems are discussed.

1.3.1 Stokes-Einstein equation

Consider container filled with solid spheres of hydrodynamic radii R_h, which are constantly bombarded by infinitesimally small (Brownian) particles causing random displacement of spheres in a medium of interest. One should consider the time and surface-averaged pressure coming from the Brownian particles (p) causing random displacement (x) of a particle of interest. On the surface of those spheres a balance of forces is formed - one coming from temperature driven random motion, and opposing it hydrodynamic viscous drag (F_h).

$$\frac{dp}{dx} = -F_h\nu = -m\beta v\nu \tag{1.1}$$

where ν denotes average number of bombarding particles in a layer of interest ΔV, and F_h is directly related to mass of a particle and a friction coefficient, which depends on physocochemical properties of medium-particle interface (β) and reorientation rate (tranlsational and rotational velocity) of the particle. In a simplified view, for ideal gases, pressure on the surface of a particle is directly related to the number of Brownian particles (n) moving through the surface layer of volume ΔV.

$$pV = nRT = \frac{nN_0}{\Delta V}\frac{R}{N_0}T = \nu kT \tag{1.2}$$

$$\frac{dp}{dx} = \frac{d\nu}{dx}kT \tag{1.3}$$

Following Fick's law, the rate of change of particles within the layer dx is directly proportional to average velocity of particles and its number per unit volume:

$$-D\frac{d\nu}{dx} = v\nu \tag{1.4}$$

Substituting 1.5 to 1.1:

$$D\frac{d\nu}{dx}m\beta = \frac{dp}{dx} \tag{1.5}$$

From the balance of pressure and frictional forces on the surface of the spherical particle that is much bigger in comparison to the Brownian particles:

$$D\frac{d\nu}{dx}m\beta = \frac{d\nu}{dx}kT \tag{1.6}$$

which gives the Einstein's diffusion equation:

$$D = \frac{kT}{m\beta} \tag{1.7}$$

For idealized spheres of hydrodynamic radii R$_h$ immersed in liquid of viscosity η, one may apply the Stoke's relation:

$$D = \frac{kT}{6\pi\eta R_h} \tag{1.8}$$

Both equations (1.7 and 1.8) relate the diffusion coefficient with physicochemical properties of the medium and the particle of interest. The first one is more general and applicable to many systems, the latter one (Stokes-Einstein relation) is frequently used in description of macromolecules immersed in aqueous medium (166–168).

Spheres of hydrodynamic volume (V$_h$) immersed in a fixed volume interact between each other leading to transient slow-down of their rotational and translational movement.

As their concentration (c) increases so does the viscosity of the medium arising from these transient contacts depending on the size, surface to volume ratio of the interacting spheres, hydration layer of the spheres, and properties of interface between the sphere and the medium, which directly affect association and dissociation constants between the particles. One of the consequences of macromolecular crowding is overall slow-down of reorientation rate of water molecules. The interaction of solvent molecules with polymer units of crowder will lead to retardation of water molecules in its vicinity following formation of so-called hydration layer. Water molecules surrounding the macromolecule will constantly exchange between hydration layer and bulk portion of solvent leading to overall slow down of their reorientation rate. Macroscopically, those effects are visible as the increase in the viscosity, on the microscopic or rather nanoscopic point of view the description is slightly more complicated, as it depends on the nature of interaction between solvent and macromolecule of interest, as well as size and shape of the latter (169–171).

Translating equation 1.8 to more complex macromolecular systems featuring fractal dimensions, filling hydrodynamic volume only partially requires certain approximations. Efforts in that vein have been undertaken by number of researchers who observed deviations from generally valid Stokes-Einstein equation, as mobility of nano-probes depends not only on concentration of viscogen, but also on the probe size. To further describe response of probes to friction associated with frequent macromolecular collisions in crowded environment, Cai proposed classification of length-scale dependent motion of nanoscopic objects depending on relation between diameter of the probe (d) and characteristic length-scale in the polymer network, e.g. distance between a monomer of one chain and the nearest monomer of another chain ($\xi = R_G(c/c^*)^{-\beta}$), where c^* denotes concentration at which polymeric solution is in concentration overlap between dilute and semi-dilute conditions and $\beta = \nu/(3\nu - 1)$. In the first regime, when dimensions of a probe are smaller than dimensions of polymeric mesh size ($d < \xi$), diffusion of a probe is unaffected by the polymer mesh. When $\xi < d < L$, the long-time self-diffusion is affected, the probe motion experiences effective viscosity which scales with square of d/ξ. When $d > L$ motion of the probe is influenced by the polymer chains and subsequently experiences macroscopic viscosity of the solution (172). Further improvements to the description of Cai, following observations of deviations from Stokes-Einstein equation with regard to size of the probe experiencing motion, have been proposed by Holyst et al., who implemented an effective radius of the probe (R_{eff}) by taking into account the hydrodynamic or gyration radii of viscogen into description of length-scale dependent viscosity. In that case, the effective viscosity changes, such that:

$$\frac{D_0}{D} = \frac{\eta}{\eta_0} = f(r_p) = \exp\left(b\left(\frac{R_{eff}}{\xi}\right)^a\right) \text{ with } R_{eff} = \sqrt{\left(\frac{r_p^2 R_h^2}{r_p^2 + R_h^2}\right)} \tag{1.9}$$

In that case, *the size of the probe sets the scale of the flow, and the viscosity experienced by a particle in motion is not a property of the object but rather a property of the flow inside the fluid and the length-scale of that flow.* The R_{eff}/ξ ratio describes the extent to which the hydrodynamic flow is screened at distance ξ. In the model, value of stretching exponent 'a' is related to hydrodynamic properties of an object and scales approximately as R_h/R_G, which oscillates between 0.7 for extended polymers in entangled regime (where macromolecular chains interpenetrate each other) to 1.3 for solution of hard spheres. Importantly, the model reproduces the viscosity dependence on crowder concentration in a wide range of experimental conditions with size of the probe ranging from 0.5 to 6.9 nm, molecular weight of viscogen from 3 to 1000 kg/mol, and temperature ranging from 278 to 315 K, as observed from joint rheometry and Fluorescence Correlation Spectroscopy measurements. The description uniformly describes relationship between microscopic and nanoscale viscosity for probes and viscogens of different dimensions, and opens up an avenue in description of disordered protein dynamics in the crowded, native-like conditions (173, 174).

1.3.2 Thermodynamics of crowding

The particles filling the container do not perfectly stick to each other - due to steric repulsion arising occurring as a result of both hard and soft interactions, additional portion of volume surrounding particles is created for each of them giving rise to the excluded volume effect - the overall volume of particles will be larger than the one predicted from simple summation of the individual van der Waals radii of atoms composing the particle. Referring concentration of particles, or, more precisely macromolecules in the crowded environment to available volume rather than total volume puts an emphasis on surface-dependent effects such as solvation and potential specific or non-specific interactions with other solutes, stemming from van der Waals, electrostatic and hydrogen bonding interactions. Hence, those interactions depend not only on the size, shape and topology of macromolecules but also on charge and polarity of macromolecular surface. Folded proteins may be easily approximated by spheres, or ellipses, for which descriptions based on geometrical considerations, such as scaled particle and lattice theory typically well describe the excluded volume effect (169–171, 175, 176). Both of the theories consider thermodynamic effects associated with introduction of a convex particle in a bath of particles, the first one focusing on the size of introduced particle, the latter on the shape of it.

The description complicates a bit for unfolded proteins, which do not fill its hydrodynamic volume perfectly. In that case, apart from surface to volume ratio one needs to consider effects associated with reconfiguration of the macromolecular chain and potential compaction. To illustrate it, one should consider two types, one involving local rearrangement of macromolecule influencing intramolecular interactions and associated with it changes in the number of degrees of freedom and global effects associated with excluded volume and inter-particle interactions. In that case, the entropic effects associated with crowding-induced compaction may be counterbalancing each other - in attempts to maximize the number of degrees of freedom, the system will tend to increase the available volume by decreasing overall size of the macromolecule, while maintaining a large number of possibilities for the chain fragments to move independently. Additionally, enthalpically-driven specific or non-specific interactions may lead to transient intra and intermolecular interactions further promoting chain compaction or potential complex formation. In that case, the latter will strongly depend on the amino acid sequence, ionic strength, pH, temperature and presence of cosolvents. For a number of proteins, crowding-dependent compaction has been observed in single molecule fluorescence experiments in a manner dependent on molecular weight and physicochemical nature of the crowder determined mainly by the quality of its interaction with solvent. The larger the viscogen, the greater compaction effects of the unfolded protein of interest observed, with reduction in R_G on the level of up to 30% for the largest viscogen used. Conveniently, the experimental data obtained for a variety of conditions is explained by renormalized Flory-Huggins theory, which relates changes in macromolecular volume fraction associated with crowding to free-energy of mixing of the macromolecule and the solvent. Additionally to excluded volume effects, depletion interactions associated with creation of impenetrable for other particles layer around the solvated macromolecule of interest may be observed in crowded environments. These effects impact association rates of unfolded proteins as has been observed experimentally for a range of crowding conditions (177–181).

1.3.3 Phase separation of proteins

A special case of crowding are concentrated solutions of proteins, which apart from solvent-coupled effects directly affect each other by virtue of frequent specific or non-specific intra and intermolecular interactions. Transitioning from diluted to more concentrated solutions may occur either gradually, by obtaining more and more concentrated solutions of protein, or spontaneously via liquid-liquid phase separation, whereby next to the phase of low protein concentration, a separate liquid phase containing high volume fraction of macromolecules is spontaneously formed. The formation of separate phase proceeds occurs via collapse of small spherical discontinuities (droplets, containing considerable concentration of macromolecules in comparison to the one in surrounding them dilute phase) within the medium of interest. Droplets are thermodynamically unstable states and subject to either dissolution or growth. As recruitment of material depends on capture radius of the droplet, while dissolution depends on its surface-to-volume ratio, the newly formed states will undergo Ostwald's ripening process - smaller droplets will dissolve at the cost of larger phase-separated macromolecular moieties leading to droplets growth and eventually formation of separate continuous phase.

The phenomenon of phase separation is typically dependent on temperature, ionic strength, concentration of macromolecules in solution, as well as, the character of interactions between macromolecules and the solvent. The most basic description of LLPS may be represented with an aid of Flory-Huggins theory, which allows to approximate thermodynamic demixing associated effects using combination of interaction parameter χ and volume fraction (ϕ) representing enthalpic and entropic contributions to free energy of demixing respectively, originating mainly from interaction potential between the macromolecule fragments and excluded volume terms (182):

$$\frac{\Delta F}{Nk_BT} = \chi\phi_1\phi_2 + \frac{\phi_1}{x_1}ln(\phi_1) + \frac{\phi_2}{x_2}ln(\phi_2) \tag{1.10}$$

where x_1 and x_2 are coefficients which depend on the evaluated system: for either a pair of solvents (both x_1 and x_2 are equal to 1) solvent-solute (x_1 is 1, and x_2 is much larger) and solute-solute (both x_n are large) (183).

The origin of concentrated phases is strictly dependent on reaching a local saturation concentration of material (c_{sat}), which may be achieved either via self-interactions of a protein or a set of proteins of interest, in which case the process is mainly enthalpically driven, due to (self-)crowding with dominant non-specific interactions (corresponding to entropic terms), or a combination of both. Observations of phase separation for proteins rich in highly charged and aromatic residues puts the spotlight on the first of the processes with cation-anion (Arg-Glu), cation-π (Arg-Tyr or Arg-Phe) and π-π interactions. Although, weak and of transient nature, a number of the interactions may locally increase the concentration of macromolecules leading to formation of phase separating (droplet) nuclei. Further, depending on concentration of macromolecules in continuous phase and the surface-to-volume ratio of the droplet, the spontaneously formed droplet may recruit additional macromolecules from surrounding medium fostering its growth or spontaneously dissolve. The most basic description of LLPS is usually encapsulated in a phase diagram, which describes the thermodynamical state of the solution as a function of volume fraction and temperature, pH or salt concentration, within which one can distinguish a number of points characterized by spinodal and binodal decomposition lines. The region enclosed by binodal defines physicochemical conditions within which two stable phases are formed, while region between spinodal and binodal defines metastable region where a thermodynamical balance between formation and dissolution of droplets is formed. Critical point on the phase diagram denotes the lowest (or highest) temperature (salt concentration or pH) for which the LLPS may occur (184, 185).

Similarly to the crowded environment, the functional characteristics of proteins within the demixed phase are very likely perturbed by high concentration of solutes. In the case of unfolded proteins, instead of concentration-dependent compaction, one should expect interpenetration of the macromolecular chains maximizing the number of inter-molecular contacts, which maintain integrity of the droplet, while allowing for a maximal number of degrees of freedom. With concentrations reaching astounding numbers of hundreds of g/L within the condensate, one can expect that the majority of water molecules constituting the droplet will belong to the hydration layer and viscosity within it will be significantly greater in comparison to the one in the dilute phase, which in turn, should have visible effects on the dynamics of the protein within the demixed phase (186, 187).

1.4 Hypothesis

Following considerations of researchers across the multi-disciplinary subject of disordered protein dynamics in the native-like conditions, in the Thesis, I attempt to describe the limitations of methods applied for description of stochastically modulated motions from NMR spin relaxation and present potential means of representation of limitations on the reorientation process of macromolecular fragments associated with internal and external degrees of freedom featured by the polypeptide unit due to the intra and intermolecular interactions. More specifically, I attempt to describe the limits of experimental sensitivity associated with estimation of dynamic modes from Model-Free analysis of NMR relaxation rates as they relate to the underlying distributions of correlation times describing reorientation of polypeptide fragments. I attempt to quantify the influence of sensitivity of NMR relaxation rates to timescales of motions on values of estimated dynamic modes and parameters associated with their scaling across a range of conditions (temperature and viscosity) influencing bulk and local properties of the medium. I present a systematic study on the effect of crowded native environment on the unfolded protein dynamics represented by ^{15}N relaxation rates, and using simplistic terms quantitatively describe the manner in which timescales of motions occurring on ps to ns timescales are affected by the crowded native-like environment. I present an efficient mean of description of influence of viscosity and temperature both in the dilute, semi-dilute and concentrated phases on the local ps to ns unfolded proteins dynamics represented by apparent dynamic modes estimated from Model-Free analysis of NMR relaxation rates. I evaluate the response to temperature and viscosity-induced modulations of apparent dynamic modes observed for unfolded proteins featuring distinct dynamic fingerprints. I discuss the role of water dynamics and neighboring side-chain interactions on effective deceleration of unfolded polypeptide units. I attempt to qualitatively evaluate changes in distributions of correlation times underlying reorientation of macromolecular fragments of varying length occurring as a result of modulation to temperature and viscosity of the medium, and describe the quality of Model-Free representation of unfolded chain dynamics from NMR relaxation rates in the form of apparent dynamic modes of polypeptide fragment reorientation processes.

CHAPTER 2

Theoretical principles of NMR relaxation

Nuclear magnetic resonance (NMR) spectroscopy describes the complexity of interactions of nuclei with external magnetic field. Exposure of a nuclei to electromagnetic field creates electrodynamic force leading to oscillations of fundamental particles and formation of fluctuating magnetic fields within the nuclei leading to formation of partial magnetic moments, which further interaction with electromagnetic field causes their redistribution, formation of spin (collinear with the average magnetic moment of a nuclei, $\hat{\mu}$) and its precession around the axis of the external magnetic field with Larmor frequency. The oscillations of magnetic field within the nuclei and electronic density associated with reorientation of a molecule within the external magnetic field cause its perturbation leading to changes in the nuclear magnetic momentum represented by angular momentum operators. The relaxation active spin transitions representing redistribution of partial nuclear magnetic moments occurring on pico to nanosecond timescales as a result of inter and intranuclear dipolar interactions are represented by spin or transition operators. The spin operators are represented by the frequency and coherence order associated with number of magnetic quanta involved in the overall change in the energetic states of the spin system. The distribution of nuclear magnetic moments may be represented by non-reducible bulk magnetization, which evolution may be efficiently probed experimentally by application of series of radio-frequency pulses to influence the processes associated with its evolution on μs to ms timescales. The connection between the stochastically modulated motions and perturbations to spin system evolution encapsulating fluctuations of nuclear magnetic moments may be described using the semi-classical Bloch-Wangness-Redfield theory.

2.1 From spin to spin density operator

The redistribution of partial nuclear magnetic moments arising due to interaction of electromagnetic fields associated with oscillations of fundamental particles within nuclei with external magnetic field gives rise to its initial angular acceleration leading to an oscillatory-rotatory movement, which may be represented by precession. The electrodynamic force generated within the nuclei leading to its reorientation and formation of average nuclear magnetic moment ($\hat{\mu}$) depends on its composition - number and type of fundamental particles. The magnitude of coupling between nuclear magnetic moment and external magnetic field is encapsulated within the gyromagnetic ratio of a nuclei, γ, an intrinsic property unique for each nuclei, defined by a ratio between the frequency of its oscillatory-rotatory movement (Larmor frequency) and magnitude of external magnetic field. As the exact description of energy levels arising as a result of the interaction of the nuclear magnetic moments with the external magnetic field is difficult, one typically describes the spin system using discrete distribution of partial magnetic moments represented by linear combination of wave functions describing intranuclear electromagnetic flux with quantized energy levels. Nuclear spin, defining direction of average nuclear magnetic moment and spin operators describing distribution of partial magnetic moments within the nuclei are represented by quantum numbers: spin quantum number (l) and magnetic quantum number (m), which dependent on the composition of the nuclei take values from 0 to 8 in steps of 1/2 (l), and from -l to l (m) respectively. The ensemble of processes describing stochastic modulation of average nuclear magnetic moments within nuclei ensemble represented by bulk nuclear magnetization vector on one hand, and spin transitions on the other is encapsulated within the semi-classical Bloch-Wangness-Redfield NMR relaxation theory.

The stochastically modulated motions of molecule fragments cause changes in relative orientation of the nuclear magnetic moments with respect to the external magnetic field leading to redistribution of nuclear magnetic moments and alignment of portion of spins in the available spin bath along the lines defining direction of external magnetic field, and small portion against the direction of external magnetic field in a process called Zeeman interaction.

$$\frac{N_\beta}{N_\alpha} = exp\left(\frac{\Delta E}{k_B T}\right) = exp\left(\frac{\hbar\omega_0}{k_B T}\right) = exp\left(\frac{\hbar\gamma B_0}{k_B T}\right) \tag{2.1}$$

The time-dependent modulation to intranuclear partial magnetic moments distribution associated with inter and intranuclear interactions may be represented by the spin tensor operators (I), while description of their evolution with respect to energy levels associated with spin transitions is encapsulated within so-called Hamiltonians. Zeeman interaction is the strongest interaction of nuclei with external magnetic field (B_0) and is described by Hamiltonian:

$$\hbar\mathscr{H}_0 = -\hat{\mu}_s\mathbf{B_0} = -\gamma\hbar\hat{\mathbf{I}}\mathbf{B}_0 \tag{2.2}$$

From the perspective of range of timescales on which NMR relaxation occurs the major part of interaction between nuclear magnetic moments and magnetic field may be represented by time-dependent contribution to the time-independent Zeeman Hamiltonian.

$$\mathscr{H}(t) = \mathscr{H}_0 + \mathscr{H}_1(t) \tag{2.3}$$

$\mathscr{H}_0$ represents the Zeeman interaction energy operator, while $\mathscr{H}_1$(t) describes stochastic modulation of nuclear magnetic moments evolution associated with one of the mechanisms of interaction between nuclear magnetic moments in external magnetic field, its evolution due to the electronic density fluctuations for spin half-integer nuclei, or interaction between nuclear quadrupolar moment and electric field gradients for nuclei featuring integer spins. The information on ensemble of available spin states and their distribution within the nuclei ensemble contributing to the average nuclear magnetic moment is encapsulated within the so-called density matrix:

$$\sigma = \int P(\psi) \left|\psi\right\rangle \left\langle\psi\right| d\tau \tag{2.4}$$

with $|\psi\rangle$ and $\langle\psi|$ representing eigenvectors of spin system and P(ψ) representing probability density of a given eigenstate. Egenfunctions carrying information on spin system of interest may be decomposed into linear combination of eigenstate vectors $|n\rangle$ and corresponding coefficients (c_n), where the square of the latter is proportional to probability amplitude of spin state of interest.

$$|\psi\rangle = \sum_n c_n |n\rangle \tag{2.5}$$

In a matrix form, diagonal elements of the density operator inform on probability of finding an eigenstate in the spin ensemble (which is proportional to its population), while off-diagonal elements describe probability of occurrence of spin transitions associated with evolution of nuclear spin ensemble between different eigenstates (its coherences). For any observable $\mathscr{B}$, an expectation value may be calculated from trace of the density matrix with a spin operator describing spin transition of interest:

$$\langle \mathscr{B} \rangle = \mathrm{Tr}(\sigma \mathscr{B}) \tag{2.6}$$

which allows to translate evolution of nuclear magnetic moments represented by the spin density matrix and respective spin transition operators into experimentally observed parameters describing evolution of the average bulk nuclear magnetization.

2.2 Bloch-Wangness-Redfield theory

In order to describe the evolution of nuclear magnetic moments with respect to experimentally accessible relaxation rates, one may use second-order perturbation theory. The description of nuclear spin ensemble modulation associated with interaction between nuclear magnetic moments and fluctuating magnetic fields of surrounding nuclei and their electronic densities may be initiated from the classical Liouville-von-Neumann equation (eq. 2.7):

$$\frac{d\sigma(t)}{dt} = -i[\mathscr{H}(t), \sigma(t)] = -i[\mathscr{H}_0 + \mathscr{H}_1(t), \sigma(t)] \tag{2.7}$$

which encapsulates the description of changes in energy levels arising due to spin transitions in the ensemble as a result of one of the mechanism of interaction between nuclear magnetic moments. In order to remove the dependence on time-independent

Zeeman interaction between nuclear magnetic moments and the static magnetic field ($\mathscr{H}_0$) one may transform σ(t) and (time-dependent) stochastic part of the Hamiltonian $\mathscr{H}_1$(t) into a so-called interaction frame:

$$\begin{aligned} \sigma^T(t) &= \exp(i\mathscr{H}_0 t)\sigma(t)\exp(-i\mathscr{H}_0 t) \\ \mathscr{H}_1^T(t) &= \exp(i\mathscr{H}_0 t)\mathscr{H}_1(t)\exp(-i\mathscr{H}_0 t) \end{aligned} \tag{2.8}$$

The approach allows to reformulate the Liouville-von-Neumann equation, to the one where intrinsic dependence on $\mathscr{H}_0$ has been removed, while the time-dependence of $\mathscr{H}_1$ has been maintained:

$$\frac{d\sigma^T(t)}{dt} = [\mathscr{H}_1^T(t), \sigma^T(t)] \tag{2.9}$$

The treatment of the time dependence of the spin density matrix proceeds via solution of differential equation within the time bounds from 0 to t:

$$\sigma^T(t) = \sigma^T(0) - i\int_0^t [\mathscr{H}_1^T(t), \sigma^T(t)]dt \tag{2.10}$$

The second-order perturbation approach includes additional integration step following substitution of σ^T in equation 2.10 to 2.9:

$$\sigma^T(t) - \sigma^T(0) = -i\int_0^t \left[\mathscr{H}_1^T(t), \left(\sigma^T(0) - i\int_0^{t'} [\mathscr{H}_1^T(t''), \sigma^T(t'')]dt''\right)\right] dt' \tag{2.11}$$

$$\sigma^T(t) - \sigma^T(0) = -i\int_0^t [\mathscr{H}_1^T(t), \sigma^T(0)]dt' - \int_0^t \int_0^{t'} [\mathscr{H}_1^T(t'), [\mathscr{H}_1^T(t''), \sigma^T(t'')]]dt''dt' \tag{2.12}$$

$$\frac{d\sigma^T(t)}{dt} = -i[\mathscr{H}_1^T(t), \sigma^T(0)] - \int_0^t [\mathscr{H}_1^T(t), [\mathscr{H}_1^T(t'), \sigma^T(t')]]dt' \tag{2.13}$$

Since we are interested in timescales longer than the average correlation time of a molecule (τ on the order of tens ps to tens of ns) and shorter than the rate of decay of bulk magnetization, we can extend limits of integration to infinity. Further, one can observe that during this time evolution of spin system is negligible so that $\langle \mathscr{H}_1^T(t) \rangle$ vanishes and $\sigma^T(t-\tau)$ is approximately σ^T(t). Additionally, the density operator and effects of spin transitions on bulk magnetization evolution described from the second order perturbation theory represent average fluctuations and mutual interactions of nuclear magnetic moments represented by average spin density operator and respective Hamiltonians. To account for the fact that the spin system will seek thermal equilibrium, in order to describe the time evolution of spin system one has to modify σ^T(t) by σ^T(t)-σ_{eq}.

Taking into an account that $\mathscr{H}_1^T$(t) and σ^T(t) are uncorrelated, we can omit the first term in the equation 2.13 and rewrite it as:

$$\frac{d\overline{\sigma^T(t)}}{dt} = -\int_0^\infty \overline{[\mathscr{H}_1^T(t), [\mathscr{H}_1^T(t-\tau)}, \overline{\sigma^T(t)} - \sigma_{eq}]]d\tau \tag{2.14}$$

In the next step, following so-called secular approximation assuming relatively minor influence of the stochastic processes associated with reorientation of the molecule on the frequency of the spin transitions and small probability of inducing motion on timescales which are of interest for NMR relaxation as a result of interaction between the nuclear magnetic moments, one may decompose $\mathscr{H}_1^T$(t) into irreducible spherical tensor operators which depend only on spin part ($\mathbf{A}_k^q$) and time-dependent spatial variables (V$_k^q$(t)):

$$\mathscr{H}_1(t) = \sum_{q=-k}^{k} (-1)^q V_k^{-q}(t)\mathbf{A}_k^q \tag{2.15}$$

where k is rank of the tensor. Spin dependent part is decomposed into linear combination of basis operators, which encapsulate information on frequencies associated with interactions between neighboring nuclear magnetic moments and modulations thereof by electronic density modulation:

$$\mathbf{A}_k^q = \sum_p \mathbf{A}_{kp}^q \tag{2.16}$$

Therefore, the basis operators of spin dependent part of the stochastic Hamiltonian associated with spin interactions ($\mathbf{A}_{kp}^q$) become eigenfunctions of the stochastic Hamiltonian commutation superoperator and ω_p^q its eigenfrequencies:

$$[\mathscr{H}_0, \mathbf{A}_{kp}^q] = \omega_p^q \mathbf{A}_{kp}^q \tag{2.17}$$

Subscript p allows to distinguish operators of the same order q but of different eigenfrequency (p). In other words, interactions between nuclear magnetic moments occurring concurrently to stochastically modulated motions described by spin transition operators at frequencies associated with reorientation of nuclear magnetic moments (spin precession) lead to separation of energy levels associated with alignment of portion of spins along and against the lines of external magnetic field. The incorporation of spin interaction part into the Master equation (2.14) describing perturbation to the time evolution of spin ensemble proceeds with transformation of the spin dependent part into the interaction frame:

$$\mathbf{A}_k^{Tq}(t) = \exp(i\mathscr{H}_0 t)\mathbf{A}_k^q(t)\exp(-i\mathscr{H}_0 t) = \sum_p \exp(\omega_p^q t)\mathbf{A}_{kp}^q \tag{2.18}$$

Therefore, one is considering evolution of nuclear magnetic moments due to spin interactions encapsulated within $\mathbf{A}_{kp}^q$ from perspective of a nuclear spin reorienting in an external magnetic field B_0 at Larmor frequency ω_p^q. We are now in position of inserting the transformed and decomposed Hamiltonian (equation 2.15 and 2.18) into equation describing evolution of spin density operator in the interaction frame (equation 2.14). Assuming that the stochastically modulated time-dependent variables are uncorrelated with tensor spin operators (secular approximation), we can extract the double commutator $\hat{\mathbf{A}}_{kp}^q$ before the integral:

$$\begin{aligned}\frac{d\sigma^T(t)}{dt} =& \sum_{q,q'} \sum_{p,p'} (-1)^{q+q'} \exp\{i(\omega_p^q + \omega_{p'}^{q'})t\} [\mathbf{A}_{kp}^{q'} [\mathbf{A}_{kp}^q, \sigma^T(t) - \sigma_{eq}]] \\ & \int_0^\infty \overline{V_k^q(t) V_k^{-q}(t-\tau) \exp(-i\omega_p^q \tau)} d\tau\end{aligned} \tag{2.19}$$

The terms $(\omega_p^q + \omega_{p'}^{q'}) = (\omega_p^q + \omega_{p'}^{-q}) = (\omega_p^q - \omega_{p'}^q)$ survive only if ω_p and $\omega_{p'}$ are sufficiently different leading to observation of *non-secular* terms, which average to zero due to oscillating magnetic fields generated by the stochastically modulated motions much faster than the ones contributing to the relaxation, and do not contribute to long-term value of σ^T. Practically, frequency modulation due to perturbation of nuclear magnetic moments by fluctuating magnetic fields and electronic density in the vicinity of the nuclear spin associated with transient modifications to molecular environment of the spin is either relatively small or short-lived that the term $\exp\{i(\omega_p^q + \omega_{p'}^{q'})\mathrm{t}\}$ may be omitted (*secular approximation*).

Consequently, one may redefine the Master equation describing evolution of the spin density operator:

$$\frac{d\sigma^T(t)}{dt} = \sum_{q,q'}\sum_{p,p'}[\mathbf{A}_{kp}^{q'}\,[\mathbf{A}_{kp}^{q}, \sigma^T(t) - \sigma_{eq}]] \int_0^\infty \overline{V_k^q(t)V_k^{-q}(t-\tau)}\exp(-i\omega_p^q\tau)d\tau \tag{2.20}$$

We will now recognize that product of spatial part of spin-lattice interaction tensor and its complex conjugate denotes average correlation function associated with reorientation of the molecule or its fragment:

$$C(t) = \overline{V_k^q(t)V_k^{-q}(t-\tau)} \tag{2.21}$$

while the real part of integral of the correlation function:

$$j^q(\omega_p^q) = 2Re\left\{\int_0^\infty \overline{V_k^q(t)V_k^{-q}(t-\tau)}\exp(-i\omega_p^q\tau)d\tau\right\} \tag{2.22}$$

is a power SDF which describes probability of observing spin transitions at certain frequency associated with a nuclear spin interaction arising due to the stochastically modulated motions. The imaginary part of the integral contributes to second-order frequency shifts which description is beyond the scope of this work. Defined in that manner Master equation may be transformed into the laboratory frame:

$$\frac{d\sigma(t)}{dt} = -[\mathscr{H}_0, \sigma(t)] - \hat{\hat{\Gamma}}(\sigma(t) - \sigma_{eq}) \tag{2.23}$$

with relaxation superoperator ($\hat{\hat{\Gamma}}$) composed of SDF describing probability of observing allowed spin transitions and respective spin transition double commutator encapsulating perturbations to the evolution of nuclear magnetic moments:

$$\hat{\hat{\Gamma}} = \frac{1}{2} \sum_{q=-k}^{k} \sum_{p} [\mathbf{A}_{kp}^{-q}\, [\mathbf{A}_{kp}^{q},]] j^{q}(\omega_{p}^{q}) \tag{2.24}$$

The Master equation may be applied to spin systems of interest under condition of non-vanishing spin interaction double commutator $[\mathbf{A}_{kp}^{-q}\, [\mathbf{A}_{kp}^{q},]]$, and non-negligible SDF describing reorientation of a molecule occurring at different timescales. The latter is maximized when timescales of reorientation of a molecule or its fragment is equal to inverse of frequency associated with spin interaction ω_{p}^{q} - Larmor frequency or linear combination of Larmor frequencies of the interacting nuclei. Derivation of relative contributions of spin interactions to relaxation rate constants for a spin system follows a two step process: first double commutators of respective spin-tensor operators are calculated, then a trace of the double commutation product with coherence order of interest is taken. The final step in evaluation of relaxation rate constant is calculation of SDF $j^{q}(\omega_{p}^{q})$ which depends on reorientation rate of (macro)molecule fragment containing spin(s) of interest (188).

2.3 Mechanisms of relaxation

Fluctuating magnetic fields associated with reorientation of the molecule lead to spin transitions may be encapsulated within description of dipolar interactions between nuclear magnetic moments of (located in the vicinity) nuclei featuring spin half-integer quantum number (direct dipolar interaction), as a result of spatio-temporal perturbation to electronic density surrounding the nuclei represented by chemical shift anisotropy (CSA) influencing nuclear magnetic moment evolution (indirect dipolar interaction), or as a result of interference between the two mechanisms (CSA-dipole-dipole induced relaxation). Additionally, nuclei of integer spin quantum number feature nuclear quadrupolar moments and their interaction with local electric field gradients may lead to additional relaxation contributions. If reorientation rate of a (macro)molecule fragment is proportional to the inverse of the eigenfrequencies associated with transition between the spin states due to one of the relaxation mechanisms, one may observe non-negligible contribution to evolution of bulk magnetization represented by relaxation rate constants. For coherences in spin systems with quantum numbers l equal to 1/2, of major interest are single and zero quantum operators describing transverse and longitudinal relaxation rates accessible experimentally (44, 188).

2.3.1 Dipole-dipole induced relaxation

The evolution of nuclear magnetic moments occurring concurrently to reorientation of a (macro)molecule leads to modulation of fluctuating magnetic fields within interacting nuclei. Considering a spin pair featuring half-integer quantum numbers (further denoted as IS spin pair) separated by a distance (r_{IS}), the partial nuclear magnetic moments of nuclei I and S influence one another, in turn leading to modulation of electromagnetic fields within the nuclei associated with redistribution of partial magnetic moments (populations of fundamental particles composing them and their orientations), and consequently to transient perturbation to direction of the resultant nuclear magnetic moments with respect to the external magnetic field. For nuclei featuring half-integer quantum numbers the process of orientation randomization of nuclear magnetic moments in the spin ensemble associated with their mutual interactions leads to loss of coherence of magnetization evolution, which may be represented by bulk nuclear magnetization decay observed experimentally. The magnetic dipole-dipole interaction occurs through space and its magnitude depends on gyromagnetic ratios of the coupled nuclei I and S and distance between them:

$$d = \frac{-\sqrt{6}\mu_0\hbar\gamma_I\gamma_S}{4\pi} r_{IS}^{-3}(t) \tag{2.25}$$

For an IS spin pair, the strength of interaction (d) decays with an inverse sixth power of distance. In some cases, the information may be used to calculate distance between the interacting nuclei. The Hamiltonian for dipolar interaction is formally written as:

$$\mathscr{H}_0^{dip} = -d\left(3\hat{\mathbf{I}}\hat{\mathbf{r}}\hat{\mathbf{S}}\hat{\mathbf{r}} - (\hat{\mathbf{I}}\cdot\hat{\mathbf{S}})\right) = -d\left(3\hat{\mathbf{I}}(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}^T)\hat{\mathbf{S}} - (\hat{\mathbf{I}}\cdot\hat{\mathbf{S}})\right) = -d\hat{\mathbf{I}}\left(3\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}^T - \hat{1}\right)\hat{\mathbf{S}} \tag{2.26}$$

where $\hat{\mathbf{r}}$ describes a unit vector. Formally, the dipolar interaction tensor $\left(\hat{\mathbf{I}}\left(3\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}^T - \hat{1}\right)\hat{\mathbf{S}}\right)$ may be grouped into six products of direction cosines, forming a so-called dipolar alphabet. It is generally more useful to express these terms in a form of irreducible spherical harmonics describing redistribution of partial nuclear magnetic moments associated with respective spin transitions.

Table 1. Spin tensor operators of respective transition operators according to their frequency (p) and coherence order (q).

q	p	A_{2p}^{q}	$\mathrm{A}_{2p}^{-q} = (\text{-}1)^q \mathrm{A}_{2p}^{-q\dagger}$	ω_p^q
0	0	$(2/\sqrt{6})I_zS_z$	$(2/\sqrt{6})I_zS_z$	0
0	-1	$-(1/2\sqrt{6})I^-S^+$	$-(1/2\sqrt{6})I^+S^-$	$\omega_S - \omega_I$
0	1	$-(1/2\sqrt{6})I^+S^-$	$-(1/2\sqrt{6})I^-S^+$	$\omega_I - \omega_S$
1	0	$-(1/2)I_zS^+$	$(1/2)I_zS^-$	ω_S
1	1	$-(1/2)I^+S_z$	$(1/2)I^-S_z$	ω_I
2	0	$(1/2)I^+S^+$	$(1/2)I^-S^-$	$\omega_S + \omega_I$

In order to estimate relaxation rate constants, for instance, of longitudinal magnetization evolution $\mathrm{I}_z\mathrm{S}_z$, one needs to perform a double commutator operation followed by trace calculation with respective spin tensor operator. To exemplify operations required to perform to estimate the relaxation rate constant, one may consider evolution of ZQ coherence order associated with flip-flop spin transitions ($\alpha\beta \rightarrow \beta\alpha$ with p of 1 and q of 0):

$$[A_{21}^0[A_{21}^0, I_z]] = (1/24)[I^-S^+[I^+S^-, I_z]] = (1/24)[I^-S^+, I^+S^-] = (1/24)\{I_z - S_z\} \quad (2.27)$$

The inner commutator is isomorphous which simplifies the calculation. In the following step, one estimates trace of I_z with the product of double commutator (equation 2.27):

$$\begin{aligned}(1/24)\left\langle I_z|\{I_z - S_z\}\right\rangle &= (1/24)\left\langle I_z^2 - I_zS_z\right\rangle = (1/24)(\langle\alpha\alpha|I_z^2 - I_zS_z|\alpha\alpha\rangle \\ &+ \langle\alpha\beta|I_z^2 - I_zS_z|\alpha\beta\rangle + \langle\beta\alpha|I_z^2 - I_zS_z|\beta\alpha\rangle + \langle\beta\beta|I_z^2 - I_zS_z|\beta\beta\rangle) = (1/24)\end{aligned} \quad (2.28)$$

giving the quantum-mechanical constant for the flip-flop ($\omega_I - \omega_S$) transition for the longitudinal two spin order $\mathrm{I}_z\mathrm{S}_z$. In principle, to estimate the relaxation rate, one has to estimate all of the combinations of commutation relationships (44).

For the dipole-dipole induced evolution of magnetization, the relaxation rates are given by:

$$R_{Dip}(I_z) = \frac{d^2}{4}\left(J(\omega_I - \omega_S) + 3J(\omega_I) + 6J(\omega_I + \omega_S)\right) \tag{2.29}$$

$$R_{Dip}(I^{+/-}) = \frac{d^2}{8}\left(4J(0) + J(\omega_I - \omega_S) + 3J(\omega_I) + 6J(\omega_S) + 6J(\omega_I + \omega_S)\right) \tag{2.30}$$

$$R_{Dip}(I_z \leftrightarrow S_z) = \frac{d^2}{4}\left(-J(\omega_I - \omega_S) + 6J(\omega_I + \omega_S)\right) \tag{2.31}$$

where the subset notation describes the coherence order evolution - the first two equations describe dipole-dipole induced relaxation of longitudinal and transverse magnetization components respectively. The last equation describes the heteronuclear dipole-dipole cross-relaxation.

2.3.2 Chemical-shift anisotropy induced relaxation

Nuclei are surrounded by cloud of electrons represented by electronic density, which in the presence of external magnetic field ($\mathbf{B}_0$) create additional, local oscillating magnetic fields ($\mathbf{B}_{loc}$) around and within the nuclei effectively modifying the partial nuclear magnetic moments distribution, and in turn transiently affecting the nuclear magnetic moment orientation. The magnitude with which local magnetic fields generated by electronic density fluctuations modify nuclear magnetic moments is commonly referred to shielding, and allows to spectroscopically distinguish nuclei of the same kind in different molecular environments, and is commonly referred to as chemical shift.

$$\mathbf{B}_{loc} = \mathbf{B}_0(1 - \sigma) \tag{2.32}$$

The nuclear substructure of a molecule depends on fundamental particle composition of the nuclei in turn affecting electronic substructure and potential for formation of chemical bonds and non-bonding interactions arising due to imbalanced polarity within the (macro)molecule fragments, which in turn may affect the nuclear magnetic moments evolution in the external magnetic field. The molecular architecture of group(s) of atoms, namely the type of atoms and interactions between them, in the vicinity of a nuclei leads to a degree of anisotropy in the nuclear magnetic shielding (σ), which may be conveniently represented by (chemical shielding) rank 2 tensor:

$$\hat{\sigma} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} \tag{2.33}$$

where σ_{yx} represents shielding constant associated with magnetic field induced along y-axis with external magnetic field B_0 along x-axis. For further calculations $\hat{\sigma}$ is typically decomposed into isotropic part, anisotropic anti-symmetric part and anisotropic symmetric part:

$$\begin{aligned} \hat{\sigma} = \sigma_{iso} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & \sigma_{xy} - \sigma_{yx} & \sigma_{xz} - \sigma_{zx} \\ -\sigma_{xy} + \sigma_{yx} & 0 & \sigma_{yz} - \sigma_{zy} \\ -\sigma_{xz} + \sigma_{zx} & -\sigma_{yz} + \sigma_{zy} & 0 \end{pmatrix} \\ + \Delta\sigma \begin{pmatrix} -0.5(1-\eta) & 0 & 0 \\ 0 & -0.5(1+\eta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \tag{2.34}$$

with isotropic part of the shielding tensor $\sigma_{iso} = (\sigma_{xx} + \sigma_{yy} + \sigma_{zz})/3$, chemical shielding anisotropy $\Delta\sigma = (\sigma_{zz} - (\sigma_{xx} + \sigma_{yy})/2)$ and rhombicity: $\eta = (\sigma_{xx} - \sigma_{yy})/\sigma_{zz}$. By convention:

$$|\sigma_{yy} - \sigma_{iso}| \leq |\sigma_{xx} - \sigma_{iso}| \leq |\sigma_{zz} - \sigma_{iso}| \tag{2.35}$$

When rhombicity vanishes ($\sigma_{xx} = \sigma_{yy} = \sigma_{\perp}$), the tensor becomes axially symmetric with symmetry axis ($\sigma_{zz} = \sigma_{\parallel}$), then $\Delta\sigma = \sigma_{\parallel} - \sigma_{\perp}$. The oscillating magnetic fields generated by reorientation of a molecule temporally modify effects of shielding leading to perturbation of evolution of spin system in turn inducing relaxation.

The description of chemical shift anisotropy influence on the local magnetic properties of a nuclei are represented by CSA Hamiltonian:

$$\mathscr{H}_{CSA} = \hbar\gamma_I B_0 \hat{\mathbf{I}}\hat{\sigma}\hat{\mathbf{B}} \tag{2.36}$$

where $\hat{\sigma}$ is the chemical shift anisotropy tensor, and $\hat{\mathbf{I}}$ and $\hat{\mathbf{B}}$ are respective spin and (external) magnetic field tensors. Of primary interest is symmetric and anisotropic part of the interaction tensor as fluctuations encapsulated within the isotropic part contribute to overall NMR relaxation to a small extent and can be treated as a part of the Zeeman interaction Hamiltonian. To evaluate effect of CSA on relaxation, one needs to estimate contributions associated with spin interactions at frequencies associated with relaxation active spin transitions. The irreducible spherical tensor representation of spin transitions with regard to coherence order of the CSA Hamiltonian is defined in Table 2.

Table 2. Spin tensor operators of respective transition operators according to their frequency (p) and coherence order (q).

q	p	A_{2p}^{q}	$A_{2p}^{-q} = (-1)^q A_{2p}^{-q\dagger}$	ω_p^q
0	0	$(2/\sqrt{6})I_z$	$(2/\sqrt{6})I_z$	0
1	0	$-(1/2)I^+$	$(1/2)I^-$	ω_I
2	0	-	-	$2\omega_I$

Calculation of double commutators and trace with respective spin tensor operator, allows to estimate the CSA-induced contribution to the relaxation rate constants:

$$R_{CSA}(I_z) = c^2 J(\omega_I) \tag{2.37}$$

$$R_{CSA}(I^{+/-}) = c^2 \left(4J(0) + 3J(\omega_I)\right) \tag{2.38}$$

The magnitude of CSA interaction strength (c) is directly proportional to the magnitude of chemical shielding anisotropy and strength of the external magnetic field:

$$c = \frac{\Delta\sigma\gamma_I B_0}{\sqrt{3}} \tag{2.39}$$

2.3.3 Chemical-shift anisotropy-dipole-dipole interference

Under certain set of circumstances, when the evolution of nuclear magnetic moments of nuclei pair due to the dipolar interaction between them and the time-dependent fluctuations of electronic density surrounding the nuclei are correlated in time and the associated spin transitions do not commute, the two stochastic processes give rise to the cross-correlated cross-relaxation. An example of the phenomenon is CSA-dipole-dipole interference (189). From the point of view of Bloch-Wangness-Redfield theory, the cross-correlation function describing the interference is composed of spatial part of spin-lattice interaction tensors describing the chemical-shift anisotropy and the dipolar interaction mechanisms. The comparison of mutual orientation of spin-lattice interaction tensors describing the chemical-shift anisotropy and direct interaction between nuclear magnetic moments, leads to introduction of component allowing to take into an account relative orientation of the two, so that the correlation function associated with interference between the two mechanisms may be approximated by an auto-correlation function with maintained magnitude of the interactions. For axially symmetric CSA tensors featured by the ^{15}N-^{1}H nuclei pair possessing relatively small degree of anisotropy (η is approximately 0.22), one may achieve the comparison of cross and auto-correlation functions by application of $\mathrm{P}_2(\cos(\theta))$, where θ angle defines orientation of principal axis of the CSA tensor with respect to the dipolar interaction tensor (which is collinear with the NH-backbone vector)(190–194). From the point of view of magnetization coherence evolution leading to the cross-correlated cross-relaxation, initially formed in-phase magnetization coherence evolves into an anti-phase (or *vice versa*), with a loss of magnetization proportional to the cross-relaxation rate associated with interference between the two mechanisms.

$$\frac{d\langle I^+\rangle}{dt} = -i\pi J_{IS}\langle 2I^+S_z\rangle(t) - (R_{DD}+R_{CSA})\langle I^+\rangle(t) - \eta_{xy}\langle 2I^+S_z\rangle(t) \tag{2.40}$$

$$\frac{d\langle I^+S_z\rangle}{dt} = -i\pi J_{IS}\langle 2I^+\rangle(t) - (R_{DD}+R_{CSA})\langle I^+S_z\rangle(t) - \eta_{xy}\langle 2I^+\rangle(t) \tag{2.41}$$

The equations show that for the anti-phase magnetization coherence of the IS spin system, one of the peaks in the doublet will be broader, while the other narrower, both by a factor proportional to η_{xy}, the latter defining the cross-relaxation rate constant between the two mechanisms. This phenomenon is a basis of a Transverse Relaxation Optimized SpectroscopY (TROSY) experiment, which allows for reduction of the apparent relaxation rate of one of the components of the doublet, and allows for spectroscopic evaluation of the nuclei in larger macromolecules featuring faster intrinsic transverse relaxation rate due to relatively long reorientation timescales (exceeding 20-30 ns).

2.3.4 Relaxation rates summary

Following application of the semi-classical second-order perturbation theory within the bounds of secular approximation for ^{15}N nuclei located in the protein backbone, the experimentally accessible relaxation rates featured in the following analyses may be approximated by following expressions (equations 2.42), (195):

$$R_2 = \frac{d^2}{20}[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + \frac{c^2}{15}[4J(0) + 3J(\omega_N)] \tag{2.42}$$

$$R_1 = \frac{d^2}{10}[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)] + \frac{2c^2}{15}J(\omega_N) \tag{2.43}$$

$$nOe = 1 + \frac{d^2}{10}\frac{\gamma_H}{R_1\gamma_N}[-J(\omega_H - \omega_N) + 6J(\omega_H + \omega_N)] \tag{2.44}$$

$$\eta_{xy} = \frac{\sqrt{3}}{15}cd\left(\frac{3cos^2\theta - 1}{2}\right)[4J(0) + 3J(\omega_N)] \tag{2.45}$$

$$\eta_z = \frac{2\sqrt{3}}{5}cd\left(\frac{3cos^2\theta - 1}{2}\right)J(\omega_N) \tag{2.46}$$

$$\text{with } c = \frac{\Delta\sigma\omega_N}{\sqrt{3}} \text{ and } d = \frac{\mu_0}{4\pi}\frac{\hbar\gamma_N\gamma_H}{r_{NH}^3} \tag{2.47}$$

where $\Delta\sigma$ denotes chemical shift anisotropy (CSA) for ^{15}N-^{1}H system (with experimentally determined values in the range of -160 to -180 ppm), ω_{H} and ω_{N} denote ^{1}H and ^{15}N Larmor frequencies, γ_{H} and γ_{N} correspond to ^{1}H and ^{15}N gyromagnetic ratios, μ_0 is the magnetic permeability of free space, $\hbar$ is the reduced Planck constant, r_{NH} is the ^{15}N-^{1}H internuclear distance (1.015Å) and θ is the angle between principal axis of the CSA tensor and the NH-backbone vector. R$_1$ and R$_2$ represent auto-relaxation rates (longitudinal and transverse respectively) and describe the bulk nuclear magnetization evolution arising due to dipolar interactions of both direct and indirect origin - between nuclei located in the vicinity or modulations of nuclear magnetic moments by the surrounding electronic density respectively. R$_1$ denotes rate at which bulk nuclear magnetization comes back to equilibrium (with bulk magnetization vector aligned along the lines of external B$_0$ field), R$_2$ is rate at which bulk nuclear magnetization decays in the transverse plane (perpendicular to the lines of B$_0$ field). nOe describes the steady-state heteronuclear Overhauser effect, which arises due to cross-relaxation between longitudinal coherence orders of ^{1}H and ^{15}N. The magnetic relaxation interference between the two mechanisms originating due to direct and indirect dipolar interaction for ^{15}N nuclei in the transverse and longitudinal planes are represented by η_{xy} and η_z respectively (equations 2.45 and 2.46).

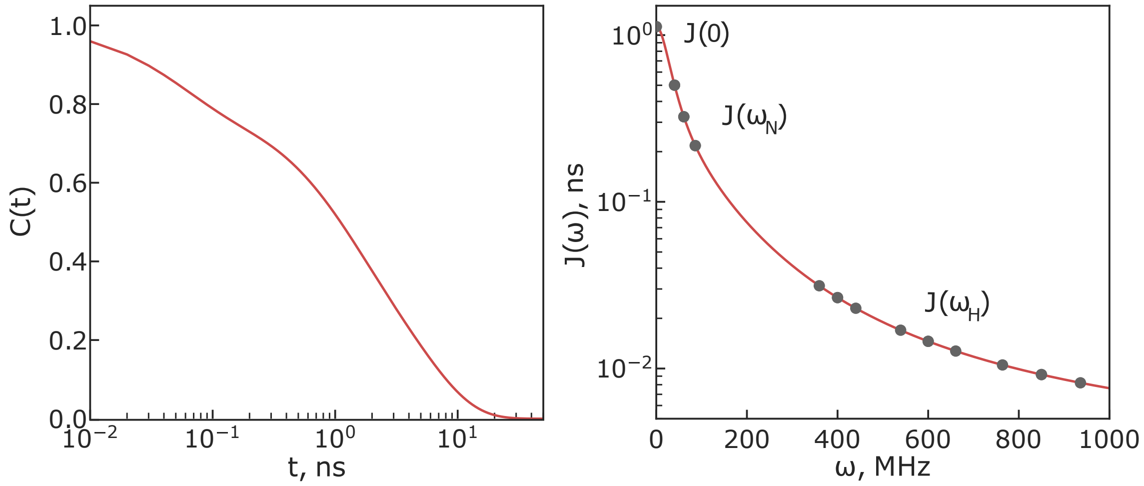

Figure 2.1: Comparison of exemplary correlation function and corresponding SDF. Points represent SDF at frequencies associated with nuclear spin interactions, i.e. Larmor frequencies of ^{15}N and ^{1}H at different magnetic field strengths of 9.4, 14.1 and 20T.

Considering expressions provided by solving Bloch-Wangness-Redfield theory for ^{15}N-^{1}H system (equations 2.42 to 2.46) and an exemplary SDF for ^{15}N backbone amide such as the one in figure 2.1 one may realize that different relaxation rates are sensitive to motional processes occurring on different decades of timescales. For instance, R_2 being dominated by J(0) contains information about the total amount of dynamics (and is equal to the integral of the correlation function) which is sensitive to motions occurring on longest relaxation active timescales (tens of ns). R_1 is dominated by its $J(\omega_N)$ component, and, consequently, is more sensitive to reorientational motions occurring on 100 ps to 1 ns timescales. Steady-state heteronuclear nOes depend on $J(\omega_H)$ components, thus reporting on dynamic modes occurring on tens of ps to hundreds of ps.

CHAPTER 3

Analysis of protein dynamics from NMR relaxation rates

^{15}N spin relaxation may be used to characterize reorientational properties of peptide units on a range of timescales, from tens of ps to tens of ns. The conformational heterogeneity of unfolded proteins, and the number of degrees of freedom available to backbone vectors constituting amide planes suggests presence of multiple reorientational processes occurring on several length and timescales, in turn giving rise to distributions of timescales of motions, rather than discrete values representing (average) contributions to the SDF. The experimental sensitivity, as well as, the complexity of motions underlying the distributions limits Our understanding of nature of the reorientation processes. Model-Free representation of NMR relaxation rates allows to characterize apparent dynamic modes of the distribution(s) of timescales of motions, characterizing average reorientational rate of the peptide units at different length-scales with respect to experimental sensitivity of a set of relaxation rates at different magnetic fields. The chapter presents limitations in description of reorientation processes of polypeptide fragments represented by average dynamic modes using Model-Free analysis associated with experimental sensitivity and complexity of motions exhibited by unfolded polypeptide chain. Model-Free treatment of NMR relaxation rates originating from arbitrary distributions of correlation times, as well as, comparison of MD-derived NH-backbone dynamics with information originating from Model-Free treatment of the underlying distributions of dynamic modes allows to obtain potential route of interpretation of unfolded protein dynamics within the bounds of adiabatic approximation.

3.1 Introduction

3.1.1 The fundamental aspects of Model-Free analysis

Historically, one of the most common approach in description of NMR relaxation rates arising due to allowed spin transitions occurring concurrently to stochastically modulated motions is so-called Model-Free analysis. Originally introduced in 1981, on an exemplary analysis of lysine side-chain motions, Model-Free analysis allows to represent time-dependent fluctuations of spatial functions of spin-lattice interaction tensor associated with reorientation of polypeptide fragments in a form of dynamic modes of distributions of correlation times with respect to sensitivity of the relaxation rates, which depending on type of dynamics exhibited by the (macro)molecule fragment containing spin system of interest may be modeled as a sum or a product of sequentially decaying exponential functions. These in turn, are characterized by parameters describing spatial restriction of a bond vector of interest (so-called order parameter, S) and timescale associated with reorientation (τ) of entire macromolecule, fragments thereof or a bond vector of interest to which spin-lattice tensor is attached to (79).

The application of Model-Free analysis to NMR relaxation rates describing average evolution of spin system of interest requires establishment of connection between local and global degrees of freedom limiting stochastically modulated motions leading to reorientation of spatial parts of spin-lattice interaction tensor. If reorientation of the latter leads to allowed spin transitions at respective frequencies and coherence orders associated with spin interactions, the information on stochastically modulated motions is encoded in experimentally accessible relaxation rates describing result of the interactions on evolution of coherence. In order to derive relationship between the relaxation rates and the stochastically modulated motions, one typically begins considerations by representing cumulative fluctuations of spin-lattice interaction tensor $\mathbf{V}$ of one of the relaxation mechanisms as auto-correlation function (ACF) (78):

$$C(t) = \langle \mathbf{V}(0) : \mathbf{V}(t) \rangle \tag{3.1}$$

where brackets indicate time averaging, while semicolon a scalar contraction which can be evaluated in any fixed coordinate system (e.g. in laboratory fixed frame with director pointing along the lines of external magnetic field). The rank 2 trace-less symmetric tensor in laboratory frame $V_m^{(L)}$ may be decomposed into spherical harmonics describing the distribution of partial nuclear magnetic moments associated with electromagnetic flux originating due to oscillations of fundamental particles within the nuclei:

$$C(t) = \sum_{m=-2}^{2} \left\langle V_m^{(L)*}(0) V_m^{(L)}(t) \right\rangle \tag{3.2}$$

Coupling of laboratory-frame component $V_0^{(L)}$ to tensors describing the reorientation processes is achieved via a series of coordinate system transformations via additional coordinate frames. In descending degrees of complexity of the reorientational processes under evaluation, transition from Laboratory frame to spin-lattice interaction tensor proceeds via transformation through global mobility tensor (G), local potential of mean torque (POMT), internal mobility tensor (I), and a reference frame for POMT describing reorientation of localized internal motions with respect to global mobility tensor (D). Each mobility tensor may be either isotropic, uniaxial or biaxial (with one, two or three distinct principal components). These transformations are typically achieved by successive application of rank 2 Wigner rotation tensors, which align axes of coordinate system of choice in-between subsequent frames:

$$V_0^{(L)} = \sum_n \sum_p \sum_q \sum_r \times D_{0n}^{2*}(\Omega_{\mathrm{LG}})\ D_{np}^{2*}(\Omega_{\mathrm{GD}})\ D_{pq}^{2*}(\Omega_{\mathrm{DI}})\ D_{qr}^{2*}(\Omega_{\mathrm{IF}})\ V_r^{(F)} \tag{3.3}$$

where sums represents different projection index components ranging from -2 to $+2$ due to distinct influence of motions of different complexity (isotropic, symmetric or asymmetric diffusion) on evolution of nuclear magnetic moments. The Wigner rotation matrices may be described using Euler formalism with angles Ω_{AB} representing constant parameters which define relative orientation of respective mobility tensors. Together these coefficients may be expressed as:

$$\sigma_q = \sum_q D_{qr}^{2*}(\Omega_{\mathrm{IF}})\ V_r^{(F)*} \tag{3.4}$$

where σ_q is defined as:

$$\sigma_q = \left(1 + \frac{\eta^2}{3}\right)^{-1/2} \left(D_{q0}^{2}(\Omega_{\mathrm{IF}}) + \frac{\eta}{\sqrt{6}} \left(D_{q-2}^{2}(\Omega_{\mathrm{IF}}) + D_{q2}^{2}(\Omega_{\mathrm{IF}}) \right) \right) \tag{3.5}$$

with normalization requirement: $\sum_q |\sigma_q|^2 = 1$. The internal and global motions are then expressed as time-dependent Euler angles $\Omega_{\mathrm{DI}}(\mathrm{t})$ and $\Omega_{\mathrm{LG}}(\mathrm{t})$ respectively. With explicit time dependence one can now write:

$$V_0^{(L)}(t) = \sum_n \sum_p \sum_q \times\ D^{2*}_{0n}(\Omega_{\mathrm{LG}}(\mathrm{t}))\ D^{2*}_{np}(\Omega_{\mathrm{GD}})\ D^{2*}_{pq}(\Omega_{\mathrm{DI}}(\mathrm{t}))\sigma^*_q \tag{3.6}$$

Substituting equation 3.6 to equation 3.2 one may obtain an ACF of general form composed of time-dependent and time-independent components associated with relative orientation of reference frames describing the reorientational processes of different complexity:

$$\begin{gathered} C(t) = 5 \sum_n \sum_{n'} \sum_p \sum_{p'} \sum_q \sum_{q'} D^2_{np}(\Omega_{\mathrm{GD}}) D^{2*}_{n'p'}(\Omega_{\mathrm{GD}}) \sigma_q \sigma^*_q \\ \left\langle D^2_{0n}(\Omega^0_{\mathrm{LG}}) D^{2*}_{0n'}(\Omega_{\mathrm{LG}}) D^2_{pq}(\Omega^0_{\mathrm{DI}}) D^{2*}_{p'q'}(\Omega_{\mathrm{DI}}) \right\rangle \end{gathered} \tag{3.7}$$

where the time-independent Wigner rotation matrices representing transformation between fixed coordinate systems were excluded from the ensemble averaging. The calculation of joint ensemble average of time-dependent components describing relative orientation frames of reference is a difficult task, and for successful formulation of any result, one has to decompose it into a product or a sum of two separate ensemble averages within bounds of either superposition or adiabatic approximation respectively.

Superposition approximation

Stochastic processes may be described by a set of time-dependent probability distributions usually expressed in form of propagators (e.g. $\mathrm{P}(\Omega_{\mathrm{LG}},\mathrm{t}|\Omega^0_{\mathrm{LG}})$). The propagators represent probability amplitude of finding orientation represented by Euler angles within certain angle interval at time t (e.g. $\mathrm{d}\Omega_{\mathrm{LG}}$ within Ω_{LG} given that the angle was Ω^0_{LG} initially). More specifically, when two motions occurring concurrently (e.g. global and internal) are investigated, a joint propagator $\mathrm{P}(\Omega_{\mathrm{LG}},\Omega_{\mathrm{DI}},\mathrm{t}|\Omega^0_{\mathrm{LG}},\Omega^0_{\mathrm{DI}})$ has to be considered with limiting values defined by:

$$P(\Omega_{\mathrm{LG}}, \Omega_{\mathrm{DI}}, 0|\Omega^0_{\mathrm{LG}}, \Omega^0_{\mathrm{DI}}) = \delta(\Omega_{\mathrm{LG}} - \Omega^0_{\mathrm{LG}})\delta(\Omega_{\mathrm{DI}} - \Omega^0_{\mathrm{DI}}) \tag{3.8}$$

$$\lim_{t\to\infty} P(\Omega_{\mathrm{LG}}, \Omega_{\mathrm{DI}}, t|\Omega^0_{\mathrm{LG}}, \Omega^0_{\mathrm{DI}}) = P_{eq}(\Omega_{\mathrm{LG}}, \Omega_{\mathrm{DI}}) \tag{3.9}$$

where δ denotes Kronecker delta, which is equal to value of 1 at x equal to 0 and is 0 otherwise. If one of reorientation processes describing global reorientation (G) in laboratory frame (L) or internal motion (I) in director frame (D) occurs, the joint propagator at initial time will yield zero. At the infinitely long time intervals the joint probability approaches equilibrium distribution, which is typically represented by normalized Boltzmann distribution:

$$P_{eq}(\Omega_{\mathrm{LG}}, \Omega_{\mathrm{DI}}) = \frac{exp\left[-U(\Omega_{\mathrm{LG}}, \Omega_{\mathrm{DI}})/k_B T\right]}{\int d\Omega_{\mathrm{LG}} \int d\Omega_{\mathrm{DI}} exp\left[-U(\Omega_{\mathrm{LG}}, \Omega_{\mathrm{DI}})/k_B T\right]} \tag{3.10}$$

where U(Ω_{LG},Ω_{DI}) is a potential of mean torque (POMT), which describes the energetic landscape associated with mutual orientation of I and L frames. The potential describing joint reorientation process is typically decomposed into a sum of potentials, which allows to factorize the joint equilibrium distribution (equation 3.10):

$$U(\Omega_{\mathrm{LG}}, \Omega_{\mathrm{DI}}) = U(\Omega_{\mathrm{LG}}) + U(\Omega_{\mathrm{DI}}) \tag{3.11}$$

$$P_{eq}(\Omega_{\mathrm{LG}}, \Omega_{\mathrm{DI}}) = P_{eq}(\Omega_{\mathrm{LG}}) P_{eq}(\Omega_{\mathrm{DI}}) \tag{3.12}$$

If the joint equilibrium distribution may be factorized this way, the two sets of stochastic variables are said to be statistically independent. In other words, if the energetic potential describing global reorientation processes is relatively featureless with respect to local reorientation potential - the joint potential may be decomposed into simple sum of potentials (with features of different degree of complexity) describing processes associated with reorientation of entire macromolecule and fragments thereof represented by fluctuations of global and local mobility tensors respectively. The factorization of the equilibrium probability distribution is a necessary but insufficient condition for the statistical independence of motions. If the reorientation potential describing energetic transitions between conformational states of a protein is subject to time evolution due to modifications in number of degrees of freedom of reorienting polypeptide fragment associated with inter or intramolecular interactions, the stochastic processes described by fluctuations of angles Ω_{LG}(t) and Ω_{DI}(t) will be statistically dependent on one another, which is not the case otherwise. In other words, if by virtue of intra or intermolecular interaction the local mobility tensor reorients slower, with timescale comparable to reorientation of global roerientation rate the two reorientation processes are not fully statistically independent.

The factorization of the joint propagator is a basis of superposition approximation stating that total reorientation of I frame relative to L is a pure superposition of motion of I relative to D and G relative to L with D frame fixed with respect to G frame:

$$P(\Omega_{\mathrm{LG}}, \Omega_{\mathrm{DI}}, t | \Omega^0_{\mathrm{LG}}, \Omega^0_{\mathrm{DI}}) = P(\Omega_{\mathrm{LG}}, t | \Omega^0_{\mathrm{LG}}) P(\Omega_{\mathrm{DI}}, t | \Omega^0_{\mathrm{DI}}) \tag{3.13}$$

With that in mind, one can go back to the equation 3.7 and decompose average Wigner rotation matrix describing the two stochastic process into product of averages of two independent processes, of which the first ensemble average is associated with global reorientation process, while the latter with local, internal motions:

$$\left\langle D^2_{0n}(\Omega^0_{\mathrm{LG}}) D^{2*}_{0n'}(\Omega_{\mathrm{LG}}) \right\rangle = \delta_{nn'} exp(-t/\tau_{global}) = \delta_{nn'} C_{global}(t) \tag{3.14}$$

If C$_{global}$ is independent of G-frame projection index n, the correlation function C(t) may be factorized into a product of correlation functions associated with global and local reorientation processes (C$_{global}$(t) and G$_{internal}$(t)) with the latter given as:

$$G_{internal}(t) = \sum_q \sum_{q'} \sigma_q \sigma^*_{q'} \sum_p \left\langle D^2_{pq}(\Omega^0_{\mathrm{DI}}) D^{2*}_{p'q'}(\Omega_{\mathrm{DI}}) \right\rangle \tag{3.15}$$

At infinite time Ω_{DI} must be independent of its initial orientation (Ω^0_{DI}), thus G$_{internal}(\infty)$ is equal to S^2:

$$S^2 = \sum_p \left| \sum_q \sigma_q \left\langle D^2_{pq}(\Omega_{\mathrm{DI}}) \right\rangle \right|^2 \tag{3.16}$$

Inserting the equations 3.14 and 3.15 into 3.7 yields the most classic Model-Free formula:

$$C(t) = exp(-t/\tau_{global}) \left[S^2 + (1 - S^2)C_{internal}(t)\right] \tag{3.17}$$

If no symmetry is present the generalized order parameter S depends on 25 independent partial order parameters $\left\langle D^2_{pq}(\Omega_{\mathrm{DI}})\right\rangle$. If the local POMT is uniaxial in the I frame, only five order parameters $\left\langle D^2_{p0}(\Omega_{\mathrm{DI}})\right\rangle$ are non-zero:.

$$S^2 = \sum_p \left|\left\langle D^2_{p0}(\Omega_{\mathrm{DI}})\right\rangle\right|^2 \tag{3.18}$$

If the local POMT is uniaxial in the D frame, only five terms with p equal to 0 survive.

$$S^2 = \left|\sum_q \sigma_q \left\langle D^2_{0q}(\Omega_{\mathrm{DI}})\right\rangle\right|^2 \tag{3.19}$$

Adiabatic approximation

The adiabatic approximation relies on time-scale separation of motions featuring different degrees of freedom. The stochastic variables Ω_{LG}(t) and Ω_{DI}(t) are said to be timescale separated if there exist a time interval t* during which Ω_{LG}(t) remains constant, while Ω_{DI}(t) has become independent of its initial value. The boundary conditions imposed on the joint propagator are given by:

$$P(\Omega_{\mathrm{LG}}, \Omega_{\mathrm{DI}}, t|\Omega^0_{\mathrm{LG}}, \Omega^0_{\mathrm{DI}}) = \begin{cases} \delta(\Omega_{\mathrm{LG}} - \Omega^0_{\mathrm{LG}})P(\Omega_{\mathrm{DI}}, t|\Omega^0_{\mathrm{DI}}), & t < t^* \\ P(\Omega_{\mathrm{LG}}, t|\Omega^0_{\mathrm{LG}})P_{eq}(\Omega_{\mathrm{DI}}), & t > t^* \end{cases} \tag{3.20}$$

which means that on short timescales, as long as the stochastic variable Ω_{LG}(t) does not change, the motion is governed almost entirely by the internal motion propagator defined as P(Ω_{DI},t|Ω^0_{DI}). At longer times, when Ω_{DI}(t) is assumed to reach its equilibrium value, the overall motion propagator governs the stochastic reorientation process, which defines condition for statistical independence of the stochastic processes commonly referred to as adiabatic approximation. In other words, the time-scale separation implies statistical independence. Practically, from the point of view of Model-Free, one begins considerations by splitting spin-lattice interaction tensor in laboratory frame in two parts:

$$V_0^{(L)}(t) = V_{0,s}^{(L)}(t) + V_{0,f}^{(L)}(t) \tag{3.21}$$

$$V_{0,s}^{(L)}(t) = V_0^{(L)}(t) - \overline{V_{0,f}^{(L)}(t)} \tag{3.22}$$

with overbar indicating averaging of stochastic processes associated with reorientation of polypeptdie fragments limited by internal degrees of freedom (Ω_{DI}). Substitution of equation 3.21 to equation 3.1 leads to a combination of two auto and two cross-correlated correlation functions:

$$C(t) = G_{ss}(t) + G_{sf}(t) + G_{fs}(t) + G_{ff}(t) \tag{3.23}$$

Adiabatic approximation requires that the reorientation rate associated with internal motion is much faster in comparison with the one exhibited by overall reorientation (global motion) so that $V_{0,s}^{(L)}$(t) and $V_{0,f}^{(L)}$(t) fluctuate on distinct timescales, which simplifies the correlation function to:

$$C(t) = G_{ss}(t) + G_{ff}(t) \tag{3.24}$$

$$G_{ff}(t) = 5\left\langle V_{0,f}^{(L)} * (0) V_{0,f}^{(L)}(t) \right\rangle \tag{3.25}$$

$$G_{ss}(t) = 5\left\langle \left[V_0^{(L)}(0) - \overline{V_{0,f}^{(L)}(0)}\right]^* \left[V_0^{(L)}(t) - \overline{V_{0,f}^{(L)}(t)}\right] \right\rangle \tag{3.26}$$

which means that fluctuations of spin-lattice interaction tensor represented by reorientation of global mobility tensor about the average value of the internal motion tensor yield component of the global motion auto-correlation function. In that case, the entire correlation function may be represented as a sum:

$$C(t) = S^2 C_{global}(t) + (1 - S^2) C_{internal}(t) \tag{3.27}$$

The representation of reorientation of polypeptide fragments following decomposition of correlation function into overall global rotation and local internal motion may, however, lead to underestimation of faster internal dynamics representing average over timescales spanning tens of ps to single ns. Further extension of the analysis, to model assuming three reorientation processes has been proposed relying on superposition approximation of overall rotation defined by G tensor with two statistically independent local motions allowing to better represent experimentally accessible ^{15}N relaxation rates (78, 123). A variation of the approach may be proposed considering three statistically independent reorientation processes. In that case, additional frames should be considered as an expansion of the set (LGDIF) into (LGDID'I'F), where additional two primed frames D' and I' are associated with faster internal motions, while unprimed D and I are reference point for slower internal motion. Analogical considerations as in equations 3.21 to 3.27 apply with additional decomposition of $V_{0,f}^{(L)}(t)$, leading to:

$$V_0^{(L)}(t) = V_{0,f}^{(L)}(t) + V_{0,s}^{(L)}(t) + V_{0,t}^{(L)}(t) \tag{3.28}$$

$$V_{0,s}^{(L)}(t) = V_0^{(L)}(t) - \left\langle V_0^{(L)}(t) \right\rangle_{\mathrm{D'I'}} = V_0^{(L)}(t) - \overline{V_{0,f}^{(L)}}(t) \tag{3.29}$$

$$V_{0,t}^{(L)}(t) = V_0^{(L)}(t) - \left\langle V_0^{(L)}(t) \right\rangle_{\mathrm{DI}} = V_0^{(L)}(t) - \overline{V_{0,s}^{(L)}}(t) \tag{3.30}$$

Similarly to considerations in equations 3.21 - 3.22, $\mathrm{V}_p^{(L)}(\mathrm{t})$ is decomposed within bounds of adiabatic approximation, so that cross-correlated correlation functions vanish and $\mathrm{G}_{internal}$ is a simple sum of two internal reorientation processes:

$$C(t) = G_f(t) + G_s(t) + G_t(t) \tag{3.31}$$

The adiabatic approximation requires that the motions occur on separate timescales, which allows to skip the cross-correlated correlation functions.

$$G_f(t) = 5\left\langle V_{0f}^{(L)*}(0)V_{0f}^{(L)}(t)\right\rangle \tag{3.32}$$

$$G_s(t) = 5\left\langle \left[V_0^{(L)}(0) - \overline{V_{0f}^{(L)}}(0)\right]^* \left[V_0^{(L)}(t) - \overline{V_{0f}^{(L)}}(t)\right]\right\rangle \tag{3.33}$$

$$G_t(t) = 5\left\langle \left[V_0^{(L)}(0) - \overline{V_{0s}^{(L)}}(0)\right]^* \left[V_0^{(L)}(t) - \overline{V_{0s}^{(L)}}(t)\right]\right\rangle \tag{3.34}$$

The subsequent partial correlation functions $\mathrm{G}(t)$ describe evolution of spatial parts of spin-lattice interaction tensor associated with processes occurring on different decades of timescales. Without going into too thorough considerations, the correlation function is represented as a sum of correlation functions:

$$C(t) = S^2 C_{slow}(t) + (S_\mathrm{f}^2 - S^2)C_{inter}(t) + (1 - S_\mathrm{f}^2)C_{fast}(t) \tag{3.35}$$

The exact derivation of order parameters within the framework of Wigner rotation matrices for multiple timescale-separated reorientation processes requires different sets of assumptions and is beyond the scope of this work. The Fourier transform of the correlation function yields SDF determining probability of observing motions at different frequencies associated with allowed spin transitions:

$$J(\omega) = S^2 J(\omega) + (S_\mathrm{f}^2 - S^2)J(\omega) + (1 - S_\mathrm{f}^2)J(\omega) \tag{3.36}$$

Further assumptions on distribution of timescales of motions exhibited by the reorienting polypeptide fragments lead to additional averaging steps further increasing complexity of the model and limiting application of adiabatic approximation requiring timescale separation. Assuming exponentially decaying correlation function, a minimum timescale separation (ratio between modes of proximal distributions of correlation times) should be on the order of four to five for reasonably compact distributions. Additionally, the interpretation of the result of classical and extended Model-Free analysis is hampered by sensitivity of experimental measurements to different decades of timescales, necessitating supplementation of the analysis with alternative approaches for accurate and precise description of distributions of timescales underlying estimated dynamic modes.

3.1.2 Alternative approaches in modeling of NH-backbone dynamics

The NMR relaxation rates are proportional to linear combination of SDFs (SDF, Fourier transform of correlation functions of the spatial part of spin-lattice interaction tensor), which determine the probability of observing stochastically modulated motions occurring at different frequencies associated with allowed spin transitions. Thus, SDFs determine sensitivity of relaxation rates to different decades of correlation times representing reorientation rates of (macro)molecule fragments, remaining within range of tens of ps to tens of ns with maxima located at inverse of Larmor frequency of the interacting nuclei. Experimental studies of protein reorientation using NMR relaxation rates as a source of information are limited by loss of memory of initial spatial distribution of partial nuclear magnetic moments encoded within the correlation function. Transverse relaxation rate is dominated by SDF at zero frequency, while SDFs at Larmor frequencies of ^{1}H and ^{15}N determine experimental sensitivity of heteronuclear nOes and R_1 with values of tens to hundreds of ps and around single ns respectively. Therefore, the relative sensitivity of relaxation rates to different decades of timescales of motions depends directly on SDF at different timescales and frequencies associated with allowed spin transitions in relation to SDF at maximum probability, i.e. at timescale equal to inverse of linear combination of Larmor frequencies of interacting nuclei:

$$L(\tau,\omega) = \frac{J(\tau,\omega)}{J(\omega^{-1},\omega)} = \frac{2\tau\omega}{1+\tau^2\omega^2} \tag{3.37}$$

In magnetic filed strength ranging from 14.1 to 22.3T, the experimental sensitivity of relaxation rate experiments ranges from 100 to 500 ps for nOe experiments with maximum sensitivity to fast timescales of motions centered at 150 ps for 1046 MHz ($\omega_{\mathrm{H}} - \omega_{\mathrm{N}}$ at 22.3T) and 250 ps centered at 539 MHz ($\omega_{\mathrm{H}}+\omega_{\mathrm{N}}$ at 14.1T). The sensitivity to intermediate motions is the highest for correlation times ranging from 1.7 to 2.7 ns. For timescales faster than 500 ps or longer than 12 ns the experimental sensitivity decreases following Lorentzian distribution (figure 3.1). Similarly, the characteristics of SDF on timescales longer than 10-12 ns at magnetic field range above 14T may lead to difficulties associated with accurate estimation of dynamic parameters characterizing reorientation processes occurring on slower timescales.

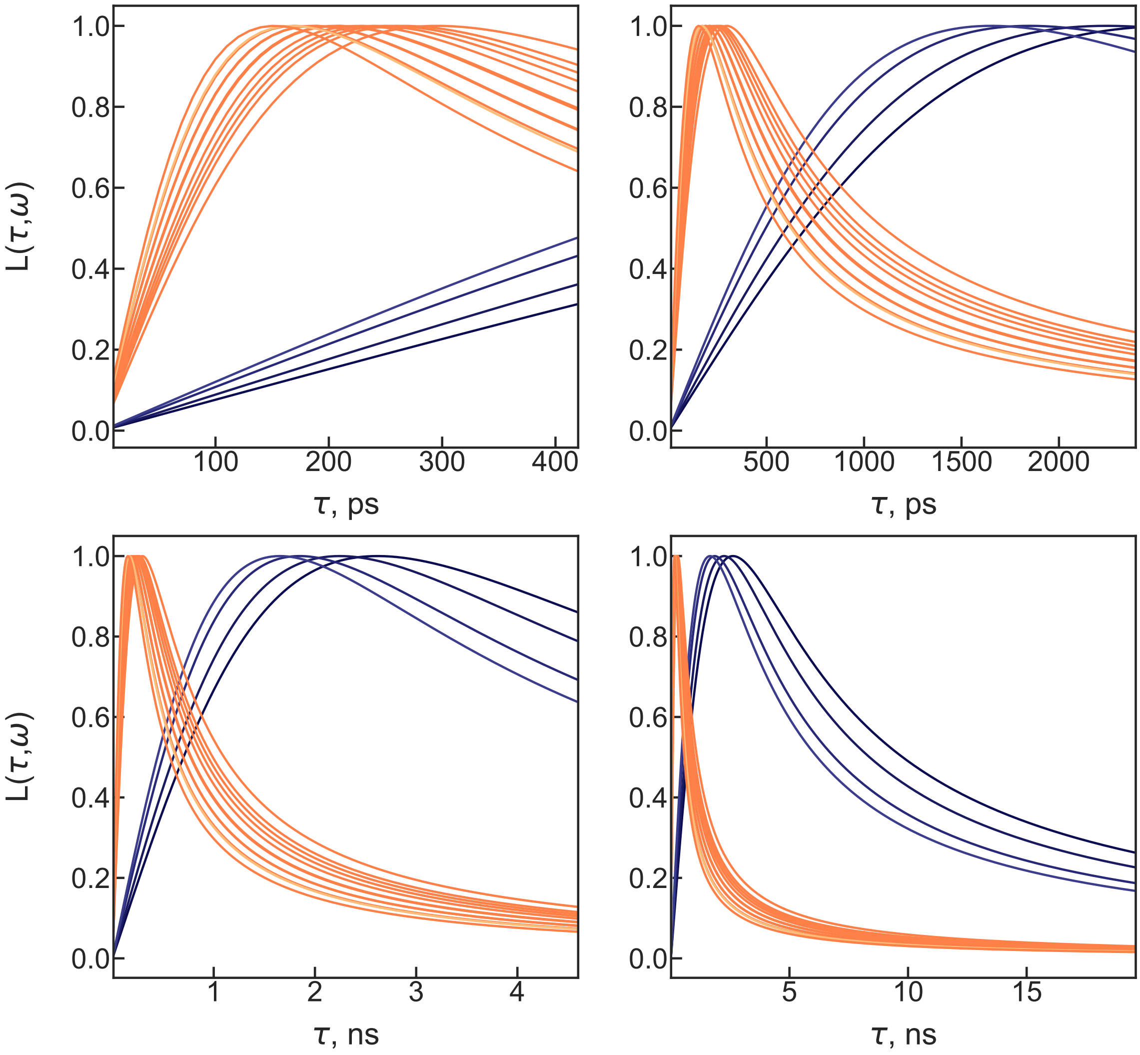

Figure 3.1: Sensitivity of different $J(\omega)$ terms to different timescales of motions at magnetic field strengths ranging from 14.1 to 22.3T defined by value of SDF at different timescale and frequency with respect to value of SDF at timescale of inverse of linear combination of Larmor frequencies associated with spin interaction. Orange and blue lines represent the values of $L(\omega_H)$ and $L(\omega_N)$ respectively.

One of the most accessible and coarse approach in modeling of the backbone dynamics from NMR relaxation rates is so-called reduced spectral density mapping (RSDM). Following derivations of Peng and Wagner, one can solve a linear combination of ^{15}N relaxation rate equations to gauge relative input of SDF components to the relaxation rates, representing distributions of timescales of motions in relation to probability of observing stochastically modulated motions at frequencies associated with spin interactions (196–198). The RSDM of ^{15}N amide relaxation rates leads to the following formulas:

$$\sigma_{\mathrm{NH}} = \frac{\gamma_N}{\gamma_H}(nOe - 1)R_1 \tag{3.38}$$

$$J(0) = \frac{6(R_2 - 0.5R_1 - 0.454\sigma_{\mathrm{NH}})}{3d^2 + 4c^2} \tag{3.39}$$

$$J(\omega_{\mathrm{N}}) = \frac{4(R_1 - 1.4\sigma_{\mathrm{NH}})}{3d^2 + 4c^2} \tag{3.40}$$

$$J(0.87\omega_{\mathrm{H}}) = \frac{4}{5d^2}\sigma_{\mathrm{NH}} \tag{3.41}$$

The estimated SDF components provide first insight into the amount of motion within observation window associated with sensitivity of $\mathrm{J}(\omega)$ at different frequencies associated with spin transitions occurring concurrently to reorientation of peptide units allowing for initial evaluation of correlation times corresponding to conformational fluctuations of polypeptide fragments. Reorientation rates of polypeptide units may be represented by SDFs at different frequencies associated with different spin transitions as a function of SDF at zero frequency, corresponding to the longest correlation time (integral of the correlation function), which within adiabatic approximation is defined by a sum of products of amplitudes and timescales characterizing modes of distributions with respect to experimental sensitivity. Values of SDF at different frequencies evaluated as a function of overall correlation time represent reorientation rate of rigid rotator featuring unimodal distribution of timescales of motions. The comparison of SDF values from RSDM allows to obtain a first insight on dynamics exhibited by macromolecular fragments. Values of SDF within and beyond the region enclosed by $\mathrm{J}(\omega_N)$ as a function of $\mathrm{J}(0)$ point to dominance of intermediate (single ns) and slow (several ns) motions in reorientation processes of polypeptide units respectively (199, 200).

Another approach in modeling of unfolded protein dynamics is Interpretation of Motions by Projection on an Array of Correlation Times (IMPACT), (201). Assuming that the stochastic processes leading to loss of memory function of spatial part of spin-lattice interaction tensor may be represented by a correlation function modeled as a sum of exponential functions with corresponding coefficients (amplitudes) and timescales of motions, one may represent backbone motions by projecting SDF components on an array of logarithmically spaced correlation times typically sampled by the reorienting polypeptide fragment (ranging from tens of ps to tens of ns). A number of Lorentzian contributions is optimized as a function of the longest correlation time τ_{max}, which is in turn connected to the separation between the timescales α:

$$\alpha = (\tau_{min}/\tau_{max})^{\frac{1}{n-1}} \tag{3.42}$$

with n typically oscillating between 5 and 6, so that the ratio between proximal timescales of motions is on the order of 4 to 5 times satisfying adiabatic approximation. In the analysis, one optimizes amplitudes of motions with respect to predefined timescales of motions exhibited by polypeptide units, which represent relative input of polypeptide fragment reorientation rates to SDF components with assumed correlation times yielding an overall idea on distribution of timescales of motions. If sensitivity of J(ω) to different decades of timescales represented by L(ω) is comparable and oscillates within 0.9-1.0, the resultant amplitude profile should directly inform on distribution of timescales represented by a sum of exponentially decaying components. The analysis has convincingly shown, that NH-backbone dynamics in folded region of Engrailed 2 is dominated by longer timescales of motions (above several ns), while in unfolded region reorientation dynamics is dominated by motions occurring on timescales with distributions spanning range of tens of ps and single ns.

Alternatively, by considering that SDFs represent probability of observing motions at different frequencies, one may determine relative contributions of dynamic processes represented by SDFs to relaxation rates at the maximum sensitivity of the latter (202). In the spirit of assumption that each relaxation rate experiment is dominated by SDF at frequency at which spin interactions for certain system occur (0, ω_N, ω_H for ^{15}N-^{1}H spin pair), one imposes that correlation times of internal motions are defined as inverse of linear combination of Larmor frequencies of interacting nuclei (which approximately corresponds to Larmor frequencies of the nuclei), so that relaxation rates are described by modes to which the experiment is the most sensitive to:

$$J(\omega_i) = \frac{2}{5}S_f^2\left(S_H^2 S_N^2 J(\omega_i,\tau_c) + (1-S_H^2)J(\omega_i,\tau_H) + S_H^2(1-S_N^2)J(\omega_i,\tau_N)\right) \tag{3.43}$$

$$\text{where } J(\omega_i,\tau_k) = \frac{\tau_k}{1+\omega_i^2\tau_k^2};\ \tau_H = \frac{1}{\omega_H+\omega_N};\ \tau_N = -\frac{1}{\omega_N} \tag{3.44}$$

Sacrificing optimization of timescales of internal motions allows to reduce the number of optimized parameters and more precisely determine the order parameters. Following this approach, one may obtain an insight on distribution of timescales of motions per relaxation rate experiment (202). If the stochastic processes may be represented by a sum or a product of correlation functions with average decay rates equal to the inverse of Larmor frequencies of interacting nuclei, the order parameters determined from the analysis represent relative populations of conformational transitions reflecting loss of memory function due to the reorientation process. If distributions of correlation times representing conformational fluctuations feature modes at timescales distant from inverse frequencies associated with spin interaction (Larmor frequencies or their linear combination), estimated populations may be influenced by relative proximity of modes to timescales for which maximum sensitivity of relaxation rate experiment is achieved. For instance, dominance of slower reorientation processes with distribution of correlation times at timescales beyond sensitivity window of relaxation rate experiment described by normalized SDF may lead to underestimation of population of distribution describing slower motions, in turn influencing estimated values of populations of remaining distributions by virtue of normalization of amplitudes of motions.

Instead of focusing on single values of timescales of motions corresponding to the mode or the maximum probability of observing spin transition due to the reorientation process, one may describe the polypeptide dynamics from NMR relaxation rates following estimation of distribution of correlation times. Modig and Poulsen proposed to evaluate statistical properties of distributions estimated from their central moments (203). In the approach, one decomposes information encoded within relaxation rates represented by SDFs (i.e. Fourier transform of correlation functions) into distribution of correlation times and estimates their central moments: mean, standard deviation, skewness, $\tau_{1/2}$ (time after which correlation function decays to a half) and λ (ratio between integrals of latter part of correlation function from $\tau_{1/2}$ to ∞, and entire correlation function). The approach has been applied to bovine acyl-coenzyme A binding protein (ACBP), and allowed to distinguish sequence specific features of distribution of timescales of motions. Similarly, application of Lorentzian distribution of correlation times to model reorientation processes of polypeptide fragments allowed to represent dynamic modes of NH-backbone of D2 annexin I as a continuous distribution instead of discrete values with a mode at single ns timescales, which value coincides with timescale of intermediate mode estimated from extended Model-Free analysis. The slowest timescales of motions observed in the Model-Free analysis are found in slower tail of the distribution. Faster timescales are represented by a long and relatively flat contribution on a range of timescales from ps to hundreds of ps. The apparent variance of the distribution shows minor correlation with estimated from independent extended Model-Free analysis order parameters. The comparison suggests that order parameters are representative of parameters characterizing distribution of motions in relation to sensitivity of SDF (204).

An ambitious task proposed by Arthur G. Palmer III and co-workers is to represent an entire distribution of timescales in a form of inverse Gaussian distribution of timescales (205). In general, the distribution is defined as:

$$p(\tau) = \left(\frac{\mu^3}{2\pi\sigma^2\tau^3}\right)^{1/2} \exp\left(-\frac{\mu(\tau-\mu)^2}{2\sigma^2\tau}\right) \tag{3.45}$$

where μ and σ define mean and standard deviation of distribution of timescales $\mathrm{p}(\tau)$. The Fourier transformed distribution function allows to directly encode moments of distribution of timescales into SDF, and provides more direct insight into dynamics of the system of interest. The approach has been successfully applied to the analysis of NH-backbone dynamics of GCN4 with quality of reproduction of relaxation rates indistinguishable from extended Model-Free formalism. In the analyzed example, the identified mean of the distribution corresponds to value identified as slower internal correlation time (around single ns) with a mode in the range of quarter of value suggesting an asymmetric distribution skewed towards longer timescales. The slow ns dynamics featured by the protein oscillate within several to tens of ns with smaller values featured by residues located in the termini of protein and characteristic profile coinciding with S^2 values.

The moments of distributions of correlation times characterizing reorientation rates of fragments of macromolecules, their variances, shape (represented by skewness) and composition in terms of reorientation processes exhibited by peptide units and segments thereof resultant from stochastically modulated motions are still under debate. In order to evaluate them, a quantitative framework relating sensitivity of the experiment to the timescales associated with underlying reorientation processes is needed. The efforts in that vein have been made by Meier, Ernst, Ferrage and Smith, who proposed detectors of dynamics approach. It relies on estimation of relative sensitivity of combination of relaxation rates to an arbitrary distributions of timescales of motions. In the approach, one defines a detector of dynamics, which is characterized by sensitivity and its response (206–209). By taking into an account that internal motions of peptide plane occur independently from global reorientation of a protein, one may estimate the underlying distribution of timescales of internal motions, by finding respective detector sensitivities and their responses (eigenvectors and eigenvalues of the relaxation rate matrix). The detector sensitivity to internal motions is defined as a linear combination of relaxation rate sensitivities at given magnetic field strength ($\mathrm{R}_k(\mathrm{z})$):

$$\rho^{(\theta,S)} = (1 - S^2) \int_{-\infty}^{\infty} \theta(z) \left(\sum_k a_k R_k(z) \right) dz \tag{3.46}$$

where $\theta(\mathrm{z})$ is a distribution of correlation times, and $\mathrm{R}_k(z)$ is defined by a linear combination of SDFs representing sensitivity of the experiment to different decades of timescales of motions. In that case, one may define a relaxation rate as:

$$R^{(\theta,S)} = S^2 R(z) + (1 - S^2) \int_{-\infty}^{\infty} \theta(z) R(z_{\text{eff}}(z)) dz \tag{3.47}$$

of which $\mathrm{R}_k(\mathrm{z})$ or $\mathrm{R}_k(\mathrm{z}_{\text{eff}}(\mathrm{z}))$ is assumed to be characterized by a mono-exponential decay with correlation time $\tau_c = 10^z$, s or $\tau_{\text{eff}} = (\tau_i \tau_c)/(\tau_i + \tau_c)$ respectively. The detectors response is then equal to linear combination of relaxation rate constants:

$$\rho_n^{(\theta,S)} = \sum_k a_k R_k^{(\theta,S)} \tag{3.48}$$

In practice, the approach leads to optimization of relaxation rate matrix to find the detectors sensitivity and response in relation to underlying distributions of timescales of motions. It allows to gather a deeper insight into sensitivity of experimental measurements to an arbitrary distribution of timescales of motions, to which a given combination of relaxation rates is sensitive to rather than focusing on sensitivity of particular components of SDF. Originally applied to analysis of dynamics from NMR relaxation data in solid-state (206, 207), the approach has found its application in solution state NMR (208, 209). The potential issues associated with interpretation of dynamics from detectors approach are related to finite variance of distributions of timescales in relation to the detectors sensitivity. In pursuit of maximizing relative sensitivity window and number of detectors obtained by normalization of relaxation rates sensitivities at different decades of timescales, one may obtain redundant information on underlying distribution of timescales of motions as a result of the detector overlap. For instance, distribution of timescales characterizing the detector sensitivity to timescales of motions occurring on one decade (i.e. centered at 150 ps) will be detected majorly by detector spanning timescale window of 20-200 ps with small degree of contribution to the response of the neighboring one (i.e. 150-750 ps). The resultant redundancy of detectors may lead to difficulties with determination of underlying distribution of timescales of motions necessitating further improvements in the definition of detectors sensitivity to experimentally accessible relaxation rates.

3.2 Methods

Model-Free analysis, numerical simulations

Model-Free analysis of NMR relaxation rates estimated from simulated arbitrary distributions of correlation times representing reorientation rates of polypeptide fragments have been performed assuming sets of Gaussian-shaped distributions centered at tens of ps (fast, 10 - 200 ps), 1 ns (intermediate), and several ns (slow, 1.5 to 30 ns). The populations of distributions were either equal (A_{fast}:A_{inter}:A_{slow} = 1:1:1) or with population of one of the distribution twice greater than the other two (1:2:1 or 1:1:2). Variances of the distributions (Ω) were set to 0.1, 0.3, 0.6 or 0.9 of distribution mode value (e.g. for intermediate distribution centered at 1 ns, variance of 0.3 determines full-width at half maximum of 300 ps). Each distribution was divided in 33 bins, adding up to 99 timescales and amplitudes of motions from which average relaxation rates was calculated assuming a simple sum of correlation times with amplitudes corresponding to the population for magnetic field strength ranging from 14.1 to 22.3T. For averaging purposes, relaxation rates estimated for each set of distributions were estimated 100 times with Gaussian noise contribution within 2% of the relaxation rate value. Model-Free analysis of the relaxation rates was performed assuming three statistically independent components (adiabatic approximation).

MD simulations of unfolded proteins

MD simulations of unfolded proteins have been performed with TIP4P-2005 water model and RMSFF2+ and Charmm36* force field at constant temperature of 278 and 298K for MKK4 and NTail respectively. Starting conformations for simulations were chosen from ASTEROIDS-ensemble on the basis of experimentally accessible RDCs and chemical shifts. Equilibration steps involved NVT and NPT or NVE and NPT simulations spanning 100 ps and 1 ns respectively followed by 100 and 200 ns NVE and NVT trajectories for MKK4 and NTail respectively. Correlation functions of rank 2 spherical harmonics were estimated from 50 and 100 ns blocks (starting at 0 nd 50 or 0, 50, and 100 ns). A cutoff of 9Å has been used for the Lennard-Jones and electrostatic interactions, while smoothing functions have been applied to electrostatic and Van der Waals forces beyond 7.5Å for NTail and MKK4 respectively. To calculate long-range electrostatic interactions, particle-mesh Ewald summation with a grid spacing of 1Å has been used. A Langevin thermostat with damping constant of 0.1 ps was used to keep the temperature of the system at 298K. During the NPT equilibration a Berendsen barostat with a relaxation time of 400 fs was used to control the pressure at target value of 1.01325 Bar. The dynamic modes derived from MD trajectories result from fit of average auto correlation functions (ACFs) of rank 2 spherical harmonics of all of the trajectories composed of 24 and 54, 50 and 100 ns blocks for MKK4 and NTail respectively. The ACFs have been estimated following Tikhonov regularization of ACFs projected on a grid of 128 timescales logaritmically spaced timescales between 1 ps and 25 ns. The MD simulations have been performed by Dr. Nicola Salvi (138–140).

Experimentally accessible dynamics from NMR relaxation rates

The intrinsically disordered regulatory domain of MKK4 (residues 1-86, UniProtKB accession number P45985) and C-terminal domain of Sendai Virus Nucleoprotein (residues 401-525 UniProtKB accession number Q07097, here referred as residues 1-125) has been expressed and purified as described previously (68, 70). Samples for NMR relaxation measurements have been prepared in Hepes buffer equilibrated at pH 7.0 with 150 mM NaCl and 10% D_2O. Data for modeling of NH-backbone dynamics was recorded on 4 spectrometers (magnetic field ranging from 14.1 to 22.4T) and 273K. The ^{15}N auto-relaxation experiments have been performed with HSQC-type pulse sequences with water suppression achieved with Echo-Antiecho detection scheme (210). The delays for R_1 and $R_1\rho$ experiments ranged from 40 ms to 1.4 s and 1 to 200 ms respectively. The spin-lock field strength in ^{15}N $R_1\rho$ experiment has been adjusted to approximately 2.0 kHz with adiabatic half-passage hyperbolic tangent pulses with 3 ms ramps and temperature compensation of up to 300 ms. Amide proton decoupling during relaxation decay of R_1 experiment was achieved using iBurp pulse of 2 ms centered at 8.5 ppm. Heteronuclear nOe measurements were performed with saturation and recovery periods of 6-8 s. ^{15}N CSA-dipole-dipole interference experiments (η_{xy} and η_z) have been performed with Echo-Antiecho detection scheme with delays of 60 and 100 ms for transverse and longitudinal cross-correlated cross-relaxation respectively (195, 211, 212). In all of the relaxation experiments the interscan delay was set to 1.5-2.5 s. Model-Free analysis have been performed with two or three apparent dynamic modes. Constant value of CSA-DD angle of 25°, fast timescales of motions equal to 45 ps, amide nitrogen-proton bond vector length was set to 1.015Å and ^{15}N CSA magnitude of -172 ppm has been assumed. The optimization of the parameters has been performed with Levenberg-Marquardt algorithm.

3.3 Distributions of timescales of motions, numerical simulations

To evaluate limits of statistical independence of correlation times associated with conformational averaging within limits imposed by adiabatic approximation implicit in Model-Free analysis of ^{15}N relaxation rates applied for estimation of apparent dynamic modes from experimental data, a set of numerical simulations of NMR relaxation rates was performed. For demonstration purposes, simulated data constituted sets of three Gaussian-shaped distributions of timescales of motions centered at tens of ps (fast, 10 to 200 ps), 1 ns (intermediate), and a few ns (slow, 1.5 to 30 ns). In the first set of examples, well defined distributions of timescales of motions with equal populations were considered ($\Omega = 0.1$, figure 3.2). As expected, when timescales of motion are well separated - ratio between proximal modes ($\langle\tau_i\rangle$ and $\langle\tau_{i+1}\rangle$) is at least 5, the average timescales estimated from Model-Free analysis of relaxation rates yields values which are in close comparison to modes of underlying distributions and amplitudes correspond to their populations, as long as the distributions are within sensitivity of SDFs. The impact on timescales and amplitudes of motions estimated from relaxation rates is apparent, when modes of proximal distributions are roughly within ratio of 2. The most severe example may be observed when all three of the distributions overlap - information on intermediate mode is represented by parameters characterizing neighboring distribution with resultant amplitude in a range twice greater to initially assumed average value, while estimated timescales of motion are situated on outer borders of the distributions, so that the estimated apparent modes become as much separated as possible, as it is allowed by sensitivity of $\mathrm{J}(\omega_{\mathrm{N}})$ and $\mathrm{J}(\omega_{\mathrm{H}})$. As a result, the estimated amplitudes of motions are influenced leading to representation of a part of 'timescale information' by amplitude of motion so that estimate of the relaxation rate is the most accurate. Therefore, even though the original distribution of timescales did not posses τs in a certain range, the Model-Free averaging of timescales of motions allows for them. As long as ratio between proximal modes of the distributions ($\langle\tau_i\rangle$ and $\langle\tau_{i+1}\rangle$) is at least five and the distributions do not overlap significantly ($\Omega \leq 0.3$), influence on the estimated values of amplitudes and timescales of motions is very small (figures App.1 - App.11, appendix). For intermediate motions centered at single ns, the proximity of modes of underlying distributions leads to similar values of transverse and longitudinal relaxation rates in a manner that depends on the populations.

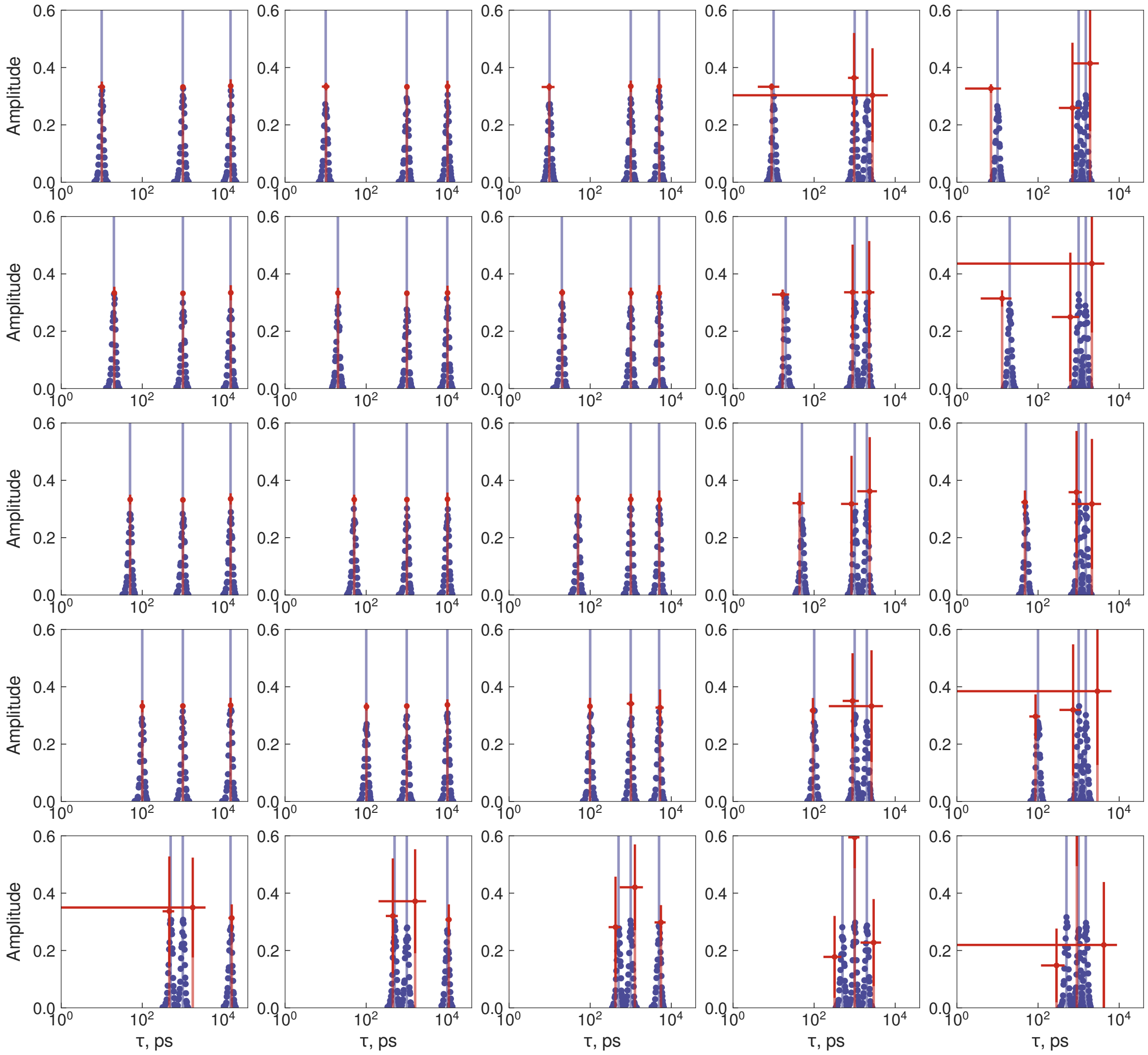

Figure 3.2: Comparison of simulated values of distributions of timescales of motions used for calculation of relaxation rates with values estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated ones (Population ratio: 1:1:1, $\Omega = 0.1$). Blue points represent distribution of simulated correlation times for each motion with average represented by blue bar. Orange points and bars represent average and standard deviation of estimated apparent dynamic modes on the basis of Model-Free analysis of relaxation rates calculated from simulated distributions of correlation times.

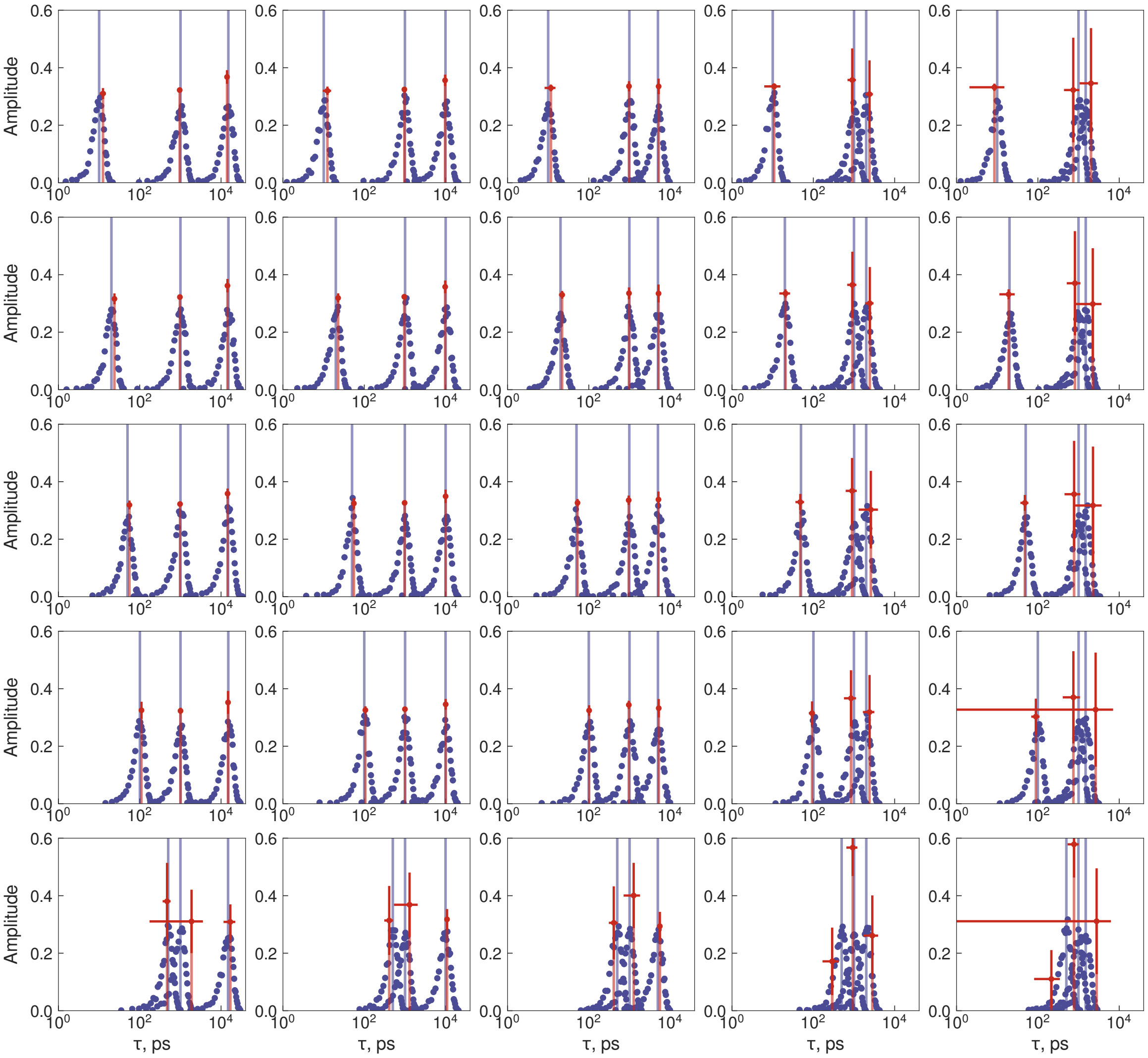

Figure 3.3: Comparison of simulated values of distributions of timescales of motions used for calculation of relaxation rates with values estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated ones (Population ratio: 1:1:1, Ω = 0.3). Blue points represent distribution of simulated correlation times for each motion with average represented by blue bar. Orange points and bars represent average and standard deviation of estimated apparent dynamic modes on the basis of Model-Free analysis of relaxation rates calculated from simulated distributions of correlation times.

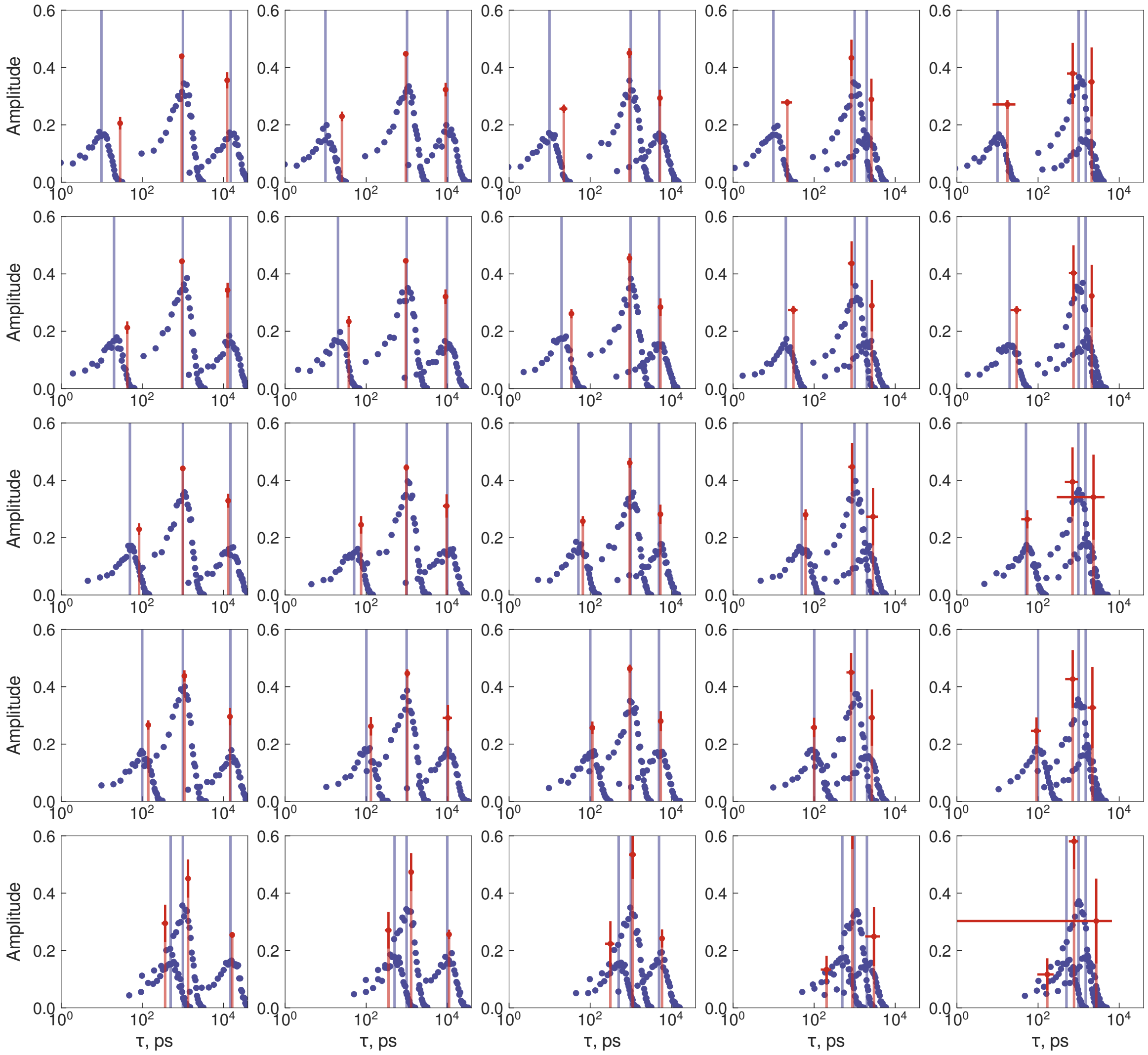

Figure 3.4: Comparison of simulated values of distributions of timescales of motions used for calculation of relaxation rates with values estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated correlation times (Population ratio: 1:2:1, $\Omega = 0.6$). Blue points represent distribution of simulated correlation times for each motion with average represented by blue bar. Orange points and bars represent average and standard deviation of estimated apparent dynamic modes on the basis of Model-Free analysis of relaxation rates calculated from simulated distributions of correlation times.

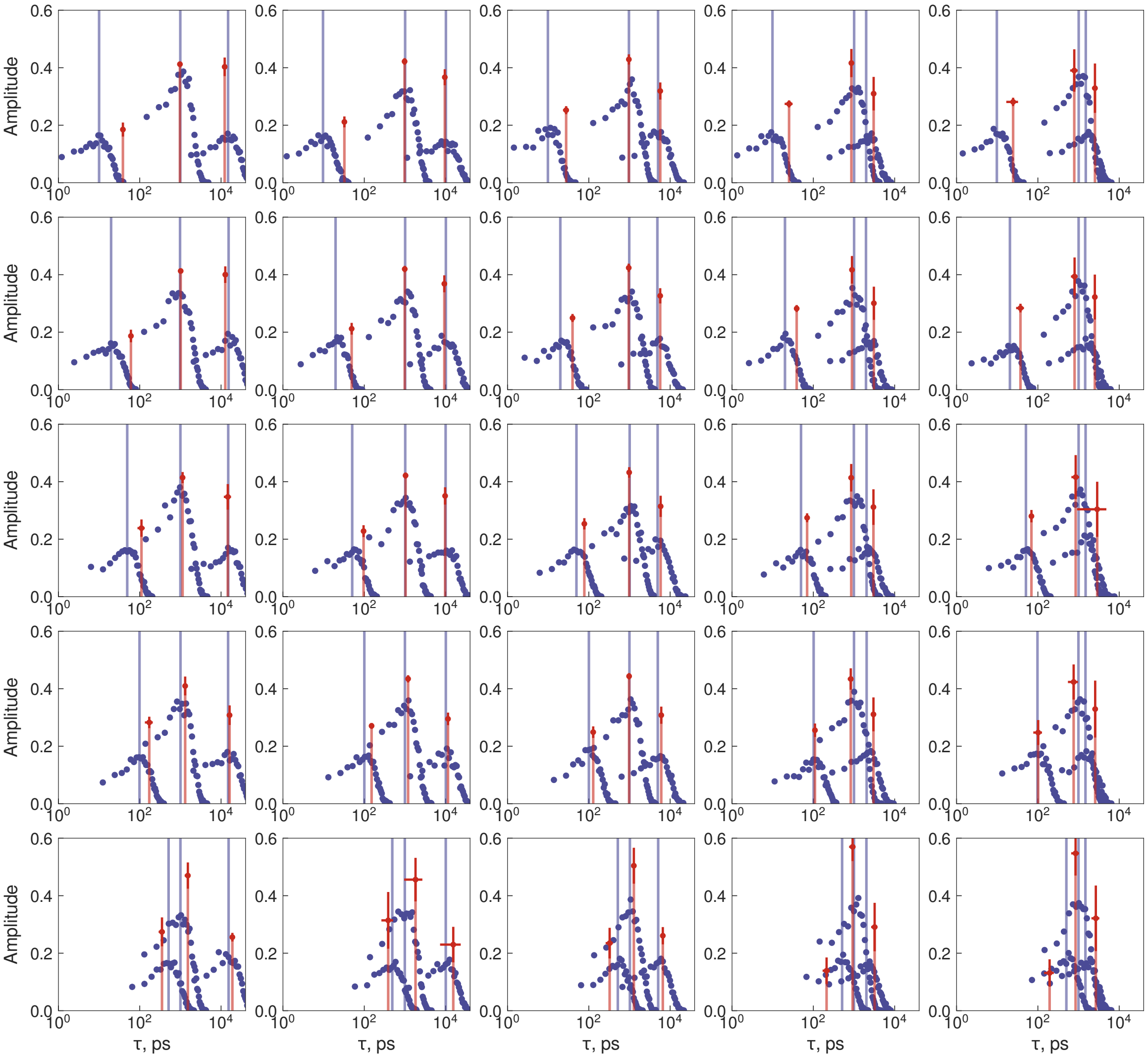

Figure 3.5: Comparison of simulated values of distributions of timescales of motions used for calculation of relaxation rates with values estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated ones (Population ratio: 1:2:1, $\Omega = 0.9$). Blue points represent distribution of simulated correlation times for each motion with average represented by blue bar. Orange points and bars represent average and standard deviation of estimated apparent dynamic modes on the basis of Model-Free analysis of relaxation rates calculated from simulated distributions of correlation times.

For equal populations of correlation times, the estimated values of amplitudes correspond to populations of underlying distributions, as long as timescales of intermediate distribution do not exceed 2.2 ns ($\Omega \leq 0.3$), which corresponds to the average of ω_N^{-1} in the applied magnetic field range. The contribution to amplitudes of slow motions from intermediate motions is apparent, when variance of intermediate distribution reaches 0.6, leading to small (less than 5%) number of correlation times exceeding 2.8 ns ($1.05\omega_N^{-1}$ at 14.1T). In other words, for broad or overlapped distributions of timescales of motions, the estimated amplitudes of slow motions may over-represent their populations as a result of sensitivity of $\mathrm{J}(\omega_\mathrm{N})$. When distribution of timescales is broad ($\Omega = 0.9$) one can think about the distributions of timescales as representative of a quasi-continuum of correlation times rather than separate distributions. The estimated average timescales of motions are remarkably similar to the ones used for calculation of relaxation rates reaffirming the assumption of representation of relaxation rates by a finite number of (apparent) dynamic modes (figure 3.5).

The similarity of information provided from overlapping distributions of correlation times leads to a greater range of allowed values for amplitudes of motions able to reproduce the relaxation rates represented by standard deviation of resultant distributions from noise-based simulation of relaxation rates. A complete set of numerical simulations containing remaining combinations of variance and populations of distributions (fast, intermediate and slow) may be found in appendix. In order to compare accuracy of Model-Free analysis in reconstructing the information on distributions of timescales of motions in a concise manner, one can evaluate changes in the root-mean-square and standard deviation of the estimated timescales of motion as compared to the average timescale of primary distribution as a function of ratio between modes of proximal distributions (figure 3.6). As should be expected, as long as degree of overlap between the distributions and sensitivity of relaxation rates (represented by normalized SDF) is large enough, the larger separation of the modes, the greater degree of accuracy and precision with which the underlying distribution is estimated. The relative RMS and standard deviation (between estimated and underlying average timescales and amplitudes of distributions) changes in an exponential-like manner as a function of separation between the modes with suggestive minimum ratio between proximal modes of underlying distributions of at least 5 for reasonable estimation of timescales. For overlapping distributions of timescales featuring variances greater than 60% value of the mode and for separation between centers of the distributions in range of 5 and above, accuracy with which one can estimate the dynamic modes from relaxation rates decreases, yielding up to 25-50% deviation in average timescale of motion from the mode of primary distribution.

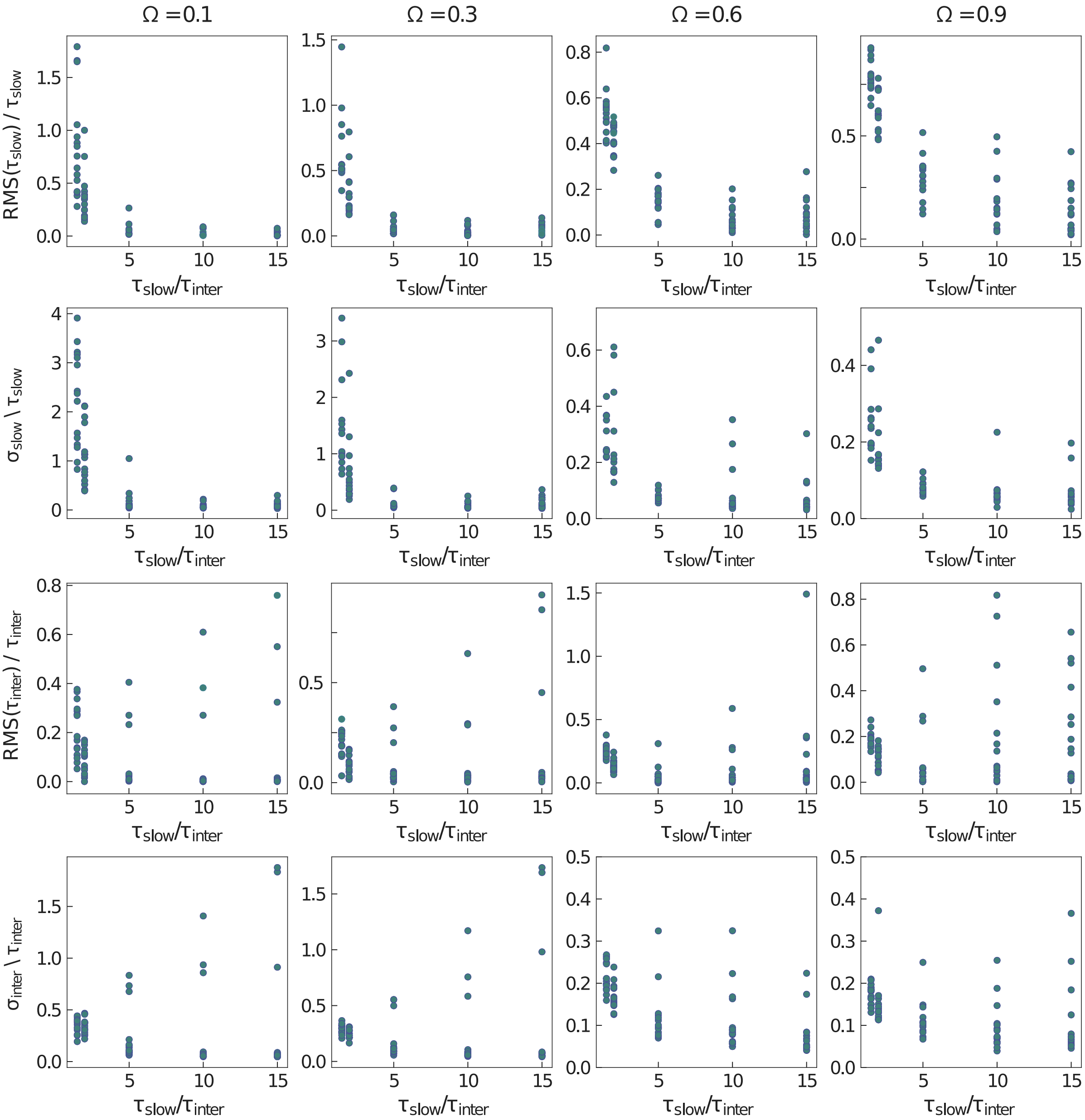

Figure 3.6: Comparison of RMS and standard deviation of distributions of timescales determining accuracy and precision of estimation of apparent dynamic modes obtained from Model-Free analysis of simulated relaxation rates relative to modes of underlying distribution with different variances ranging from 0.1 to 0.9 of the mode value.

From application perspective, it is instrumental to investigate limiting cases associated with estimation of dynamic modes from experimentally accessible relaxation rates. It is however, a non-trivial task to predict issues associated with interpretation of dynamic parameters *a priori*. Typically, a first step to any Model-Free analysis is Reduced Spectral Density Mapping (RSDM), which subjected to a combination of three relaxation rates measured at single magnetic field allows to gauge relative input of motional processes occurring on different decades of timescales to SDF. Although, one typically investigates only absolute values of SDFs, in order to draw meaningful conclusions it is more useful to consider relative input of each of the components to the longest correlation time which may be estimated from NMR relaxation rates dominated by J(0), representing integral of the correlation function of the spatial parts of spin-lattice interaction tensor. Comparison of SDF components with quality of representation of dynamic modes and populations of distributions by average correlation times (apparent dynamic modes) and associated amplitudes from Model-Free analysis of relaxation rates allows to dissect combinations of SDFs, which may be troubling in terms of accuracy of representation of the former as obtained from Model-Free analysis (figure 3.7). It appears, that slow dynamic modes are perturbed whenever intermediate distribution has a major contribution to ^{15}N-SDF ($J(\omega_N)$ to J(0) ratio of at least 0.5) and J(0) values are rather small (approximately single ns). In that case, the entire SDF appears flatter leaving very narrow range of values between J(0) and $J(\omega_N)$ for accurate estimation of remaining apparent dynamic modes. The predicted from numerical simulations relative error in estimation of τ_{slow} values may range from 0.5 to 1.5 with maximum deviations for simulated timescales of motions observed for the set of most compact distributions, i.e. with dominant intermediate motion component and very small contribution of fast and slow motions (high $J(\omega_N)$ to J(0) ratio and low J(0)), visible experimentally in the form of low R_2 and high R_1 values. Upon estimation of dynamic modes from relaxation rates the combination of SDFs leads to a relatively shallow χ^2 surface in parameters space composed of amplitudes and timescales, in turn making it difficult to accurately (and precisely) estimate slow and intermediate modes (figure 3.7). One could expect a similar outcome for intermediate or fast modes, when their contribution to SDF is small - experimentally, this is visible for high R_2 coupled with low R_1 values, and typically high heteronuclear nOes. When overall dynamics of NH-backbone is fast (J(0) values below 500 ps), the distributions of correlation times corresponding to slow and intermediate motions begin to overlap (figure App.5) yielding estimation of apparent slow dynamic mode representing correlation times at boundaries of the distribution (away from the intermediate one). In that case, information about timescale or amplitude of motion may carry a degree of error, which is difficult to estimate (on the order of 50%, figure 3.7).

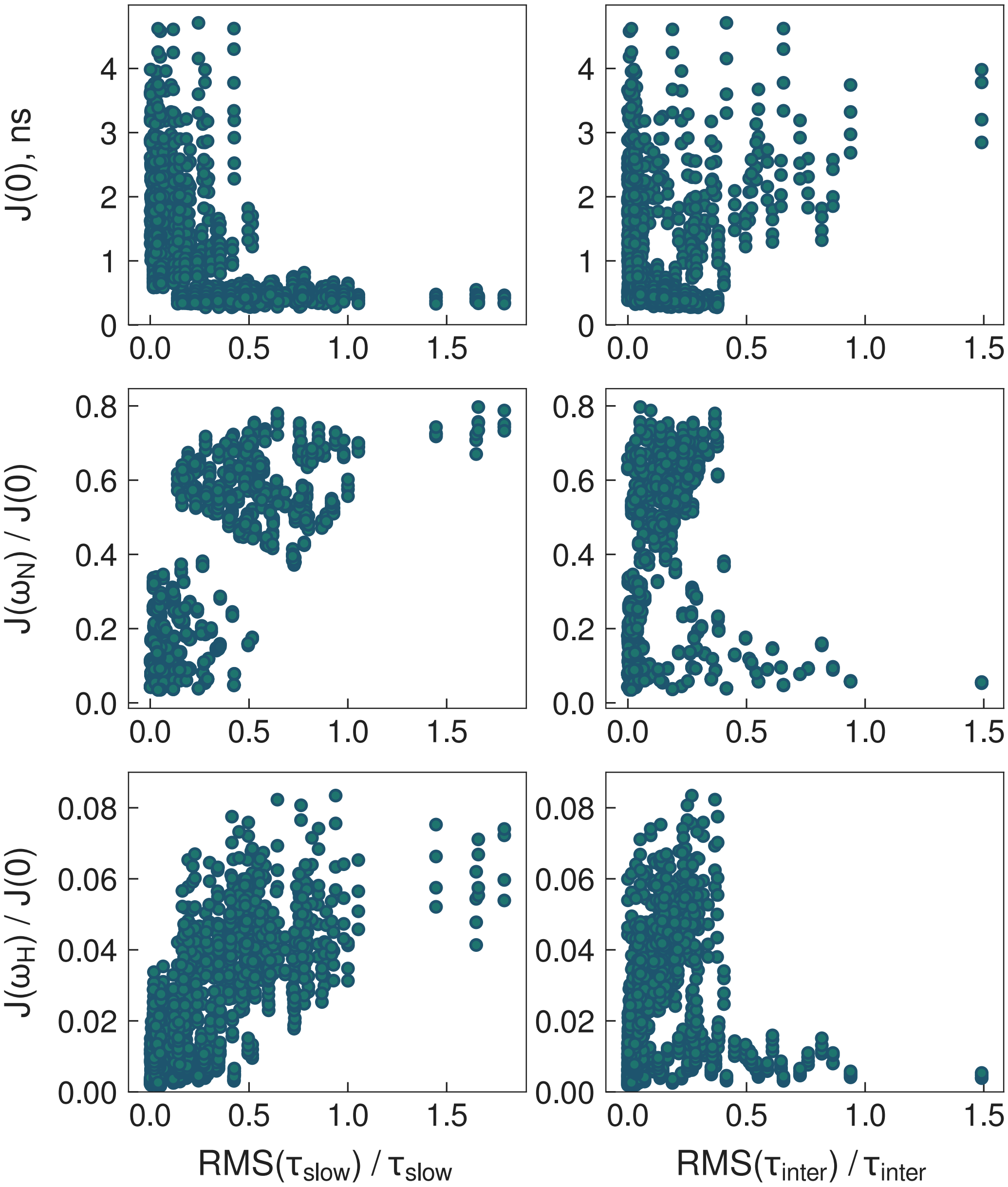

Figure 3.7: Comparison of SDF components estimated for simulated relaxation rates at characteristic frequencies (at magnetic field strength of 14.1T) which dominate values of slow, intermediate and fast motions with relative root-mean square deviations of obtained distributions with respect to modes of originally imposed distribution of timescales.

3.4 From relaxation to dynamic parameters

NMR relaxation rates are a rich source of information on dynamics of the system. In this section, I will present an analysis workflow of relaxation rates leading to interpretation of complex dynamics exhibited by an unfolded protein. I begin my considerations with ^{15}N backbone relaxation rates acquired at magnetic field ranging from 14.1 to 22.3T. Figure 3.8A shows ^{15}N relaxation rates of a model system Map kinase kinase 4 (MKK4) recorded at 273K revealing non-uniform sequence dependence profile. The R_2 values (dominated by J(0)) for residues at the termini are relatively small (approximately 2 s^{-1}), concurrently with low values of R_1 and heteronuclear nOes point to dominance of intermediate motions with population of correlation times representing slow motions increasing gradually. The field-dependence of R_1 suggests that the majority of correlation times representing intermediate motions (single ns) is within window of maximum sensitivity of $J(\omega_N)$ for the N-terminal part of the protein (represented by a relatively small degree of field dependence with slowest intermediate timescales faster than 2.5 ns). The relative increase of nOe values with magnetic field in the N-terminal part of the protein suggests significant influence of hundreds of ps dynamics to NH-backbone fluctuations with significant population of correlation times in a range of 200-300 ps (corresponding to sensitivity maxima of $J(\omega_H)$). The central part of the protein posseses dynamics, which contribution to $J(\omega_N)$ and $J(\omega_H)$ represented by R_1 decreases gradually with magnetic field suggesting greater contribution of slower than 2.5 ns dynamics, or smaller population of hundreds of ps to single ns motions. A more direct insight into conformational fluctuations of the polypeptide fragment may be obtained by performing RSDM, which directly informs on relative population of correlation times at timescales equal to inverse of frequencies associated with nuclear spin interactions, which corroborates initial interpretation of the relaxation rates (196, 198). Comparison of SDF values (figure 3.8B) allows to realize that J(0) is quite small for residues situated in the termini of the protein (residues 1-18, 80-86), hinting potential issues with accurate and precise estimation of the NH-backbone dynamics (figure 3.7). To separate parameters representing effective memory loss due to the stochastic processes and rate at which the processes occur contributing to the $J(\omega)$ values, one may perform Model-Free analysis starting with a model of lowest complexity - featuring smallest number of dynamic modes.

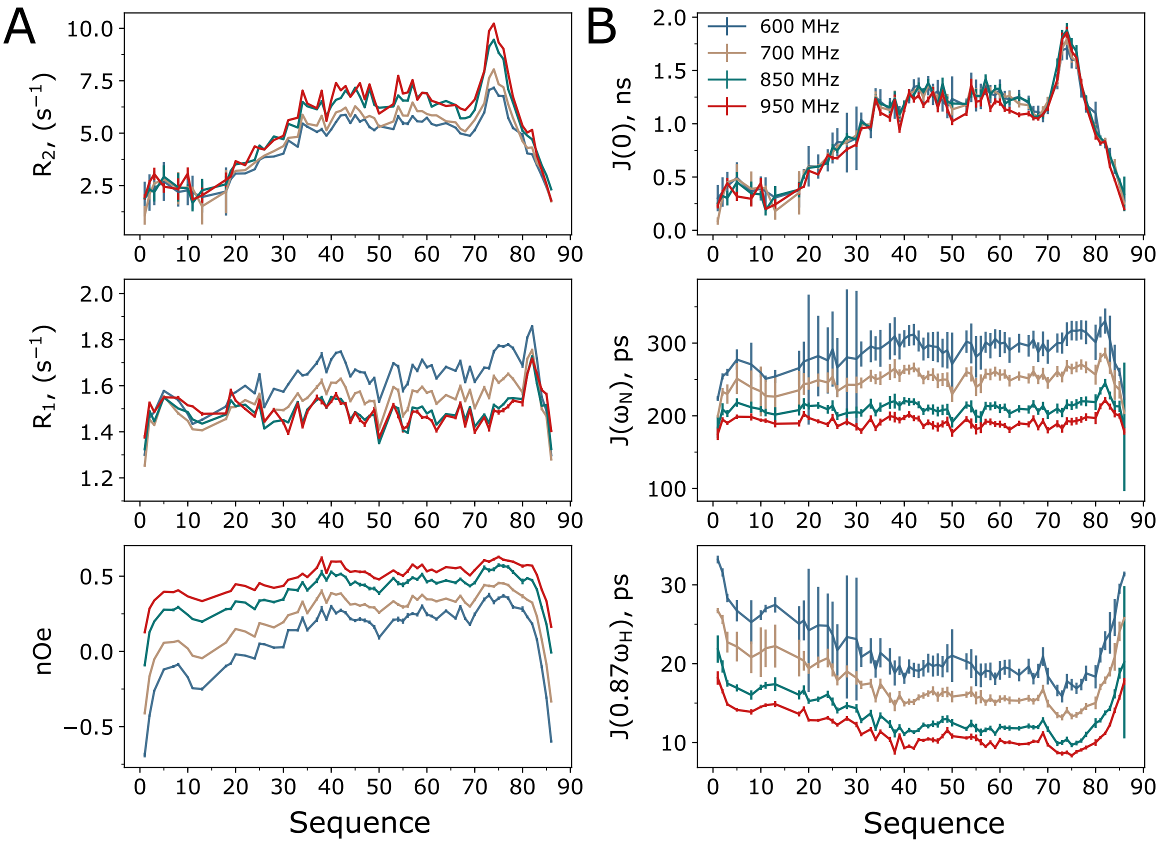

Figure 3.8: Residue resolved ^{15}N *backbone relaxation rates of MKK4 at 273K (A) and associated with it SDF values estimated from RSDM (B) at frequencies associated with spin interactions (B).*

Interpretation of dynamics for residues dominated by fast to intermediate correlation times does not change drastically, since population of the slowest component is close to zero. Model-Free analysis of the relaxation rates with varying number of apparent dynamic modes representing statistically independent timescale separated motions (adiabatic approximation) has been applied. Figure 3.9 shows comparison of reproduction of experimental data with Model-Free analysis assuming two and three apparent dynamic modes. Model-Free analysis with two modes clearly under-represents the relaxation rates for major part of the polypeptide sequence. With an exception of several residues in the N-termini, the estimated from Model-Free analysis values assuming two independent contributions to NH-backbone dynamics do not correspond to the relaxation rates - assumption of third independent reorientation process is necessary.

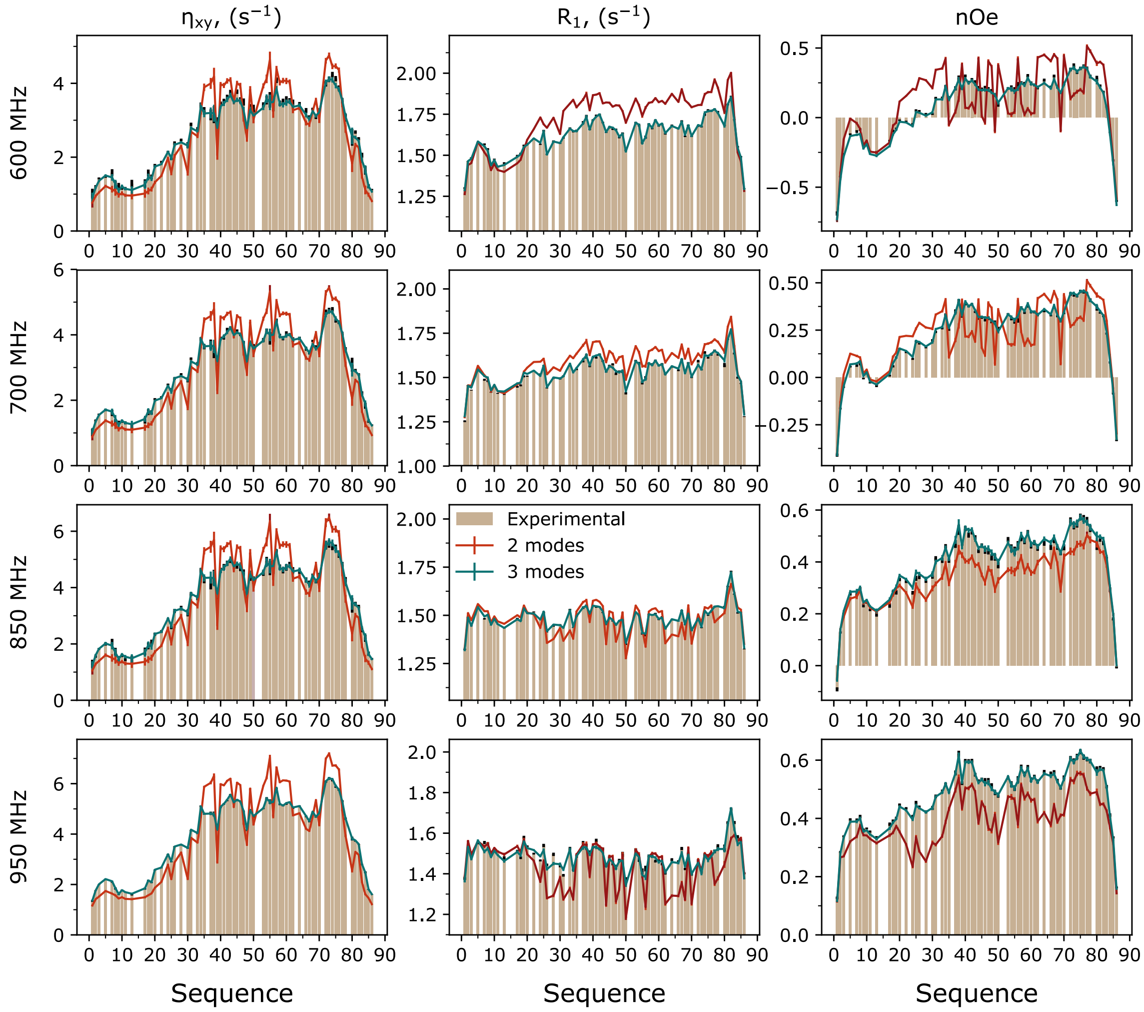

Figure 3.9: Reproduction of the experimental data (bars) using Model-Free analysis assuming two (orange) or three (green) statistically independent, timescale separated reorientation processes.

The transition from Model-Free analysis assuming two to three independent contributions to relaxation rates seems to occur without major differences in reproduction of the relaxation rates describing reorientation dynamics of residues in the ultra-flexible N-terminus, with an exception of additional information on relative population of correlation times in a range five times greater than the average intermediate correlation time. The apparent convergence of the analyses with two and three modes is a direct consequence of (slow and intermediate) distributions overlap represented by relatively small J(0) values. Comparison of experimental and calculated data additionally suggests that distribution of slow motions is relatively broad, effectively contributing to intermediate motions amplitude.

3.5 Distributions of correlation times in unfolded proteins

Reproduction of dynamic modes of distributions of correlation times from numerical simulations spanning tens to hundreds of ps, single ns and several ns reaffirms equivalence of representation of conformational fluctuations in terms of ensemble averaged correlation times, i.e. derived from MD simulations without losses on information content with regard to the NMR relaxation rates, as long as dynamic modes in the underlying distributions may be distinguished. The averaging of experimental relaxation rates informing on distributions of timescales of motions centred at multiple decades of timescales leads to at least three apparent dynamic modes obtained from Model-Free analysis (figure 3.9). Considering an idealized set of relaxation rates calculated directly from a fit of auto-correlation function (ACF) of nitrogen amide spin-lattice tensor fluctuations represented by correlation times, without experimental uncertainties associated with traces of contribution from exchange, sample to sample variations, differences in temperature calibration between spectrometers and spectral noise, Model-Free analysis is limited only by dynamic range (experimental sensitivity) of relaxation rate experiments characterized by $\mathrm{L}(\omega_\mathrm{N})$ and $\mathrm{L}(\omega_\mathrm{H})$, which depends on magnetic field strength at which relaxation rates are acquired. If reorientation of polypeptide units is characterized by a distribution of correlation times with more than three dynamic modes in absence of additional contributions associated with experimental limitations, Model-Free analysis of relaxation rates in magnetic field strength ranging from 14.1 to 22.3T should lead to accurate estimation of apparent dynamic modes (their timescales and amplitudes) with distributions spanning range of tens of ps to tens of ns. Figures 3.10 and 3.11 exemplify comparison of dynamic modes accessible from ACF fit representing average fluctuations of polypeptide unit orientation derived from 50 to 100 ns long MD simulations of MKK4 and NTail with dynamic modes estimated on the basis of Model-Free analysis of relaxation rates calculated from ACF-derived dynamics assuming three or four dynamic modes with imposed adiabatic approximation. Within magnetic field range of 14.1 to 22.3T and timescale range not exceeding 10-12 ns above which $\mathrm{L}(\omega_\mathrm{N})$ becomes less dependent on timescale at which the motion occurs (figure 3.1), Model-Free analysis with four dynamic modes performed on the simulated relaxation rates dataset accurately reproduces all of the correlation times exhibited by the reorienting polypeptide units. For timescales longer than 10 ns, the sensitivity of R_1 experiment at 600 MHz is relatively small and information on timescales of motion is hidden within amplitudes of slow motions (Chapter 3.3).

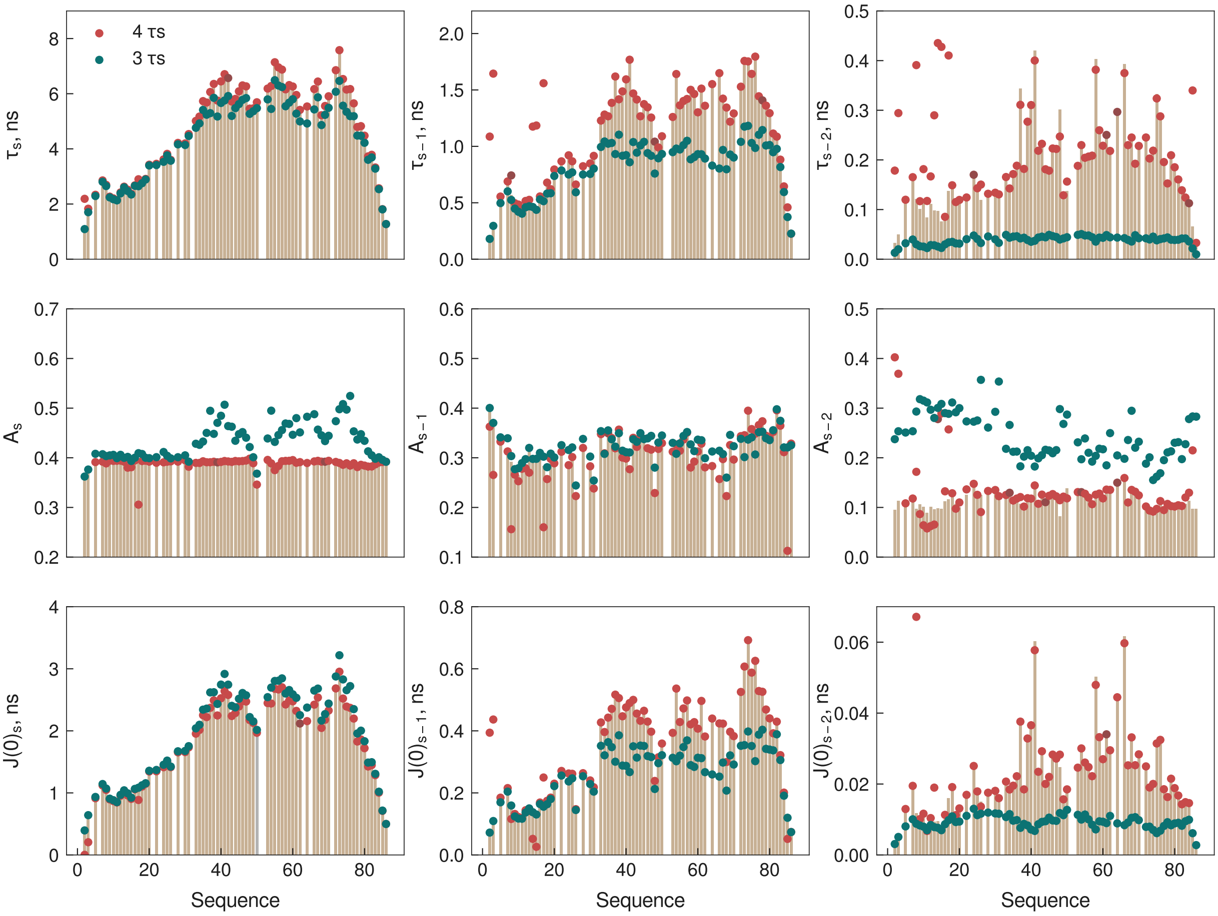

Figure 3.10: Comparison of dynamic modes for MKK4 (amplitudes and timescales of motions as well as their product - J(0) components, starting from the slowest one (s)) obtained from the fit of average NH-backbone correlation function representing average over 40, 50 ns blocks of trajectories, and Model-Free analysis with three or four apparent dynamic modes of relaxation rates calculated on the basis of MD-derived NH-backbone dynamics with magnetic field strength ranging from 14.1 to 22.3T, at 278K.

The estimated apparent timescale of slow dynamic mode sets the limit on the apparent mode of intermediate distribution, so that motions represented by the estimated (average) correlation times are statistically independent, which from practical point of view leads to a minimum ratio between modes of proximal distributions of five. As the amplitudes of motions represent relative memory loss due to n-th motional process within assumption of exponential decay of correlation functions with respect to sensitivity of SDF at frequencies associated with allowed spin transitions, they are subject to normalization with sum of their components equal to 1. Therefore, an increase in

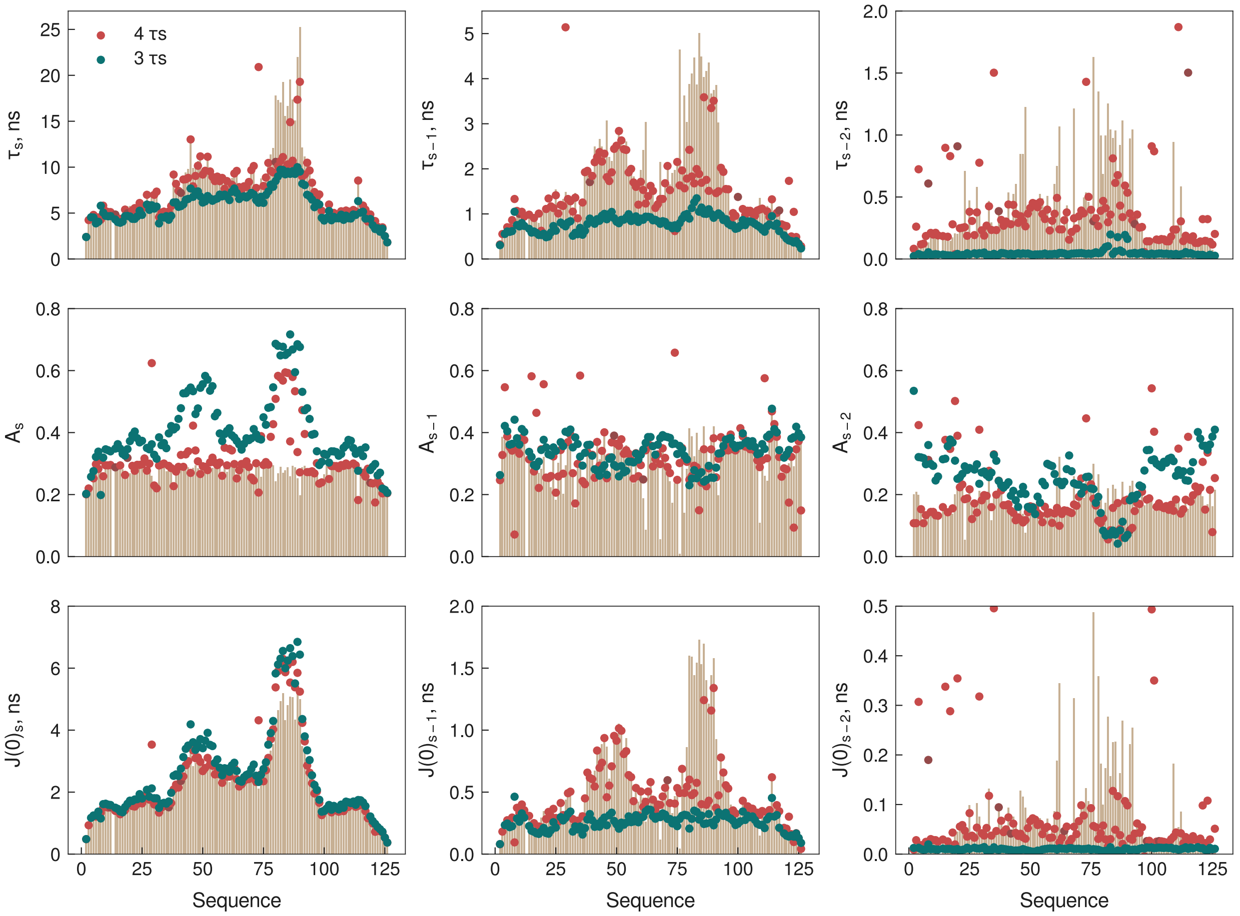

Figure 3.11: Comparison of dynamic modes for NTail (amplitudes and timescales of motions as well as their product - $J(0)$ components, starting from the slowest one (s)) obtained from the fit of average NH-backbone correlation function representing average over 54, 100 ns blocks of trajectories, and Model-Free analysis with three or four apparent dynamic modes of relaxation rates calculated on the basis of MD-derived NH-backbone dynamics with magnetic field strength ranging from 14.1 to 22.3T, at 298K.

amplitudes of slow motions will lead either to a decrease in amplitudes of intermediate or fast motions. The optimization of parameters space performed by d'Auvergne and Gooley for models constituting three timescales, two amplitudes of motion (order parameters) reveals high degree of correlation between order parameters for slow and fast motions within range of timescales of fast motions not exceeding 100 ps for combination of relaxation rates acquired at magnetic field strength of 11.8 and 14.2T. The correlation disappears for fast motions with timescales exceeding 100 ps. In other words, the bigger degree of overlap of timescales of fast motions with $\mathrm{L}(\omega_\mathrm{H})$ influencing accuracy of estimation of average correlation times, the smaller degree of correlation between slow and fast order parameters. If timescales of intermediate motions are underestimated due to lower limiting value of τ_{slow} arising due to timescale separation associated with adiabatic approximation, the value of intermediate timescales of motions contributing to J(ω_{H}) will correspond to population-weighted average correlation time of distribution spanning range of hundreds of ps to single ns, in turn influencing values of faster apparent dynamic mode, which contributes to J(ω_{H}). On the other hand, if the system may be characterized by dynamics with significant population of slow intermediate motions on timescales exceeding ω_N^{-1}, the average values of τ_{inter} and τ_{slow} will influence amplitudes of slow motions leading to underestimation of population of fast motions (213). Inclusion of relaxation rates at lower magnetic field strengths should allow to increase the dynamic range of the relaxation rates by allowing J(ω_{N}) and J(ω_{H}) to probe motions at hundreds of ps, single and several ns more accurately (209, 214).

Inspection of the quality of fit of the experimental relaxation rates (figure 3.9) allows to conclude that three apparent dynamic modes are satisfactory to describe the relaxation rates within the experimental precision. The purity of relaxation rates in terms of contributions from dipolar mechanism to relaxation of ^{15}N located within amide plane, i.e. from water protons in vicinity of peptide planes, which magnitude is comparable with experimental precision with which one can estimate the relaxation rates (215). The phenomenon in turn leads to difficulties in estimation of four dynamic modes from experimentally accessible relaxation rates following Model-Free analysis. In other words, the fourth mode, which could allow estimation of information on the reorientation processes occurring on hundreds of ps timescales is hidden within the experimental sensitivity of R_1 and nOe, precision with which they are measured and potential relaxation rate impurities. Temporal averaging of timescales of motions with three dominant contributions leads to partial representation of information on correlation times spanning range of single and tens of ns by amplitude of slow motions.

Secondary contributions to the estimated values of amplitudes of motions from Model-Free analysis assuming three dominant components arise due to sensitivity of different SDFs (J(ω_{H}) and J(ω_{N})) to different decades of timescales exhibited by reorienting peptide units (figure 3.1). To evaluate the phenomenon, consider an arbitrary peptide unit (NH), which reorientation process produces distribution of correlation times with modes at 50 and 500 ps corresponding to apparent fast and fast intermediate motions, which amplitudes and timescales derived from Model-Free analysis are estimated with similar precision due to similar sensitivity of J(ω_H) and J(ω_N) (NH1). If the peptide unit is reorienting faster due to transient changes in its molecular architecture or increase in temperature, the respective profile of amplitudes of motions might change (NH2). The distribution of fast timescales of motions is expected to be centered at smaller values - i.e. between 20 and 200 ps instead of 50 and 500 ps. In that case, the shift from NH1 to NH2 to a very small extent will decrease amplitude of fast motions by a factor associated with 'acceleration' of the fast mode (to 20 ps) and increase as a result of contribution of correlation times with distribution now centered at 200 ps. The same peptide unit pairs NH1 and NH2 might possess distributions with modes at longer timescales, i.e. 2.5 and 10 ns (NH1) and 1 ns and 4 ns (NH2). Motions represented by conformational transitions distributions centered at 1 and 10 ns should contribute to apparent intermediate mode in a similar manner, while motions at 4 ns (NH2) should be estimated with smaller accuracy than the 2.5 ns motions (NH1), in both cases due to the sensitivity of J(ω_N) (figure 3.1). In Model-Free analysis with three apparent dynamic modes, the NH1 is expected to posses intermediate dynamic mode in a range of approximately 2.0 to 2.5 ns and slower dynamic mode in a range of 8 to 12 ns, while the NH2 should posses modes centered at 700-800 ps and 3.5-4.5 ns, with partial timescale dependent compensation of slow and intermediate apparent dynamic mode values by amplitudes of slow and intermediate motions respectively. The overall shift of distributions from NH1 to NH2 should cause redistribution of amplitude profile resulting in greater amplitudes of fast motions and smaller amplitudes of slow motions. The expected changes in amplitudes of motions, due to acceleration from NH1 to NH2 might stem from shift or changes in population of reorientation rates spanning range of hundreds of ps (500 to 200 ps), now detected to a greater extent by J(ω_H) instead of J(ω_N), that would not be compensated by increase in J(ω_N) associated with acceleration of slower motions (from 10 to 4 ns and 2.5 to 1.0 ns). An example of the dynamics may be found for residues 40 and 10 of MKK4 (figure 3.10) for NH1 and NH2 respectively.

3.6 Conclusions

Model-Free analysis of Gaussian-shaped distributions of correlation times spanning three orders of magnitude (from tens of ps to tens of ns) confirms observed in the analysis of experimental data necessity of assumption of three apparent dynamic modes corresponding to statistically independent dominant components of correlation function representing reorientation processes encapsulated within ensemble and time-averaged spin relaxation rates. As long as ratio between modes of proximal distributions of timescales is at least five, independently on variances of distributions (with up to 90% of the mode), the obtained from Model-Free analysis timescales of motions are representative of modes of the underlying distributions. With increasing span of the distributions, represented by its variance (beyond 30% of its average value), information encoded in amplitudes of motion is perturbed in a manner that depends on ratio between averages of proximal distributions of timescales and difference between apparent dynamic mode and that of the underlying distribution. In that case, a portion of information originating from timescales of slowest motions is represented by the apparent amplitude of slow and intermediate motions for high and low ratios between modes of proximal distributions respectively. When ratio between modes of distributions is smaller than five - the distributions are more likely to overlap and the resultant from Model-Free analysis timescales of motions represent which values may be found in the tails of the joint distribution. Instances of the phenomenon may be observed for systems under experimental investigation in the study, i.e. ultra-flexible N-terminal fragment of MKK4, and are possible to predict *a priori* Model-Free analysis from RSDM. Peptide units exhibiting small J(0) and large J(ω_N) to J(0) ratio suggest overlap of distributions spanning single and several ns timescales. As long as the distributions do not overlap significantly, Model-Free analysis (within the bounds of adiabatic approximation) informs on apparent dynamic modes representative of correlation function describing reorientation processes of polypeptide fragments. The additional comparison between dynamic parameters derived from Model-Free analysis of NMR relaxation rates (calculated on the basis of MD trajectory dynamics) with three and four apparent dynamic modes points to averaging of timescales of motions, and their semi-equivalent representation as amplitudes of motions for fast and slow intermediate motions occurring on hundreds of ps and single ns, as a result of sensitivity of different J(ω) contributing to the experimentally accessible relaxation rates.

CHAPTER 4

Temperature dependence of unfolded protein dynamics from NMR relaxation rates

Stochastically modulated motions acting on polypeptide units cause their reorientation on multiple length and timescales. The observed by NMR relaxation ensemble of spin transitions allows to estimate parameters characterizing reorientation processes occurring on ps to ns timescales, which for unfolded proteins are coincident with in-and-out of balance movement of bond vectors (librational motions), as well as dihedral angle fluctuations limited by internal and external degrees of freedom associated with reorientation of several peptide units. Observation of NMR relaxation rates as a function of temperature allows to modulate distributions of correlation times of polypeptide fragment reorientation processes allowing to more efficiently probe their dynamics. The presented previously temperature-dependent analysis of NMR relaxation rates has been applied to an unfolded system featuring relatively broad or overlapped distributions of correlation times spanning range of single to several ns. Temperature dependent analysis of amplitudes of motions representing number of correlation times associated with polypeptide reorientation occurring concurrently to the allowed spin transitions allows to gain initial insight on their distribution(s). Characteristic temperature coefficients calculated from the analysis have been parametrized leading to a general estimate of correlation times distributions exhibited by reorienting peptide units of an unfolded protein of interest on different decades of timescales.

4.1 Introduction

Stochastically modulated fluctuations lead to variations in interatomic distances between fragments of polypeptide chains. When grouped together, the fluctuations cause modulations of internal and external degrees of freedom of peptide planes resulting in changes in $\phi\psi$ angles which may be represented by transitions within and between minima of Ramachandran space. These collective motions require larger energetic investment than sole back-and-forth movement of a backbone vector. Depending on molecular architecture of a macromolecule, the motions may be associated with larger scale polypeptide fluctuations resulting in segmental motions as typically observed in unfolded proteins and polymeric materials or in overall reorientation for globular proteins (126, 136, 139, 140, 216–219). The transitions may be described using a so-called energy landscape of conformational space. The concept first described by Eyring in 1935, has its beginning in defining transitions between accessible states, and is typically used to describe reaction rates based on interdependence of kinetic and thermodynamic information:

$$k = \frac{k_B T}{h} \exp\left(-\frac{\Delta H}{RT}\right) \exp\left(\frac{\Delta S}{R}\right) \tag{4.1}$$

where k is rate of reaction, which in the view of conformational fluctuations may be represented by a rate of transition from one state to another, while ΔH and ΔS represent thermodynamic parameters associated with changes to interaction strength within or between molecule fragments, as well as medium and the latter (enthalpic term), as well as changes in its number of degrees of freedom associated with transtion (entropic term), allowing to estimate energy difference associated with the transition - effective barrier height between accessible states (220–222). The equation may be understood as follows: the lower energetic barrier, the higher probability of transition between two states, the faster reorientation process. Translating kinetic rate theory to protein dynamics, low complexity and very fast motions, i.e. NH-backbone bond librations, which occupy low energetic states are similar in nature and may be represented by a relatively flat potential. The motions are encapsulated within a minimum describing energy-state dependence of more complex motion (e.g. reconfiguration within and between Ramachandran basins) requiring more time to occur. For motions of greater complexity, one can observe transitions between Ramachandran minima, which depending on composition of a macromolecule, here defined by peptide bond geometry and steric hindrance of protein residues, may be represented by similar differences between energetic states. Beyond single residue level, one can expect much broader energetic landscape representing collective transitions of groups of atoms (125, 126, 139, 201, 217). These concerted motions, first observed in polymers during heat activated processes, such as glass transitions are result of large-scale reconfiguration of entire macromolecule of interest (126, 136, 216). Overall, the reorientation processes give an idea of hierarchical ordering of protein dynamics ranging from small to large scale motions featuring gradient of complexity (107, 223–225). The ensemble of these processes has been observed experimentally for folded proteins from an analysis of solid-state NMR relaxation rates of different nuclei located in distinct polypeptide fragments of monocrystalline GB1, which motions have been gradually awakened by systematic increase of temperature of the system (226).

The majority of monomeric folded proteins posses tight molecular architecture featuring a number of intramolecular hydrogen bonding and electrostatic interactions between polypeptide units, which lead to their correlated reorientation resulting in an overall tumbling of entire macromolecule (125, 126). The dynamic modes representing reorientation rates of polypeptde fragments and relative population of correlation times with respect to sensitivity of NMR relaxation experiment feature clear timescale separation, with slower timescales representing reorientation processes occurring on timescales of few to tens of ns, and faster dynamics, which observation is possible for protein fragments exhibiting greater number of degrees of freedom, i.e. residues located in loops and close to the protein termini occurring on tens to hundreds of ps (79, 123). The collective nature of motions exhibited by folded proteins leads to dominance of slow dynamics in reorientation processes leading to the allowed spin transitons sets the limit on accurate estimation of all of the parameters characterizing distribution(s) of correlation times representing them from NMR relaxation rates. The information on local dynamics exhibited by folded protein fragments is in turn represented by so-called order parameters, which defined analogously to amplitudes of motions (with S^2 equal to A_{slow}) are proportional to number of correlation times with respect to sensitivity window of SDF components at linear combination of Larmor frequencies of interacting nuclei (typically ^{15}N and ^{1}H for backbone amide order parameters). In terms of stochastic reorientation of macromolecule fragments, estimated from NMR relaxation order parameters delineate intermediate *plateau* of the correlation function representing relative memory loss of spatial part of spin-lattice interaction tensor due to slower reorientation processes with respect to sensitivity of SDFs at applied magnetic field strengths. In other words, S^2

is associated with limiting value of timescales of slow motions estimated from NMR relaxation rates and denotes their relative contribution to SDFs. Since faster motions are associated with more localized fluctuations of polypeptide unit fragments than the slow ones, for proteins featuring similar size (with similar reorientation rate), order parameter may represent limitations imposed on the reorientation of peptide unit(s) (125, 126). In turn, ^{15}N-backbone and ^{13}C' order parameters have been linked to relative degrees of freedom of reorienting peptide plane(s) - the smaller the order parameter, the greater degree of mobility of polypeptide fragment containing the nuclei which NMR relaxation is evaluated (107). One of the oldest and most popular model allowing to estimate relative degree of mobility of polypeptide fragments from Model-Free analysis of NMR relaxation rates approximates internal motions of peptide plane(s) by diffusion of a bond vector in a cone with semi-angle θ'. More advanced approaches propose to model order parameters by application of Gaussian probability distribution in three orthogonal planes with variances representing relative angular degrees of freedom of peptide plane (GAF) (113, 114, 227).

The modifications to reorientation dynamics of isotropically tumbling folded proteins associated with changes in parameters describing molecular and macromolecular environment of the medium may be represented by Stokes-Einstein equation with temperature dependent changes to overall reorientation rates proportional to relative changes in solvent viscosity. The timescales of motions occurring on tens to hundreds of ps estimated from Model-Free analysis do not reveal marked temperature dependence, or the latter is obscured due to dominance of slow dynamics in reorientation process of the macromolecule. In turn, information on temperature dependence of internal motions is represented by relative changes in order parameters, which in turn may be linked to conformational entropy following application of a potential energy function representing relative number of degrees of freedom characterizing capability of polypeptide unit fragment to reorient (99, 228–233). Temperature dependence of order parameter is relatively flat and may be approximated by a linear or exponential function, the latter defining a rate coefficient (Λ), which may be connected to inter or intramolecular restrictions limiting peptide plane reorientation in a variety of energetic potentials. As the description of thermodynamics of folded proteins in square potentials is limited to relatively small rates of change of order parameter with temperature, more complex potentials basing on maximum entropy principle have been invoked, which define contribution to reorientation rates in arbitrary potentials restricted by equilibrium probability distribution of polypeptide fragment orientation (99). The latter allows to describe temperature dependence of local mobility represented by order parameters smaller than 0.5, with an arbitrary degree of mobility defined by cone with semi-angle ranging from 0 to $\pi/2$. The parameter Λ for NH-backbone order parameters typically ranges from 0.5 to 5 in structured parts of proteins leading to estimation of relative degrees of freedom on the order of 1 to 45°with a potential well minimum of approximately 5 to 12 kT (98, 229).

Although, modeling of NH-backbone thermodynamics described by diffusion of the vector in a cone model posses several limitations associated with accuracy of estimation of influence of local librational motions and slow correlated dynamics on reorientation rates of polypeptide fragments, its application allowed to characterize temperature dependence of relative number of degrees of freedom of peptide units for a number of folded proteins, which may be further related to heat capacity associated with its reorientation. The estimated heat capacity proportional to number of degrees of freedom is expected to yield greater values in unfolded parts of the chain and smaller values in the ones featuring stable secondary structure (230). Temperature-dependence of order parameters defined by Λ for Villin Headpiece HP36, B1 domain of protein G and *E. coli* derived Ribonuclase H is similar for folded and unfolded parts of the protein with limited influence of secondary structure on estimated heat capacity of the peptide planes. The latter may be associated with a relatively high propensity for intramolecular interactions between protein fragments (segments of peptide units counting up to ten residues) influencing the relative degrees of freedom of more flexible polypeptide units (99, 228, 231, 232). Furthermore, application of temperature-dependent Model-Free analysis of NMR relaxation rates in studies of NH-backbone dynamics Basic Leucine Zipper GCN4 described by reorientation of vector in a cone model allows to estimate relative contribution of backbone reorientation rates to changes in conformational entropy associated with binding to DNA. The estimated value of 0.5 kJ mol^{-1} K^{-1} is in close comparison with values obtained from calorimetric data combined with analysis of relative contribution of backbone motions to entropy based on Monte Carlo and MD simulations. The analysis of SDF components at Larmor frequency of ^{1}H representing magnitude of observed fast motions reveals a degree of correlation with changes in the α-helical populations of the polypeptide fragments in the basic region of GCN4 represented by relative similarity to ^{13}C' chemical shift temperature dependence (234).

The complexity of motions exhibited by unfolded proteins precludes straightforward application and interpretation of analyses available for folded proteins. An elegant example of an extension of Model-Free analysis with respect to temperature dependence of apparent apparent dynamic modes (with respect to sensitivity of relaxation rates) of unfolded protein fragments has been provided by Abyzov et al. (2016) who estimated the modes from Model-Free analysis of more than 100 ^{15}N spin relaxation rates of C-terminal part of Sendai Virus Nucleoprotein (NTail) at four different magnetic fields and six temperatures ranging from 268 to 303K (235). This wealth of experimental data allowed to approximate parameters characterizing reorientation processes of polypeptide units across ps-ns timescales and their temperature dependence with single residue resolution. Very fast motional timescales (tens of ps) are both sequence and temperature-independent, while their amplitudes are modulated by both experimental temperature and position along the macromolecular chain. Consequently, the fast modes were assigned to librational motions in flat energetic basins. Motions occurring on intermediate timescales (at approximately single ns) were highly dependent on their position in the sequence and their apparent energies of activation exhibit marked differences between structured and unstructured parts of the protein suggesting their association with local backbone dihedral sampling (126, 216). The amplitudes, timescales and energies of activation associated with the slowest component (spanning range of several ns) featured a bell-shaped profile along the sequence similar to the one expected for a random polypeptide. If the dynamic modes occurring on slow timescales are associated with reorientation of segments composed of several peptide units, considering that moments of inertia of reorienting polypeptide unit(s) are directly proportional to its size, shorter constructs should exhibit faster reorientation rates (21, 22, 236–238), which has been observed experimentally (235). The ensemble of experimental observations has been confirmed from decomposition of correlation functions of rank 2 spherical harmonics for peptide planes originating from ensemble-averaged 100 ns long MD trajectories of NTail into a product of three correlation functions describing the reorientation processes (136, 139).

4.2 Methods

Estimation of temperature-dependent dynamics accessible from NMR relaxation rates

The intrinsically disordered regulatory domain of MKK4 (residues 1-86, UniProtKB accession number P45985) and C-terminal domain of Sendai Virus Nucleoprotein (residues 401-525 UniProtKB accession number Q07097, here referred as residues 1-125) has been expressed and purified as described previously (68, 70). Samples for the NMR spectroscopy measurements have been prepared in Hepes buffer of pH 7.0 with 150 mM NaCl and 10% D_2O. ^{15}N relaxation rates for modeling of NH-backbone dynamics were recorded on four spectrometers (with magnetic field strength ranging from 14.1 to 22.4T) and four temperatures covering range of 273 to 288K. ^{15}N auto-relaxation experiments have been performed with HSQC-type pulse sequences with water suppression achieved with Echo-Antiecho detection scheme (210). The delays for R_1 and $R_1\rho$ experiments ranged from 40 ms to 1.4 s and 1 to 200 ms respectively. The spin-lock field strength in ^{15}N $R_1\rho$ experiment has been adjusted to approximately 2.0 kHz with adiabatic half-passage hyperbolic tangent pulses with 3 ms ramps and temperature compensation of up to 300 ms. The amide proton decoupling during relaxation delay of R_1 experiment was achieved using iBurp pulse of 2 ms centered at 8.5 ppm. Heteronuclear nOe measurements were performed with saturation and recovery periods of 6-8 s. ^{15}N CSA-dipole-dipole interference (η_{xy} and η_z) have been performed with Echo-Antiecho detection scheme with delays of 60 and 100 ms for transverse and longitudinal cross-correlated cross-relaxation respectively (195, 211, 212). In all of the relaxation experiments, interscan delay length was set to 1.5-2.5 s. Model-Free analysis have been performed with two or three apparent dynamic modes for a total of 56 relaxation rates - four temperatures ranging from 273 to 288K, optimizing 12 parameters. Constant value of CSA-DD angle of 25° fast timescales of motions of 45 ps, backbone amide NH bond vector length equal to 1.015Å and ^{15}N CSA magnitude to -172 ppm have been assumed. The optimization of parameters has been performed with Levenberg-Marquardt algorithm.

4.3 Temperature-dependent analysis of NMR relaxation

As a follow-up to the work of Abyzov et al. (2016), I have undertaken efforts to apply the developed methodology to another system - fully disordered N-terminal fragment (residues 1-86) of cancer-associated protein Map kinase kinase 4 (MKK4) (235). The system offers a few advantages: high stability of samples, well dispersed ^{15}N-^{1}H spectra allowing for quick and error-minimized relaxation data processing, while providing large differences in observed values of relaxation rates across the polypeptide sequence (figure 4.2). The latter fact is associated with several features in the protein sequence: presence of ultra-flexible N-termini (residues 7-20) composed mainly of serines and glycines, followed by mixed sequence of apolar, flexible and charged residues, ending on high molecular-weight and polar residues (70-80), with several secondary structure-breaking PXP repeats (where X is any other residue) populating multiple sites of the sequence (figure 4.1).

^{1}MAAPSPSGGG ^{11}GSGGGSGSGT ^{21}PGPVGSPAPG ^{31}HPAVSSMQGK
^{41}RKALKLNFAN ^{51}PPFKSTARFT ^{61}LNPNPTGVQN ^{71}PHIERLRTHS ^{81}IESSGK

Figure 4.1: Sequence of N-terminal fragment of MKK4: blue and green mark residues belonging to docking motif and KIS domain respectively (70).

Highly heterogeneous response of NH-backbone dynamics to temperature is represented by systematic changes in ^{15}N backbone relaxation rates (figure 4.2). The temperature-dependent decrease in J(0)-dominated η_{xy} is associated with overall acceleration of backbone dynamics associated with reorientation of polypeptide units occurring on tens of ns timescales. The relatively invariant R_1 for majority of magnetic fields as a function of temperature for central part of the protein suggests small degree of perturbation to intermediate motions with respect to sensitivity of R_1 experiment. The correlated with temperature changes in R_1 and heteronuclear nOes for residues in the N-terminus for magnetic field strength range from 14.1 to 22.3T suggest overall acceleration of fast intermediate motions with distribution of correlation times centered at single ns towards hundreds of ps, giving rise to the temperature dependence of the relaxation rates.

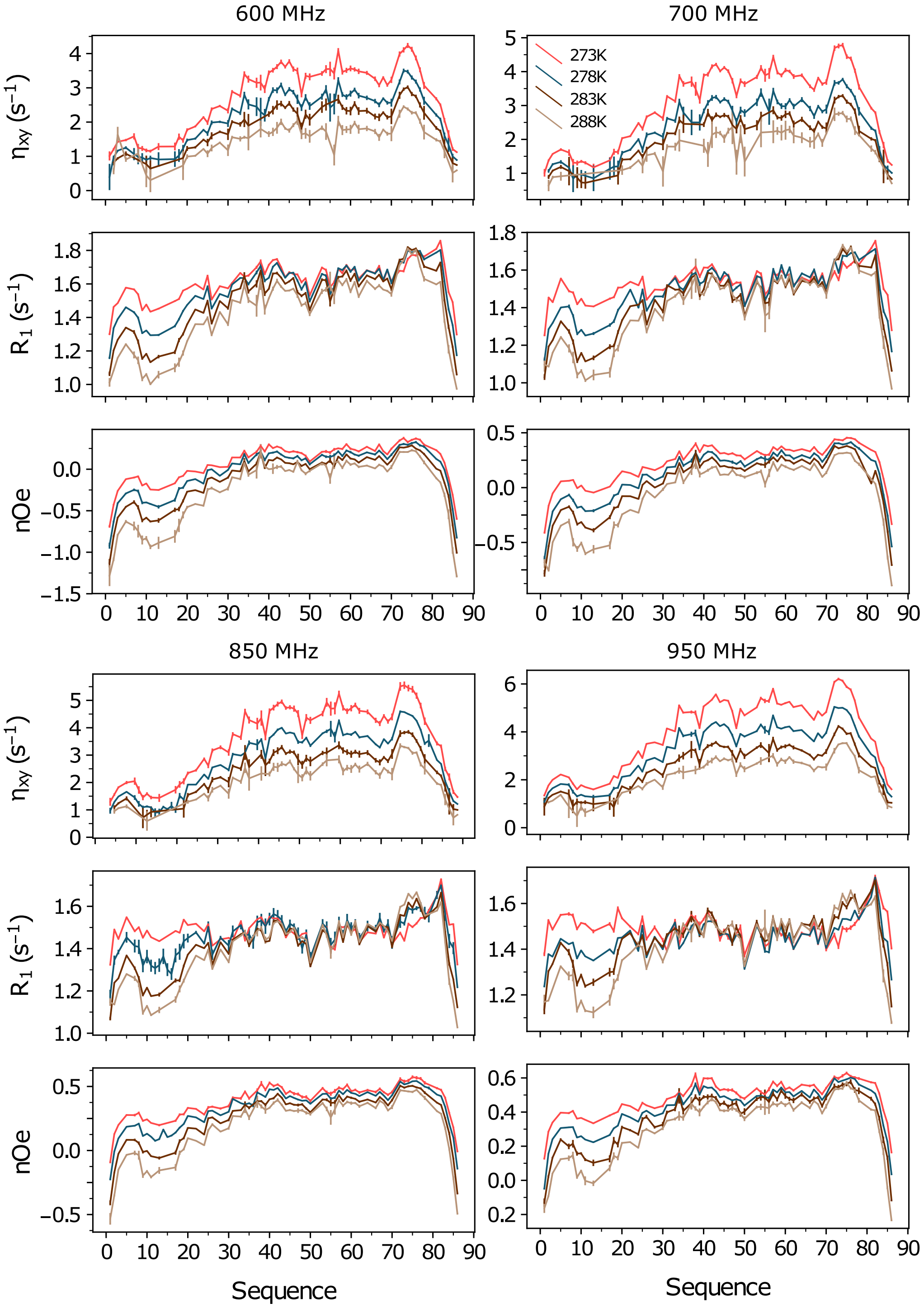

Figure 4.2: ^{15}N *backbone relaxation rates of MKK4 at multiple field strengths (ranging from 14.1 to 22.3T here marked with respect to the proton Larmor frequency at which spectrometer operates) and temperatures (ranging from 273 to 288K).*

The relatively small amount of dynamics on timescales longer than single ns for the N-terminal part of the protein suggests relatively collapsed distribution of correlation times spanning range of several ns. The C-terminal part of the protein reveals greater population of slower conformational transitions represented by correlation times spanning range of several ns, as observed in η_{xy}. The changes to R_1 with temperature at magnetic fields beyond 20T for the protein region suggest dominance of reorientation processes with correlation times beyond 2 ns. As in the previous study, several models of SDF featuring different number of apparent dynamic modes have been evaluated on datasets acquired at a single temperature (in the range of 273 to 288K) before modeling the entire dataset. Model 1, assuming two apparent dynamic modes included optimization of two timescales and one amplitude of motion. Model 2, with three dynamic modes included optimization of two timescales and two amplitudes of motions with fast timescales of motions set to a constant value of 45 ps. Model 3 included additional optimization of fast timescales of motions.

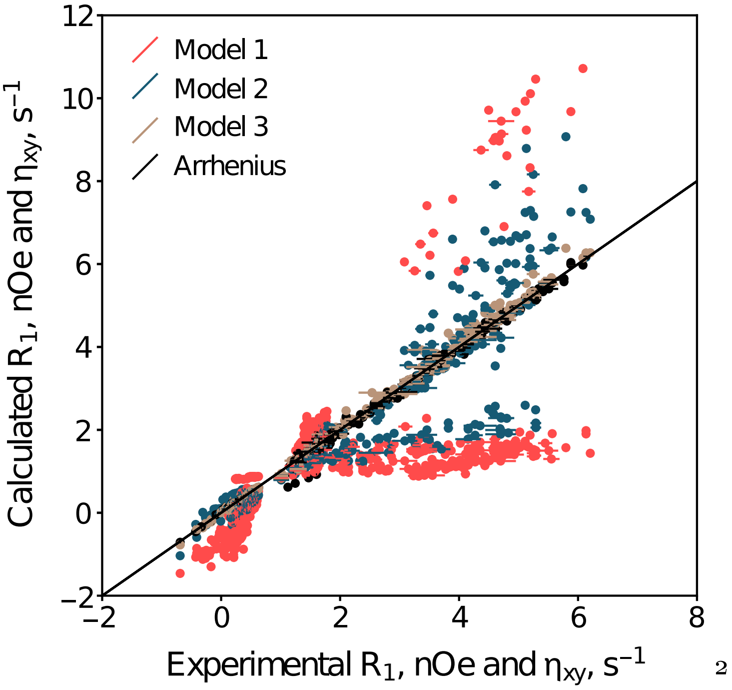

Figure 4.3: Comparison of different models for SDF at 273K. Quality of reproduction of randomly removed 30% of MKK4 relaxation rates dataset and back-calculated using remaining 70% in the analysis using different models of SDF.

Models containing only two components (1 and 2, figure 4.3) were not able to reproduce the data within experimental uncertainty, while the model with three dynamic modes reproduced the relaxation rates for all of the protein residues in the temperature range. The off-diagonal points in cross-validation analysis observed for models 1 and 2 suggest insufficient representation of reorientation rates of peptide planes leading to over or underestimation of experimentally accessible relaxation rates as has been discussed in the preceding Chapter (section 4). At 288K amplitude of slow motions is close to zero for residues 8-11 suggesting relatively small contribution of reorientation process to the ps-ns dynamics exhibited by the system. Inspection of the data reproduction from model 2 at 273K reveals quite good representation of experimental relaxation rates by the apparent dynamic modes at all of the magnetic fields for residues of the unfolded protein (Chapter 3.4). It is possible, that at temperature equal to and above 288K, contribution of slow motions to the ps-ns dynamics in the region is negligible and two modes would sufficiently describe the dynamic behavior of the polypeptide fragment. Similarly to analysis in the previous study, one of the timescales (representing fast motions) was constrained to a constant value within range of tens of ps (45 ps) (235), while the other two (representing slow and intermediate motions) were connected via pseudo-Arrhenius dependence:

$$\text{Model 3 with } \tau_i(T) = \tau_{i,\infty} \exp\left(\frac{E_{a,i}}{RT}\right) \text{ where } i = 1, 2 \tag{4.2}$$

where $E_{a,i}$ and $\tau_{i,\infty}$ correspond to i-th energy of activation and timescale at infinite temperature respectively, while R denotes gas constant (8.314 J mol^{-1} K^{-1}). The calculated parameters describing apparent dynamic modes of the system are in a range of values observed for NTail with fast timescale of motions exhibiting little dependence on temperature or molecular architecture (bulkiness and polarity of the side-chains), while significant modulations to amplitude of motions with respect to the parameters are observed (figure 4.4). Timescales of intermediate motions are in a range of single ns with marked sequence-dependent profile, while correlation times representing slow motions span range of single to few ns, and apart from the ultra-flexible N-terminus their sequence-dependence profile reminds of a bell-shaped curve.

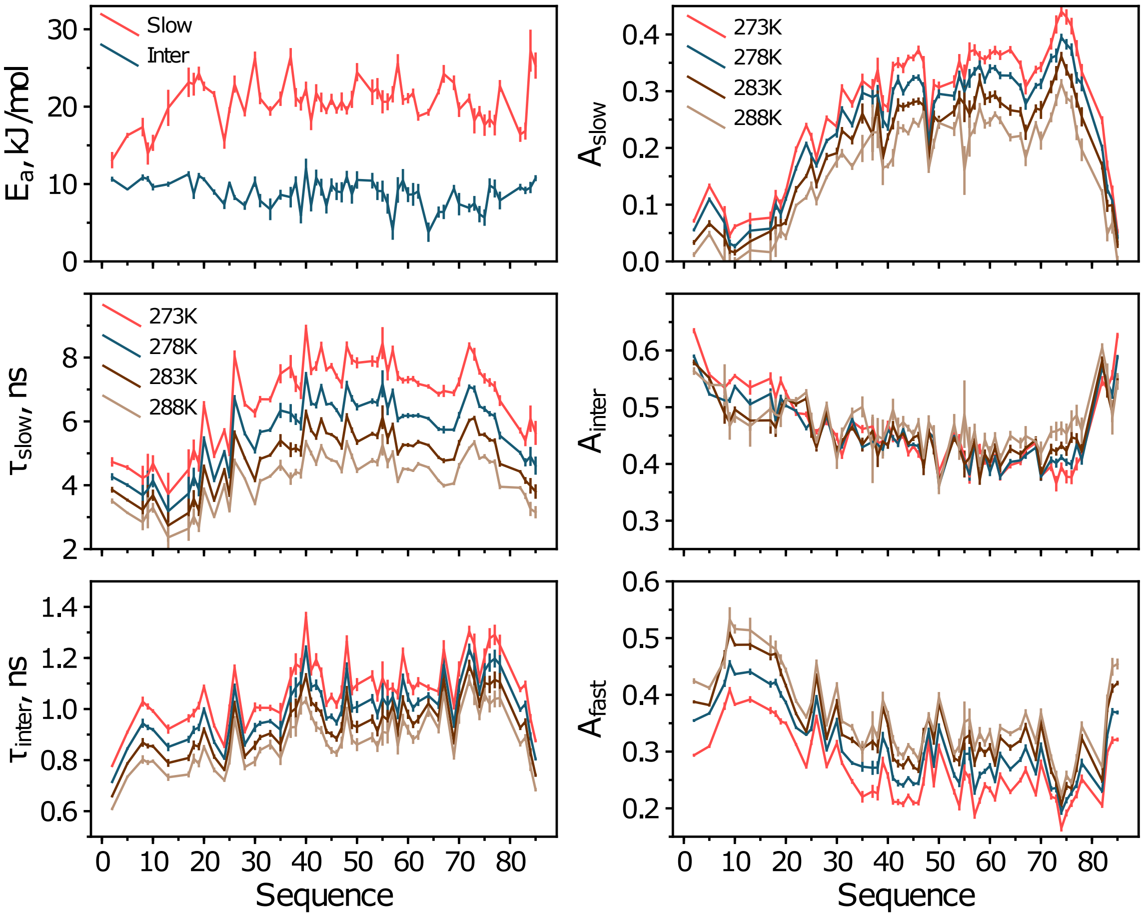

Figure 4.4: Result of the Arrhenius Model-Free analysis of ^{15}N *backbone relaxation of MKK4. All 56 relaxation rates were analyzed using Model 3 with equation 4.2 as described previously (235). This results in determination of site-specific activation energies along the sequence* (E_a) *for slow and intermediate timescales of motions, amplitudes of the three components* (A_{slow}, A_{inter}, *and* A_{fast}) *and timescales of each component as a function of temperature.*

As observed in the previous study, the estimated values of $E_{a,slow}$ are range from 20 to 25 kJ/mole and apart from two residues in the C-terminus - their sequence-dependent profile is similar to the one exhibited by timescales of slow motions (bell-shape) featuring local minima in vicinity of PXP repeats. Values of $E_{a,inter}$ are within range spanning 8 to 10 kJ/mole and posses much flatter profile to the one observed for NTail, especially around the termini (235). Amplitudes of slow and fast motions tend to exhibit overall linear dependence on temperature throughout the protein sequence with negative and positive slopes respectively. Amplitude of intermediate mode varies significantly across the protein sequence, and with increasing temperature changes overall profile from more oblique to more horizontal. Values of the amplitudes of motion estimated from the Arrhenius Model-Free analysis do not deviate significantly from the ones determined from independent Model-Free analysis (figure 4.5).

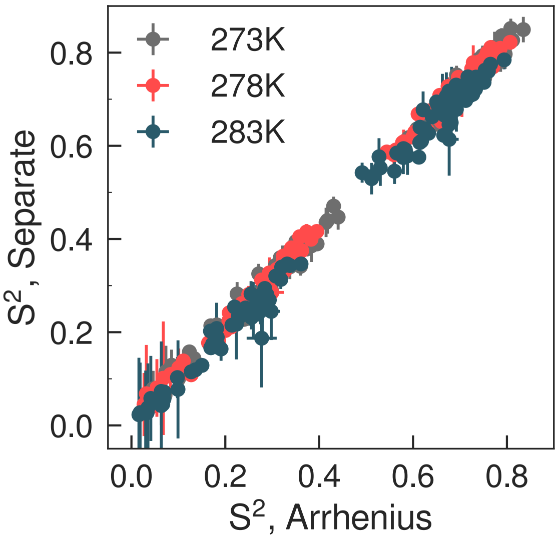

Figure 4.5: Comparison of order parameters (S^2 is A_{slow}, S^2_f is equal to 1-A_{fast}) estimated using independent or Arrhenius Model-Free analyses.

4.4 Temperature dependence of dynamic modes

Considering that timescales of motions occurring spanning range of ps to ns exhibited by the unfolded MKK4 which are within range of sensitivity of relaxation rates in the applied range of magnetic field strengths (with τ_{slow} faster than 10 ns, figure 4.6), sequence-dependent profile of amplitudes of motions estimated from Model-Free analysis is directly influenced by distribution of reorientation rates of polypeptide fragments and restrictions imposed in Model-Free analysis requiring statistical independence of motions, as well as, normalization of amplitudes of motions. If intermediate motions dominate the overall correlation function, limiting values of τ_{slow} are determined by apparent mode of intermediate distribution of correlation times relative to the $\mathrm{L}(\omega_{\mathrm{N}})$ maximum and ratio between proximal modes associated with adiabatic approximation. If distribution of slow correlation times is relatively broad with ratio between dynamic modes of distributions representing the two slower reorientation processes exceeding five, estimated value of timescales of slow motions set by the limiting value of τ_{inter} will point to faster tail of distribution of slow timescales of motions due to the characteristics of $\mathrm{L}(\omega_{\mathrm{N}})$. If distribution of slow timescales of motions is relatively narrow, but intermediate timescales of motions dominate, value of the apparent timescales of slow motions corresponds to mode of the distribution. For distribution of correlation times with major influence of slow motions with correlation times significantly overlapping with $\mathrm{L}(\omega_{\mathrm{N}})$ and high $\mathrm{J}(0)$ values, its integral of correlation function represented by $\mathrm{J}(0)$ which sets the limiting value of τ_{slow} in turn influencing value of τ_{inter} to be in a range five times smaller. For relatively broad distributions of intermediate timescales of motions it is value of $\mathrm{L}(\omega_{\mathrm{N}})$ maximum and limiting value of τ_{inter} associated with adiabatic approximation that determines the slow apparent mode. The sensitivity of relaxation experiments and statistical independence of motions imposes that for collapsed distributions of correlation times estimated apparent amplitudes characterizing joint distributions might compensate difference between apparent timescales and that of modes of the distributions.

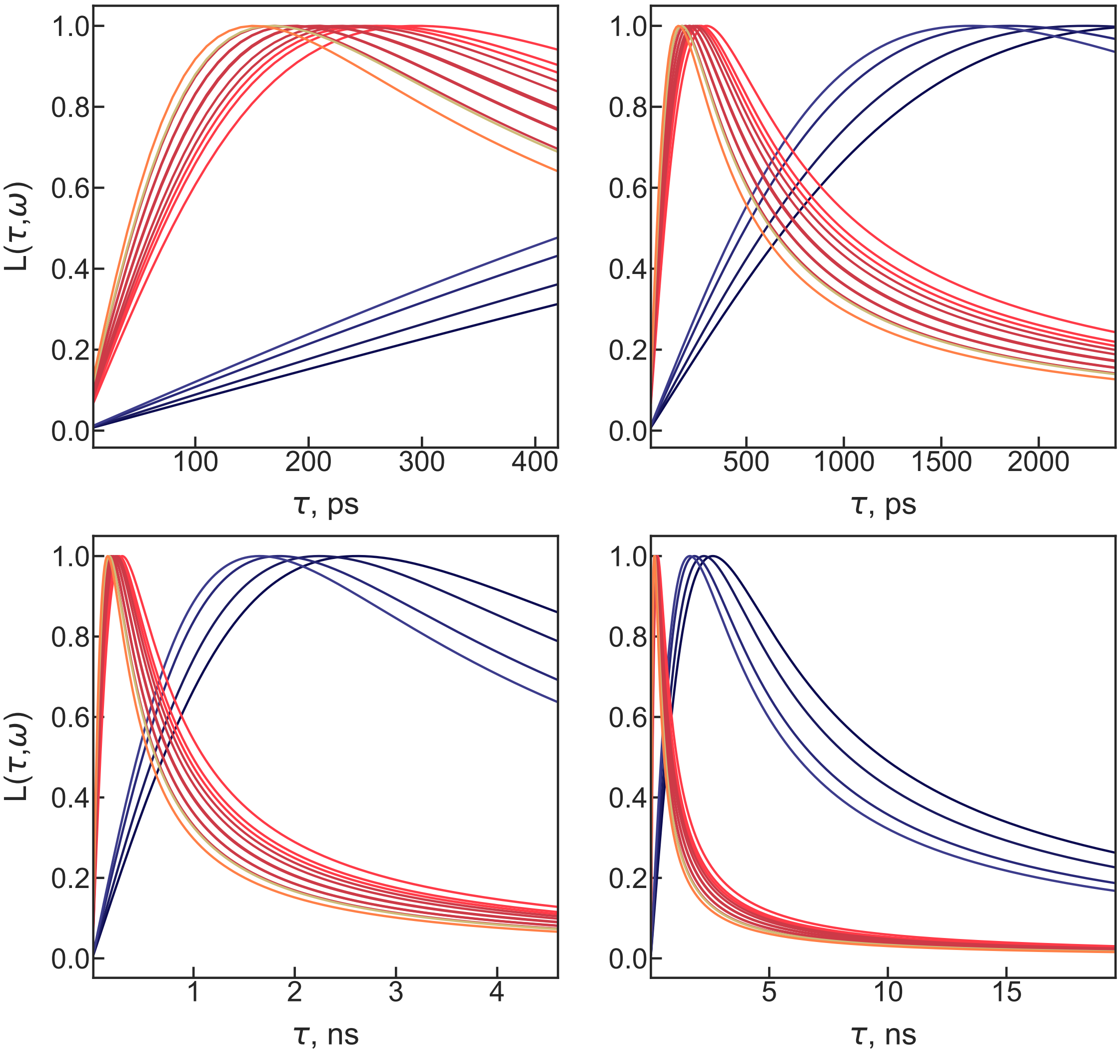

Figure 4.6: Sensitivity of different $J(\omega)$ terms to different timescales of motions at magnetic field strengths ranging from 14.1 to 22.3T defined by value of SDF at different timescale and frequency with respect to value of SDF at timescale of inverse of linear combination of Larmor frequencies associated with spin interaction. Orange and blue lines represent the values of $L(\omega_H)$ and $L(\omega_N)$ respectively.

Estimated amplitudes of slow motions for N-terminal part of MKK4 and the apparent dynamic modes suggest minor population of slower correlation times on timescales exceeding ω_{N}^{-1}. Motions exhibited by peptide units in central and C-terminal part of the protein, however, due to greater bulkiness of side-chains found therein are expected to be slower and feature distribution spanning single ns timescales of motions, which should be better separated from distribution of correlation times representing intermediate motions. Averaging of correlation times results in observed modulation of relative amplitude profile stemming from relatively small sensitivity of $\mathrm{J}(\omega_{\mathrm{N}})$ to motions at timescales longer than tens of ns and shorter than hundreds of ps (figure 4.6). Range of timescales estimated from Model-Free analysis for fast motions have its origin in population-weighted average of correlation times over range spanning tens and hundreds of ps. Separation of $\mathrm{J}(\omega_{\mathrm{H}})$ contributions from $\mathrm{J}(\omega_{\mathrm{N}})$ for timescales of fast motions leads to correlation between $\mathrm{A}_{\mathrm{slow}}$ and $\mathrm{A}_{\mathrm{fast}}$ (213). As a result, timescales of fast motions feature relative insensitivity to molecular architecture and temperature, and their relative contributions to SDF represented as $\mathrm{A}_{\mathrm{fast}}$. In other words, $\mathrm{A}_{\mathrm{fast}}$ increases as a function of additional correlation times at hundreds of ps, which overlaps with the sensitivity range of $\mathrm{J}(\omega_{\mathrm{H}})$ to an extent that is allowed by variation in $\mathrm{A}_{\mathrm{slow}}$ due to incomplete representation of slow or intermediate motions combined with normalization of amplitudes of motions condition. To estimate timescales of fast motions with greater accuracy, one would have to perform Model-Free analysis presumably with an arbitrary number of distributions of correlation times on a relaxation rate dataset free from contributions due to additional dipolar or scalar couplings with surrounding water protons with relative magnitude of the contribution comparable to precision of the experiment.

Upon increase of temperature, stochastic fluctuations leading to reorientation of peptide units are expected to accelerate, and, consequently change relative magnitude of observed SDFs contributing to the experimentally accessible relaxation rates. Faster intermediate motions occurring on hundreds of ps timescales contributing to, or rather, to which both $\mathrm{J}(\omega_{\mathrm{H}})$ and $\mathrm{J}(\omega_{\mathrm{N}})$ are sensitive to, are expected to shift towards smaller values increasing population of the fast apparent mode, and respective values of SDFs at ω_{H}, observed as an increase in estimated value of $\mathrm{A}_{\mathrm{fast}}$ and concomitant decrease in $\mathrm{A}_{\mathrm{slow}}$. If distributions of intermediate and slow reorientation rates overlap significantly, the values of $\mathrm{A}_{\mathrm{slow}}$ should decrease as a result of proximity of the distribution to $\mathrm{L}(\omega_{\mathrm{N}})$ maximum, or in other words, its degree of overlap with $\mathrm{L}(\omega_{\mathrm{N}})$.

The statistical independence of motions coupled with an increase in the values of A_{fast} imposes limitations on the rate of change of apparent dynamic modes with temperature. If the estimated apparent modes are well separated, the activation energies may vary to a greater extent yielding high activation energies with values ranging from 8 to 10 and 20 to 25 kJ/mole for intermediate and slow motions respectively. The difference between activation energies of apparent dynamic modes in the protein termini, or rather values estimated for entire flexible N-terminal part of the protein might originate from that condition. In other words, the difference between the activation energies in the studied case is negatively correlated with the degree of overlap between the distributions of the timescales of motions. Temperature dependent increase observed for amplitudes of fast motions in the N-terminal part of the protein suggests significant population of faster motions occurring on hundreds of ps contributing to $J(\omega_{\mathrm{H}})$ (100-500 ps), much greater than in the latter part of the unfolded chain. If reorientation rates spanning range of hundreds of ps have similar temperature dependence, rate of increase of A_{fast} with temperature should depend on relative proximity of the mode to maximum sensitivity of the $J(\omega_{\mathrm{H}})$, or in equivalent terms, to degree of overlap between the normalized SDF $L(\omega_{\mathrm{H}})$ and distribution of timescales of fast motions, as well as on relative population of correlation times spanning single to several ns timescales, leading to apparent dominance of distribution of correlation times spanning range of hundreds of ps. The increase in fast amplitudes of motions is accompanied by a decrease in amplitudes of slow motions to an extent limited by non-redundant $J(0)$ leading to temperature dependent decrease in values of A_{inter}, which is associated with overlap of distribution of intermediate and slow reorientation rates. On the other hand, the C-terminal part of the protein features smaller temperature dependence of A_{slow}, which might be associated with positioning of hundreds of ps distributions further away from maximum sensitivity of $J(\omega_{\mathrm{H}})$ with values above 500 ps, leading to its overlap with intermediate distribution centered around single ns. The relative increase of A_{inter} with temperature in the protein region might be associated with acceleration of slow ns motions, so that the slower correlation times posses values corresponding to greater $L(\omega_{\mathrm{N}})$. Overall, this gives an impression of quite broad intermediate distribution with correlation times ranging from 500 ps to 2 ns. Comparison of rates of change of amplitudes of motions with temperature should inform on relative span of the distributions. Correlated changes in slow and fast amplitudes of motions with temperature across the central part of the protein imply similar range of timescales sampled by hundreds of ps motions, or that relative span of the distributions in comparison to sensitivity of $J(\omega_{\mathrm{H}})$ is similar. Following similar considerations, one may interpret relative changes of amplitudes of motions of NTail. Increase in temperature leads to increase in accessibility of motions occurring on hundreds of ps timescales to J(ω_{H}) resulting in a flat increase of fast amplitudes of motions as a function of temperature within the entire studied range.

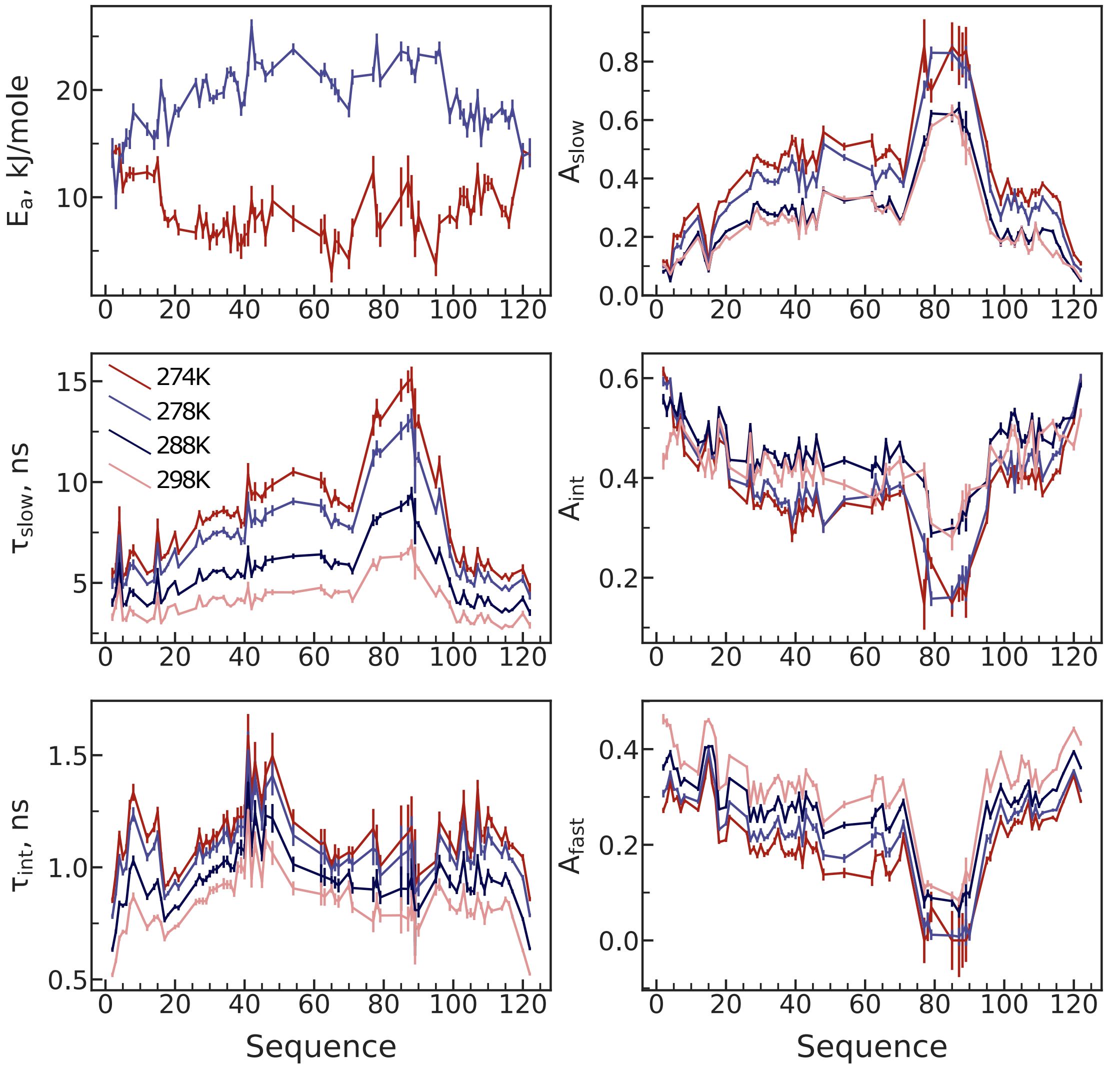

Figure 4.7: Result of the Arrhenius type Model Free analysis of ^{15}*N backbone relaxation of NTail. Reproduced on the basis of Abyzov et al., 2016. All 64 relaxation rates were analyzed using Model 3 with equation 4.2 as described previously (235). The determined site-specific activation energies along the protein sequence* (E_a) *for slow and intermediate motions, amplitudes of the three SDF components* (A_{slow}, A_{inter} *and* A_{fast}) *and apparent dynamic modes of the two slower components as a function of temperature.*

The high values of amplitudes of intermediate and fast motions, as wells as, relatively low values of apparent intermediate and slow dynamic modes suggest existence of significant population of fast intermediate motions occurring on a range of hundreds of ps with average value oscillating between 200 and 300 ps for residues in the protein termini and in the main chain respectively. Residues in the helical element and in-between two tryptophan residues W41-W49 exhibit remarkably low amplitudes of fast motions implying representation of intermediate and slow reorientation rates in a form of amplitudes of slow motions due to relatively small sensitivity of $J(\omega_N)$ to slower reorientation rates (Chapter 3.5). The magnitude of absorption of information on slow intermediate correlation times representative of difference between mean timescale and mode of the distribution to A_{slow} is inversely correlated with temperature. The latter occurs presumably due to acceleration of slow intermediate and fast intermediate motions towards maxima of $L(\omega_N)$ and $L(\omega_H)$ respectively. The amplitudes of intermediate motions in central part of the chain exhibit temperature-dependent trends observed for the C-terminal part of MKK4 and increase with temperature until 288K. The latter is coupled with a decrease in amplitudes of slow motions suggesting temperature dependent increase in the degree of overlap of distributions representing slow intermediate (single ns) and slow (several ns) motions. As a result of the increase of the two amplitudes (fast and intermediate), amplitudes of slow motions decrease until 288K. If reorientation rates spanning range of hundreds of ps accelerate leading to increase in amplitudes of fast motions due to greater amount of motions within experimental sensitivity, and population of distribution representing slow motions does not decrease their value, the amplitudes of intermediate motions are expected to decrease if all of the motions occurring on slow intermediate timescales are within sensitivity of $J(\omega_N)$. In other words, above 288K apart from the residues located in MoRE, the observed relaxation rates informing on slow intermediate distribution of timescales are already in sensitivity window of $J(\omega_N)$ with corresponding timescales of motions in a range below ω_N^{-1} at magnetic field strength accessed experimentally in the study (figure 4.6). Temperature dependence of the dynamic modes for protein residues located in termini reminds of trends observed for GS-rich region of MKK4 with overall increase in amplitudes of fast motions and decrease in amplitudes of slow and intermediate motions. The values of intermediate and slow energies of activation for residues located in the termini suggest a significant degree of overlap of intermediate and slow distributions of reorientation rates. The energies of activation in the MoRE suggest partial overlap of slow intermediate and slow timescales of motions changing with temperature much faster leading to increase in $E_{a,inter}$ with maintained separation between E_a.

4.5 Temperature dependence of motional amplitudes

As discussed above, the estimated amplitudes of motions depend on extent of averaging of correlation times with respect to sensitivity of J(ω_H) and J(ω_N) contributing to values of experimentally accessible relaxation rates, and accuracy of representation of motions occurring on slow intermediate timescales as apparent dynamic modes relative to dynamic modes of underlying distributions and their populations, imposed conditions associated with statistical independence of reorientation processes and normalization conditions of amplitudes of motions. The temperature dependence of fast motions is associated with relative number of correlation times representing fast and intermediate motions in comparison to the slow ones, and degree of proximity of fast intermediate modes occurring on hundreds of ps to maximum sensitivity of J(ω_H). The rate of change of amplitudes of motions should contain information on activation energy and number of correlation times associated with reorientation of polypeptide fragments on fast timescales (up to several hundreds of ps). The temperature dependence of slow motion amplitudes on the other hand, features more complexity due to changes in slow intermediate reorientation rates with respect to L(ω_N), while being directly affected by concurrent changes in correlation times not exceeding several hundreds of ps. Therefore, the rate of change of amplitudes of slow and fast motions should depend on rate of approach of apparent dynamic modes representing motions occurring on slow ns and hundreds of ps to values associated with maximum sensitivity of SDFs with 2.7 ns (J(ω_N)) and 300 ps (J(ω_H)) for relaxation rates measured at magnetic field of 14.1T.

If relative acceleration of fast intermediate motions (occurring at hundreds of ps) does not significantly depend on molecular architecture of polypeptide side-chains, then rate of increase of amplitudes of motions with temperature should inform on variance of distribution of intermediate timescales of motions, or rather, the rate of approach of fastest motions in fast intermediate distribution towards L(ω_H) maximum and their degree of overlap with the sensitivity function. The rate of change of amplitudes of slow motions is a bit more complex, since its value depends on the rate of change of amplitudes of fast motions with temperature, which might be increased by amount proportional to acceleration of slow intermediate reorientation rates and changes in its population, so that the correlation times feature intermediate timescales observed as contribution to J(ω_N). In the temperature range under investigation, one can easily observe that the amplitudes exhibit linear temperature dependence decreasing for slow and increasing for fast motions respectively. To interpret the phenomena, the temperature dependence of amplitudes of motions estimated from Arrhenius-type Model-Free analysis was represented by a simple two parameter linear function:

$$A_i(T) = a_i T + b_i \tag{4.3}$$

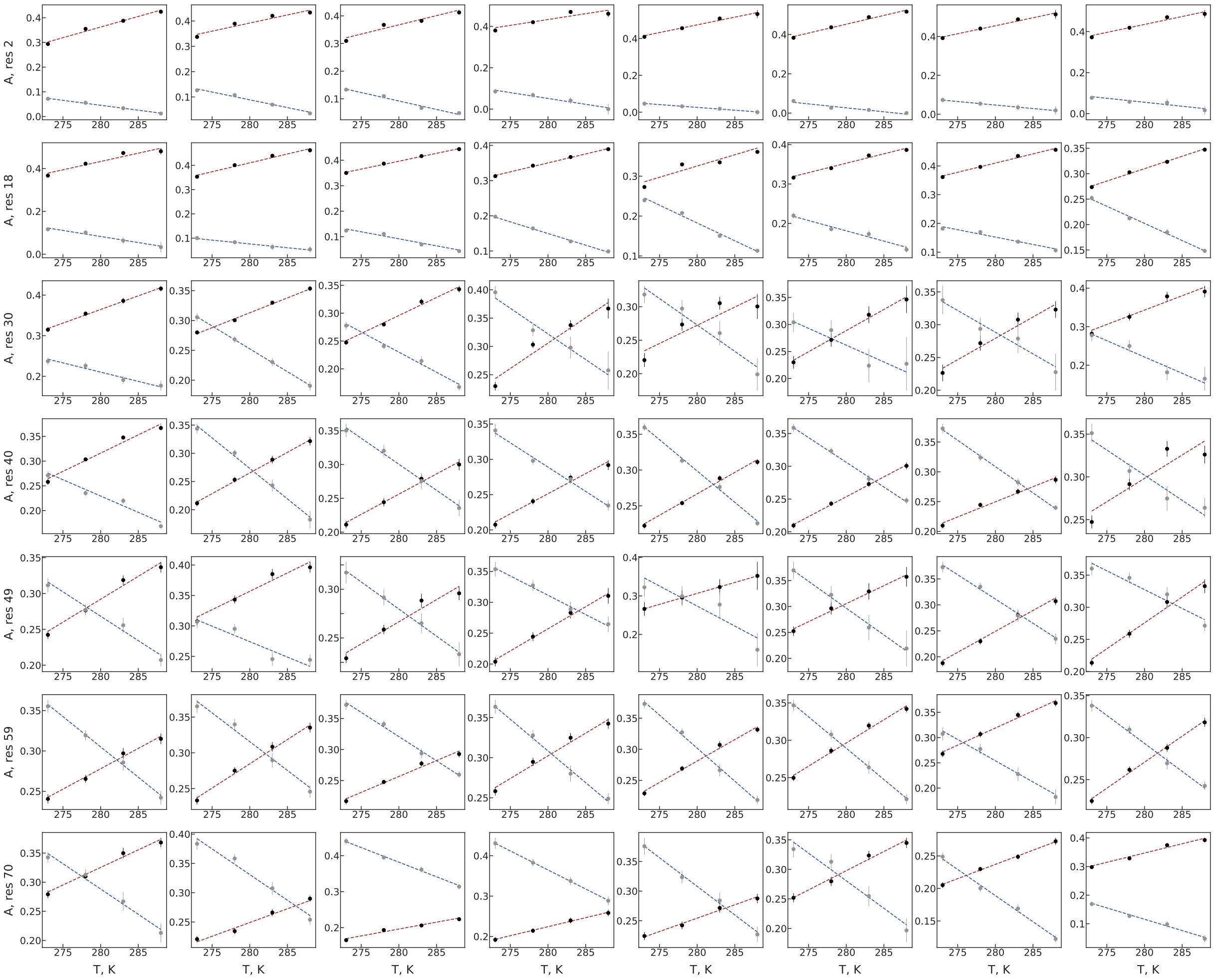

Figure 4.8: Temperature dependence of amplitudes of motion of MKK4. Black and grey points denote values derived from Arrhenius Model-Free analysis of MKK4, brown and blue lines represent fit to two parameter linear equation for fast and slow motions respectively. The fits have been performed in temperature range of 273 to 288K.

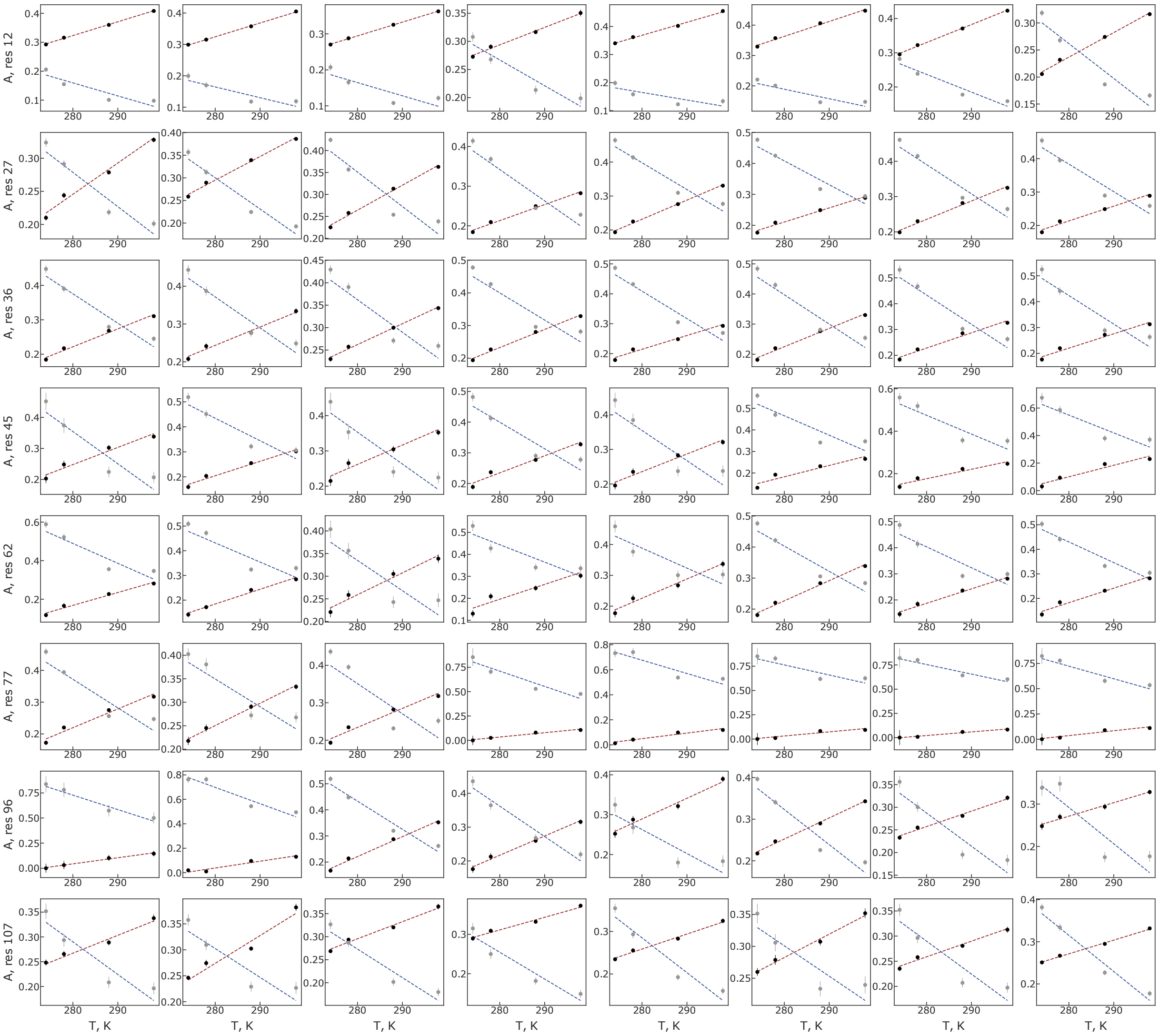

Figure 4.9: Temperature dependence of amplitudes of motion of NTail4. Black and grey points denote values derived from Arrhenius Model-Free analysis of MKK4, brown and blue lines represent fit to two parameter linear equation for fast and slow motions respectively. The fits have been performed in temperature range of 274 to 298K.

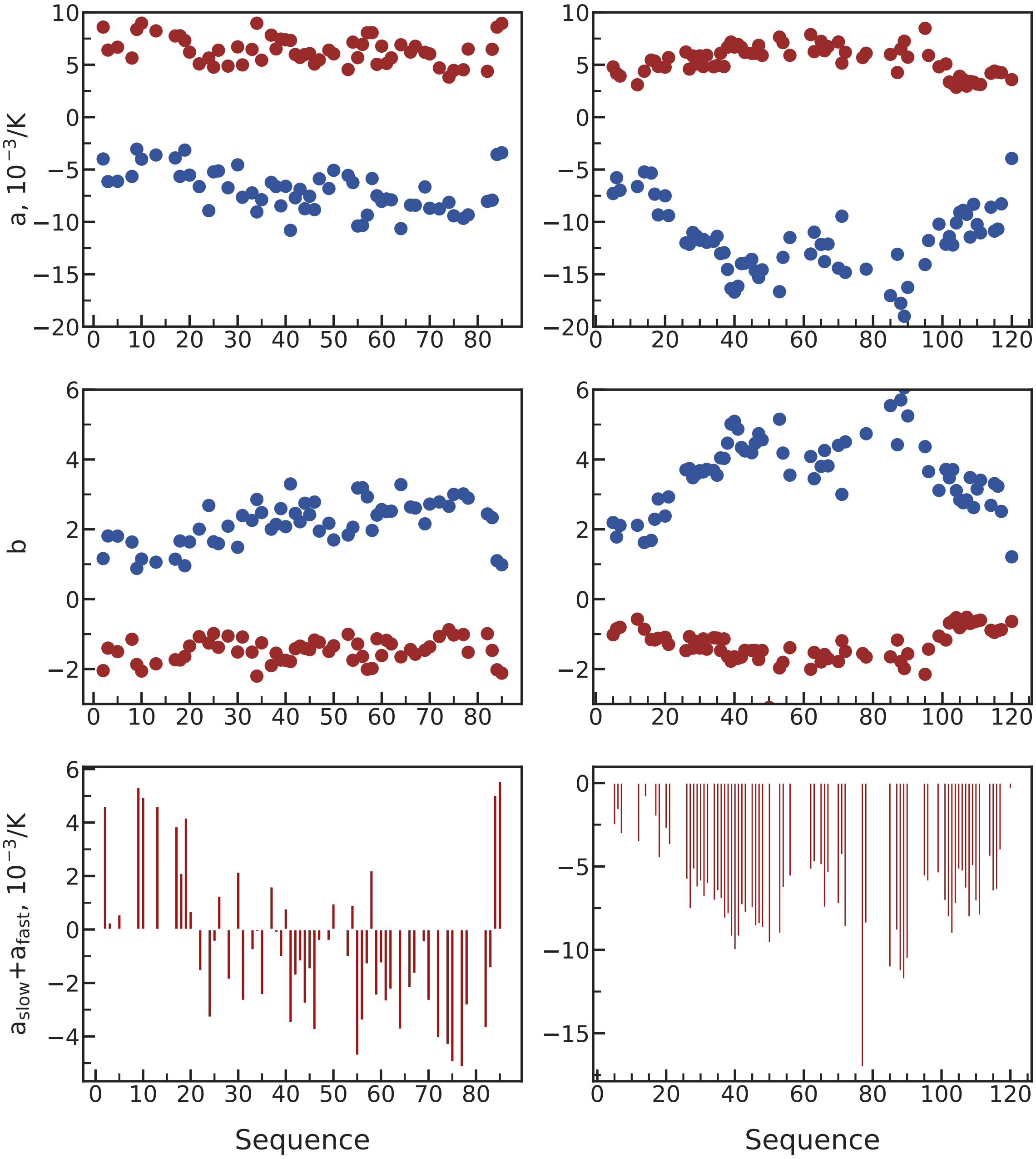

Figure 4.10: Comparison of parameters characterizing rate of change of amplitudes of motions with temperature for MKK4 and NTail model systems in temperature range of 273 to 288K and 274 to 288K for MKK4 and NTail respectively, brown and blue points represent parameters for fast and slow motions respectively.

The linearity of the amplitudes seems to be maintained over temperature range of 273-288K for both of the proteins and dynamic modes. For NTail, which temperature range was greater by 10K, it is more difficult to put a straight line through amplitudes of slow motions with the last point (at 298K) diverging from the linear trend. Therefore, the fit has been performed on a similar range of temperatures (274-288K). Overall, the rates of change of amplitudes of motions exhibit values in the similar range, of approximately 3 to 10, 10^{-3} K^{-1} for fast motions and from -5 to -20, 10^{-3} K^{-1} for slow motions. The similar rate of change of amplitudes of motions suggest, that motions occurring on timescales closer to the $L(\omega_H)$ maximum will saturate the sensitivity window faster than intermediate motions on hundreds of ps leading to decrease in amplitudes of intermediate motions. The latter is compensated with increase in A_{inter} associated with deceleration of slow intermediate to slow correlation times representing motions occurring on timescales ranging from single to several ns leading to a decrease in slow amplitudes of motions instead of intermediate ones. Further acceleration of fast timescales to values below 50 ps might lead to a relative deceleration of rate of change of amplitude of motion with temperature, due to a decrease in number of fast correlation times within the sensitivity window (50 - 350 ps). The high value of a_{fast} may represent either greater activation energy of the mode allowing for its faster approach to the maximum sensitivity of normalized SDF, or a well defined distribution centered around hundreds of ps, which sufficiently fulfills the sensitivity window without significant losses due to acceleration of fast tens of ps timescales. If the motions posses similar activation energies, the values of a_{fast} should be directly proportional to population of fast intermediate reorientation rates and inversely proportional to difference between the mode of underlying distribution and maximum sensitivity $(L(\omega^{-1}, \omega)$ and variance of distribution of fast motions with limiting values of 10 and 500 ps. In other words, a_{fast} is directly proportional to the rate of increment of overlap between distributions of fast timescales of motions and sensitivity window $(L(\omega_H))$.

The sequence dependence of a_{fast} for MKK4 exhibits profile observed for amplitudes of fast reorientation rates suggesting that motions exhibited by peptide units in the flexible N-terminal part of the protein are dominated by distribution with mode close to timescales spanning range of 150 to 250 ps, and increase in temperature gradually increases their availability to $J(\omega_H)$. Similar observations cannot be made for rate of acceleration of fast intermediate motions for residues located in the termini of NTail possibly due to greater number of fast motions with significant population of motions occurring on tens of ps timescales, which will accelerate with temperature leading to decrease of A_{fast}. The central parts of the polypeptide chain of both proteins features similar a_{fast} implying similar values of modes, populations and variances of fast and fast intermediate motion distribution exhibited by peptide units with presumed correlation times oscillating between 300 to 400 ps and 20 to 50 ps respectively. In other words, the increase in number of reorientation rates in sensitivity window due to fast intermediate motions is partially compensated by decrease in number of correlation times representing distribution of timescales of motions occurring on tens of ps. Residues featuring fast exchange between folded and unfolded conformations of NTail belong to the category. The relatively low values of a_{fast} for the C-terminal part of MKK4 featuring apolar and aromatic residues suggest slower reorientation rate and well defined distribution on hundreds of ps timescales of motions further away from the $L(\omega_{\text{H}})$ possibly separate from tens of ps timescales leading to loss of A_{fast} with temperature, as a result of acceleration of fast motions faster than saturation of $L(\omega_{\text{H}})$ by fast intermediate motions occurs.

The rate of change of amplitudes of slow motions with temperature is dependent on changes in amplitude of fast motions. Therefore, evaluation of rate of change of slow intermediate correlation times with respect to temperature should be performed after subtraction of relative rates of amplitude increment with temperature. The difference between a_{fast} and a_{slow} should take into an account temperature dependent decrease in A_{slow} associated with increase in A_{fast}, and allow to evaluate relative acceleration of fast intermediate correlation times, or rather degree of overlap of fast intermediate distribution of timescales of motions with $L(\omega_{\text{H}})$ in relation to slow intermediate correlation times, the latter contributing more to $J(\omega_{\text{N}})$ values. If difference between a_{fast} and a_{slow} is positive, the relative increase in number of fast intermediate reorientation rates in $L(\omega_{\text{H}})$ window is faster than that of slow intermediate ones in $L(\omega_{\text{N}})$ suggesting relatively small difference between average activation energies of the two apparent modes, that would not be compensated by increment of slow intermediate correlation times in the sensitivity window. In that case, the relative number of correlation times in the sensitivity windows defined by $L(\omega)$ dictates estimated values of amplitude of slow motions and relative changes of amplitudes with respect to temperature, the phenomenon observed for the N-terminal part of MKK4. The negative values of a_{diff} imply greater activation energy of slower intermediate motions, or rather suggest that greater rate of $L(\omega_{\text{N}})$ saturation by correlation times in comparison to $L(\omega_{\text{H}})$. The latter is observed for the central part of MKK4 and NTail with values of relative increment of amplitudes slow motions of NTail in a range twice greater than for MKK4 suggestive of greater degree of overlap distribution representing slow intermediate motions with $L(\omega_N)$ for NTail. The respective sequence dependence of a_{diff} points to the greater acceleration of slow intermediate motions in the central part of the protein with greater values for fragments featuring greater moments of inertia acting on reorienting chain (i.e. WW motif and helical residues) - motions in that range are expected to be dominated by correlation times spanning range of several ns, and relative decrease in amplitude of slow motions suggests their greater rate of the distribution overlap with $L(\omega_N)$ as a function of temperature in comparison to the motions exhibited by residues located close to the termini.

4.6 Discussion

The application of a recently developed Model-Free analysis based method for estimation of apparent dynamic modes of distributions of polypeptide fragments reorientation rates from NMR relaxation rates to an unfolded protein (MKK4), featuring distinct sequence-dependent conformational features compared to previously studied intrinsically disordered protein (NTail) allows for comparison of associated dynamic parameters and an assessment of genera applicability of the approach to studies of ps-ns unfolded protein dynamics. Extensive set of relaxation rates interpreted in terms of amplitudes and timescales reveals at least three dominant contributions to SDF featuring different degrees of complexity: fast fluctuations of bonds and small groups of atoms (vibrations associated with stretching and deformation of backbone vectors), backbone dihedral angle motion (represented by fluctuations within and between the Ramachandran minima) and correlated fluctuations of several residues in a row (segmental chain-like dynamics). The observed temperature dependence of amplitudes of motion for both proteins (NTail and MKK4) is in a similar range to the one observed generally for unfolded fragments of proteins (10^{-3} K^{-1}) and apart from residues involved in transient formation of helix of NTail and those located in ultra-flexible N-termini of MKK4 suggests rather compact distributions of polypeptide fragments reorientation rates (228, 230, 239, 240).

The average rate of acceleration of slower motions with temperature approaches energy of activation of bulk solvent reorientation (around 20 kJ/mole) with greater E_a values in center of the unfolded chain and smaller values observed for residues situated in the termini. The latter is likely associated with number of degrees of freedom in the segment and molecular architecture of side-chains influencing energetic balance between inter or intramolecular and solvent interactions. Without stable secondary structure element, the activation energies of apparent dominant intermediate modes are uniformly distributed around 10 kJ/mole, and do not exceed 15 kJ/mole limit, as it occurs for parts of the sequence of NTail featuring transiently formed α-helical recognition motif. The $E_{a,\text{ inter}}$ for residues in the protein termini is associated with imposed statistical independence of timescales of motions - the slow intermediate motions cannot approach smaller values faster then the minimum ratio of four to five between proximal averages of dominant distributions would allow for it. The intermediate timescales of motions of MKK4 as a function of protein sequence exhibit systematic increase of τ_{inter} towards more bulky part, where steric hindrance of side-chains slows down the reorientation rate (241, 242). Observation of significantly lower values of τ_{slow} for the N-terminal part and higher values for more bulky residues of MKK4, as well as, comparison of timescales of motions between unfolded parts and helical element of NTail suggests that their distribution is associated with moments of inertia acting on the reorienting peptide unit(s) associated with inter and intramolecular hydrogen bonding, electrostatic and van der Waals interactions contributing to transient formation of so-called segment. The interpretation is in agreement with observation of highly correlated reorientation of peptide units constituting polypeptide segments in MD simulations of NTail coinciding with greater values of τ_{slow} for the residues. In a similar manner, the reorienting folded elements or polypeptide segments of greater collectivity of the motion feature longer timescales of motions, which has been observed previously for a number of proteins i.e. GCN4, Engrailed 2, α-synuclein or δRNAP (125, 127, 136, 137, 201, 230, 243, 244).

With increase in temperature, slow and slow intermediate correlation times occurring on single to several ns shift towards smaller values with minor modifications to relative degree of overlap of fast intermediate and intermediate distributions of correlation times with timescales ranging from several hundreds of ps to single ns defined by $L(\omega_{\mathrm{N}})$. The latter is reflected in relatively similar rate of changes of amplitudes of fast and slow motions with temperature and almost invariant amplitude of intermediate motions for central part of the protein. The observation of highly positive a_{diff} in the N-terminal part of MKK4 is associated with acceleration of highly abundant fast intermediate motions in an absence of significant population or relatively broad distribution of motions featuring high degree of collectivity occurring on slower timescales. The phenomenon is presumably associated with relatively large number of degrees of freedom available to the peptide units located in the flexible N-terminus of MKK4 composed majorly from serines and glycines. Similar observations can be made for the residues located in the termini of both proteins. The remaining NH-backbone sites of NTail and bulky C-terminal part of MKK4 (residues 70-80) feature negative values of a_{diff} suggesting increase in number of slow intermediate motions with decrease in temperature resultant from increase in cooperativity of polypeptide unit reorientation. The latter presumably originates from relatively lower interaction energy associated with binding between neighboring side-chains with respect to the one for solvent-side-chain interaction. In order for that to happen, decrease in enthalpic term must overcome entropic penalty associated with decrease in number of degrees of freedom due to intramolecular interaction, the phenomenon presumably originating as a result of changes in hydrogen bonding interaction energy both between water molecules (reflected in differences in water clustering propensity) and with polar side-chain atoms of unfolded polypeptide chain (125, 126, 140, 245). The increase in temperature causes acceleration of intermediate and fast intermediate distributions of timescales of motions (occurring on hundreds of ps), which are presumably associated with limitations to reorientation of peptide planes due to internal degrees of freedom (127, 217, 235).

The unfolded chain of MKK4 possesses several sequence features which give rise to the observed profile of timescales and amplitudes of slow and intermediate motions reflecting significant degree of overlap of distributions of timescales occurring within $\mathrm{L}(\omega_{\mathrm{N}})$ due to faster reorientation rate of peptide units with resultant distributions centered at tens of ps, hundreds of ps and single ns. The N-terminal part of MKK4 featuring a stretch of serines and glycines separated by PXP repeats at both ends of the polypeptide unit is characterized by rather broad distribution of slow correlation times with average corresponding to farther shoulder of distribution rather than its mode. The latter suggests conformational heterogeneity of the unfolded protein fragment allowing the chain fragment to form an array conformers from very compact to very extended ones. The distribution of slow correlation times for residues in the C-terminal part of the protein featuring more bulky peptide units composed of charged and apolar residues is better defined with greater separation of slow correlation times from the intermediate ones, due to greater moments of inertia of the reorienting polypeptide units allowing for smaller degree of conformational variability (236, 246). Central part of the protein exhibits distributions of slow correlation times in-between the two extremities. Timescales of intermediate motions are relatively similar throughout the sequence and oscillate within several hundreds of ps to single ns with differences in region of the polypeptide featuring greater populations of β-proline and α-helical conformations, which are expected to yield slower reorientation rates on single ns timescales due to hydrogen bonding between side-chains of neighboring residues (41).

As reorientation rate of polypeptide units depends on its conformational sampling, exchange between conformational states on intermediate timescales may be represented by jumps within and between the Ramachandran minima. In that case, intermediate motions are expected to be sensitive to populations of conformational states (222, 247). Presence of several proline residues placed in consecutive manner leads to reduction in number of degrees of freedom of neighboring peptide planes, their slower reorientation and more defined distribution represented by local maxima in τ_{slow}, τ_{inter} and $\mathrm{A}_{\mathrm{slow}}$ observed for 20TP and 26SP. Additionally, τ_{inter} exhibits greater values in regions coinciding with enhanced β-proline sampling, i.e. in polypeptide region encompassing 39KRKA43 and 47NF with values greater by up to 20% in comparison with neighboring polypeptide fragments. The region (39K-48F) encompasses so-called docking motif and KIS domain, which are involved in binding to the folded p38α allowing for regulation of the MAP kinase cascade (70). The relative reorientation rates of peptide units on single to several ns timescales and its distribution associated with intramolecular hydrogen bonding propensity, which represented by significant population of β-proline conformations might facilitate binding of unfolded MKK4 to p38α with former maintaining the unfolded state upon binding. The relative deceleration of motions occurring on timescales spanning range of several ns in the region suggests presence of strained conformations possibly due to charge repulsion between consecutively positioned charged and bulky apolar side-chains in the subsequent part of the motif (44LKLN47). The combination of local backbone conformation and preferences for polypeptide topology in the region may facilitate binding of the unfolded MKK4 to p38α and modulate residence time of the bound state necessary for the regulatory domain to achieve its function. The region encompassing 67GVQN70 features consecutive local minima and maxima in τ_{inter} suggesting conformational heterogeneity in the peptide fragment, which might be associated with repulsion of positively charged side-chains of 69Q and 70N, in turn influencing proximity of Q69 to V68 consequently leading to increase in values of τ_{slow} and τ_{inter} as high as 20% with respect to the neighboring residues. Similarly, consecutive placement of negatively and positively charged residues 74E-75R enclosed by apolar residues 73I-75L in the 72-79 region increases probability of forming intramolecular contacts due to charge and steric interactions of neighboring side-chains leading to slower reorientation rate of the fragment on multiple timescales visible in the form of relatively greater τ_{slow} and τ_{inter}.

Similar characteristics in terms of distribution of timescales of motions to the ones found in the C-termini of MKK4 may be found for WW-motif and α-helical recognition element (MoRE) with relatively well-defined distributions of slow and slow intermediate timescales of motions ranging from single to tens of ns, which high populations represented by amplitudes of slow motions presumably originate from intra and intermolecular hydrogen bonding, electrostatic and van der Walls interactions resulting in highly correlated motions of the polypeptide fragments. Additional features associated with differences in distributions of fast and fast intermediate motions may be observed for NTail residues spanning region 30-40 (EAHADQDARG) featuring positively and negatively charged residues interspaced with apolar ones leading to oscillatory pattern in the amplitude of fast motions. The latter represent minor differences in degree of overlap of fast intermediate distribution of reorientation rates with $\mathrm{L}(\omega_{\mathrm{H}})$, which is associated with varying populations or variances of the distributions presumably resultant from varying (internal) degrees of freedom of the reorienting peptide planes associated with molecular architecture of their side-chains.

4.7 Conclusions

The discussed here approach, allows to uniformly describe apparent dynamic modes of an unfolded or transiently folded system across a range of temperatures in the same time reducing number of parameters necessary for accurate description of experimental relaxation rates. Additionally, the approach facilitates description of apparent dynamic modes under conditions which are challenging to record and analyze, i.e. high temperature influencing moments of distributions of timescales leading to their overlap, high pH increasing the amide proton exchange rate and its contribution to transverse relaxation rates effectively broadening the line shapes and limiting accuracy and precision of the measurement. The analysis allows to gain an initial insight on distributions of timescales of motions, rather than focusing on population-weighted average timescales of motions. Consequently, the analysis reveals well defined distributions of timescales of motions for NTail and MKK4 for majority of residues with dominant slow and intermediate dynamic modes in central part of the chain, and fast to fast intermediate in less bulky regions of the protein and in the termini. The temperature dependence of amplitudes of motions for residues featuring less bulky side-chains suggests partial overlap of slow and slow intermediate distributions of timescales of motions featuring either smaller population or greater variance of the former one, resultant from relatively fast reorientation rate of peptide units on longer length-scales. The latter implies relatively greater conformational heterogeneity of the reorienting peptide fragments associated with greater number of degrees of freedom in the flexible part of the protein than in the central part. The central part of the chain, which conformational heterogeneity allows for greater separation and more defined (of smaller variance) distributions of slow and slow intermediate timescales of motions. The apparent properties of the distributions are reflected in relatively high difference between energies of activation of the apparent modes and temperature dependence of amplitudes of slow and fast motions with significantly greater rates of decrement of the former ones. The temperature dependence of slow and fast amplitudes of motions represented by negative a_{diff} suggests relative increase in population of slow and slow intermediate motions with respect to $L(\omega_N)$ for all of the NH-backbone sites of NTail and central to C-terminal part of MKK4 suggesting increase in the degree of cooperativity of motions with decrease in temperature.

Chapter 5

A unified description of unfolded protein dynamics under native conditions

Reorientation of peptide planes strictly depends on their neighborhood and limitations imposed by steric and electrostatic interactions, as well as hydrogen bonding. Unfolded protein dynamics observed in macromolecular settings reveal marked differences in comparison to dilute phases. Additional limitations imposed on reorienting peptide plane due to presence of macromolecules lead to the slow-down of reorientation rate of its polypeptide unit(s) on different timescales directly representing stochastically modulated reorientation of peptide units limited by internal and external degrees of freedom. The chapter presents a systematic study on viscogen-induced solvent-mediated modifications to reorientation rate of peptide units from NMR spin relaxation. Model-Free inspired interpretation of ^{15}N relaxation rates within bounds of adiabatic approximation presents a direct evidence on coupling between stochastically modulated motions occurring on relaxation active ps-ns timescales and opposing it friction arising due to the limitations associated with intramolecular interactions between neighboring side-chains, as well as intermolecular interactions of polypeptide units with viscogen macromolecules. Limits and potential benefits of developed methodology allowing to predict viscogen-induced, solvent-mediated changes to the dynamic modes estimated from ^{15}N relaxation rates have been discussed.

5.1 Introduction

Reorientation of peptide plane depends on local degrees of freedom which may be defined by range of fluctuations of side-chains limited by hydrogen bonding, as wells as steric and electrostatic interactions with neighboring peptide units, which are encoded in their molecular architecture. According to the central limit theorem, probability distribution of distances between consecutive peptide units, which may be defined by separation between centers of their masses, increases by a factor associated with variance of distribution of distances between adjacent carbon alpha atoms. The latter depends on local degrees of freedom of polypeptide units. Therefore, steric and electrostatic interactions, as well as, hydrogen bonding contribute to modulation of degrees of freedom of peptide planes and segments thereof. The magnitude of fluctuations, which increases as a function of distance is inversely proportional to correlation between polypeptide units fluctuations of subsequent residues. The range over which motional correlations persist defines polypeptide segment length (248).

Fluctuations of a vector may be represented by its diffusion at a finite distance - the larger distance, the longer will be timescale associated with vector diffusion. The probability of vector displacement by a distance $d = |x - \mathrm{x}_0|$ is described by a diffusion propagator, which defines a probability amplitude of vector displacement over a distance d during time interval $\Delta\mathrm{t}$. If distribution of these fluctuations may be described by a Gaussian, for unrestricted motion in three dimensions the propagator is defined by:

$$p(x, x_0, \theta, t) = \frac{1}{8\sqrt{\pi (Dt)^3}} exp\left(-\frac{x^2 + x_0^2 - x x_0 cos\theta}{4Dt}\right) \tag{5.1}$$

For restricted motions additional factors are necessary to describe diffusion process (249). According to fluctuation-dissipation theorem, the diffusion coefficient is defined by energetic balance between parameters describing stochastic fluctuations giving rise to the motion in relation to opposing it viscous drag of the solvent. For a finite size spherical particles in a liquid medium $D = kT/m\beta$, where β defines friction coefficient of particles with medium in which the particles are immersed in. In principle, frictional forces acting on reorienting peptide plane to large extent should be proportional to size of its side-chains and the ones of neighboring residues reflecting solvent dependent and independent contributions to (macro)molecular friction respectively.

Any mammalian cell carries within its interior hundreds of g/L of macromolecules, which constantly interact with each other transiently modifying their properties and those of surrounding medium. The ensemble of the processes is typically described in terms of biophysical interactions on the basis of sticking and crowding, both of which contribute to maintenance of cellular homeostasis and take part in formation of a so-called quinary structure of protein(s) macromolecular ensemble, or rather, of the native in-cell environment (250). Sticking and crowding describe enthalpic and entropic contributors to thermodynamic equilibrium of the macromolecule-medium system, with the former being dominated by charge-charge, van der Waals and hydrogen bonding interactions, while the latter are described from the point of view of work necessary to incorporate a macromolecule in already existing molecular environment (179, 180, 251). As concentration of macromolecules increases so does viscosity of surrounding medium and likelihood of modifications of protein conformational fluctuations and associated with them function (252–254). The ensemble of processes is typically described either with aid of thermodynamical arguments, which show general influence of viscostiy on stability of folded or unfolded proteins (255, 256), or on physicochemical grounds with viscosity and friction used as main indicators of inter and intramolecular interactions in the crowded *milleu* (20, 178, 181).

Unfolded proteins have been extensively studied from point of view of crowding-mediated modifications to their size, shape and functional properties using single molecule fluorescence approaches with interpretation guided by polymer physics (178, 181). In addition to describing dynamic behavior of the solute, the studies show an impact of macromolecular concentration for a variety of viscogens on viscosity of the solution, which directly reflects frictional properties of polypeptide fragments represented by interactions between solvent, protein, as well as solvent and solute. In addition, a plethora of computational studies demonstrates role of hydration in protein functional dynamics. Within the scope, one may find relationships between the solute surface topology, polarity of its fragments and relative effects thereof on interactions with the solvent, and coupling between reorientational dynamics of the solvent and the solute (142, 143, 257–259). NMR spectroscopy has also contributed to these endeavors in a significant way providing an insight into crowding mediated slow-down of proteins such as GB1, α-synuclein from different macromolecular environments with complexity scaling up to the in-cell (260–269). An elegant example of modifications to reorientation potential of NH-backbone fluctuations by complex macromolecular environment is presented by

Theillet et al. (2016). Following application of several viscogens of different physicochemical nature leads to relative deceleration of NH-backbone dynamics of α-synuclein observed in form of modifications to ^{15}N backbone relaxation rates. Pronounced changes in unfolded chain dynamics for residues in the termini observed from Model-Free analysis of the relaxation rates is presumably associated with crowding induced compaction and potential inter and intramolecular interactions with the surrounding medium. The extent of changes seems to be more pronounced in the native environment (mammalian cells) suggesting additional impact of intermolecular interactions on the relative deceleration of the NH-backbone fluctuations (243). The complexity of interactions found therein, ranging from steric interactions, hydrogen bonding and electrostatic interactions in different media makes it difficult to devise a single expression, that would enable a description of specific and non-specific effects on protein dynamics associated with crowding.

From physicochemical point of view, it is useful to assume that the non-specific interactions between macromolecules of protein and crowder will lead to viscosity-mediated deceleration of protein dynamics (270). When a macromolecule is immersed in a solution, it first modifies the properties of surrounding medium leading to deceleration of the solvent molecules, and with increase of the macromolecule concentration, the surrounding medium begins to influence dynamics of the (bio)polymer itself. Overall, the crowder-induced modifications to the flow around the macromolecules is represented by increase in viscosity, which then stipulates slow-down of overall reorientation of the macromolecules, the latter being typically described by one of the forms of Einsteins diffusion equation (271, 272). More specifically, viscosity effects associated with crowding are related to relative size of crowding agent and that of the solute, or rather effective mesh size of macromolecular network defined by interstitial space for the solute to rotate and translate freely. A framework covering crowding-associated modifications to medium viscosity is encapsulated within a so-called length-scale dependent viscosity model describing at least two additional limiting cases: nano and micro-viscosity (174). As concentration of crowding agent increases, viscosity of any solution to a first approximation increases linearly as long as the amount of macromolecules does not cause permeation of each others hydration layer. Beyond a certain concentration (C*) associated with approaching so-called semi-dilute limit, average distance between macromolecules is so small that hydrodynamic volumes of the macromolecules begin to permeate each other leading to more frequent intermolecular interactions and slow-down of their reorientation rates. At the same time, each viscogen molecule reduces amount of free solvent, defined by ensemble of water molecules not associated with macromolecules at given time point, further enhancing inter and intramolecular interactions, which beyond C* are reflected in more abrupt changes in viscosity-concentration dependence (252–254).

As has been observed previously, incorporation of significant amount of macromolecules into solution containing an unfolded protein leads to perturbation of its dynamic behavior. The highly heterogeneous response of backbone dynamics to viscosity precludes estimation of a general trend with which one could describe the ensemble of processes shaping dynamics of unfolded proteins in crowded native-like environment. Herein, an effort has been made to characterize viscosity dependence of apparent dynamic modes representing average reorientational dynamics of unfolded protein fragments probed by NMR relaxation rates by systematically increasing concentration of viscogen macromolecules. The system (NTail - a 125-residue long domain of Sendai virus Nucleoprotein), which apparent dynamic modes temperature dependence has been studied previously, possesses a transiently folded recognition element (MoRE) alongside with diverse conformational fingerprint associated with compositional heterogeneity of amino acid residues across the sequence (68, 235).

5.2 Methods

Preparations of the viscogen solutions

Concentration of Dextran40 (BioChemica, AppliChem A2249) was estimated by measuring absorbance at 270 nm (A270) on a ThermoScientific Genesys 10uv Scanning spectrophotometer in UV-Cuvettes micro 70 μL. The extinction coefficient of Dextran40 was estimated by probing absorbance of its equilibrated solutions at 270 nm (A270). The solutions have been prepared by dilution of Dextran40 stock with known concentration of the polymer. Briefly, 3 g of the polymer was dissolved in 25 mL of buffer of interest in a volumetric flask overnight at approximately 318K and subsequently diluted to 120 g/L. The obtained stock was diluted to reach concentrations ranging from 12 to 120 g/L of Dextran40. The estimated extinction coefficient yielded 4.96 ml cm^{-1} g^{-1}, so that 100 g/L of Dextran40 corresponded to approximately A270 of 0.5. Highly concentrated solutions of Dextran40 were prepared by dissolving 0.5 g of the polymer in 1 mL of buffer of interest in a water bath (318 K) followed by vigorous mixing and overnight 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, concentration of equilibrated stock of the polymer could not be determined exactly and its value was assumed to yield approximately 400 g/L.

Estimation of viscosity-dependent dynamics from NMR relaxation rates

The intrinsically disordered C-terminal domain of Sendai Virus Nucleoprotein (residues 401-525 UniProtKB accession number Q07097, here referred as residues 1-125) was expressed and purified as described previously (68). Samples for NMR spectroscopy measurements have been prepared in 50 mM phosphate buffer equilibrated at pH 6.0 with 500 mM NaCl and 10% D_2O (NTail). ^{15}N relaxation rates for viscosity-dependent modeling of unfolded protein dynamics was recorded on 3 spectrometers (with magnetic field strength ranging from 14.1 to 22.4 T) and 2 temperatures: 288 and 298K. ^{15}N auto-relaxation experiments have been performed with HSQC-type pulse sequences with water suppression achieved following Echo-Antiecho detection scheme (210). Delays for R_1 and $R_1\rho$ experiments ranged from 40 ms to 1.4 s and 1 to 250 ms respectively. Spin-lock field strength in ^{15}N $R_1\rho$ experiment has been adjusted to 1.5 kHz, with adiabatic half-passage hyperbolic tangent pulses with 3 ms ramps and temperature compensation of up to 300 ms. Amide proton decoupling during relaxation delay of R_1 experiment was achieved using iBurp pulse of 2 ms centered at 8.5 ppm. Heteronuclear nOe measurements were performed with proton saturation and recovery delays in range of 6 to 8 s. ^{15}N CSA-dipole-dipole interference (η_{xy} and η_z) have been performed with Echo-Antiecho detection scheme with delays of 60 and 100 ms for transverse and longitudinal cross-correlated cross-relaxation respectively (195, 211, 212). In all of the relaxation experiments value of inter-scan delay set oscillated between 1.0 to 2.5 s. $R_{1,\text{ water}}$ has been measured using simple saturation recovery pulse sequence with 1000 (1 ms) Sinc saturating pulses centered at water resonance following recovery delay ranging from 1 ms to 12 s and small flip-angle readout pulse (500 ns).

Model-free analysis as a function of temperature and viscosity in dilute to semi-dilute regime for SeV Ntail has been performed for a Model assuming with three dominant contributions to SDF optimizing two amplitudes of motions per temperature, two timescales of motions at infinite dilution and temperature, two activation energies and two proportionality constants relating viscosity-associated slow-down of polypeptide units with relative deceleration of water molecules in presence of the crowder (friction coefficients). In Model-Free analysis of NTail, constant value of fast timescales of motions equal to 50 ps has been assumed, while angle between the long axis of the CSA tensor and the NH-backbone vector (θ) has been optimized for all of the conditions simultaneously. In all cases, NH bond vector length has been assumed to be equal to 1.015Å and ^{15}N CSA magnitude to -172 ppm. Model-Free analyses have been conducted for a total of 105 relaxation rates acquired at four temperatures ranging from 274 to 298K and four admixtures with Dextran40 with final concentration of viscogen of 50, 95, 145 and 195 g/L optimizing 15 to 17 parameters - the two additional parameters denote optimization of parameters describing fast apparent dynamic mode and its relative changes with regard to viscosity of the solution. The optimization of parameters has been performed with Levenberg-Marquardt algorithm.

5.3 Viscosity dependence of dynamic modes

To begin considerations on viscosity-dependence of apparent dynamic modes characterizing ps-ns conformational fluctuations of unfolded proteins, I turn my attention to the HSQC spectra of the protein under investigation at increasing concentrations of the crowder. As shown in the figure 5.1, relative changes of chemical shifts are within 0.025 ppm for the highest concentration of viscogen used, suggesting limited influence of potential specific interactions on unfolded chain reorientation dynamics of the studied system. Remaining non-specific interactions depend on reorientation rate of the crowder, population of bound state and extent of modifications to relaxation rates and chemical shifts of reorienting peptide units influenced by inter and intramolecular interactions of the polypeptide fragments, which are reflected in ^{15}N relaxation rates alongside with their modifications associated with changes to viscosity of the medium. To determine extent of modifications to apparent dynamic modes associated with incorporation of viscogen macromolecules to the medium, a set of ^{15}N backbone relaxation rates of NTail in admixtures with 50 to 195 g/L Dextran40, at two temperatures (288 and 298K) and three magnetic fields (with operating ^{1}H Larmor frequency of 600, 700 or 850 MHz) has been acquired. Figure 5.2 convincingly shows heterogeneous response of spin relaxation rates to crowding for both transverse and longitudinal rates of auto (R_2 and R_1) and cross-correlated cross-relaxation (η_{xy}).

The observed changes in R_2 profiles point to crowder induced deceleration of slowest detectable nanosecond conformational fluctuations represented by J(0). Comparison of R_1 at 850 and 600 MHz and 298K as a function of concentration of viscogen suggests relatively little changes to slow intermediate motions occurring on timescales above 2.7 ns ($J(\omega_N)$ at 600 MHz) with greater changes occurring within timescale window of approximately 1.7 - 2.0 ns ($J(\omega_N)$ at 850 MHz), implying rather compact distribution of intermediate motions with minor influence of reorientation processes with rates slower than 2.7 ns and faster than 5-6 ns for majority of peptide units of the protein. Viscosity mediated deceleration in the WW motif and MoRE leads to changes on timescales above 2.7 ns visible in R_1 profile at 600 MHz suggestive of greater population of correlation times in the timescale window. The crowder-associated decrease in R_1 and increase in nOe values implies deceleration of fast and fast intermediate correlation times with distributions ranging from tens to hundreds of ps, from 50 and 150 ps, towards $L(\omega_H)$ maximum at 14.1T. Comparison of rate of change of heteronculear nOes with respect to viscogen concentration at 14.1 and 20T suggests greater sensitivity to fast intermediate motions to the medium viscosity to the one observed for fast motions - $L(\omega_H)$ sensitivity window at 14.1T is filled faster than the one at 20T.

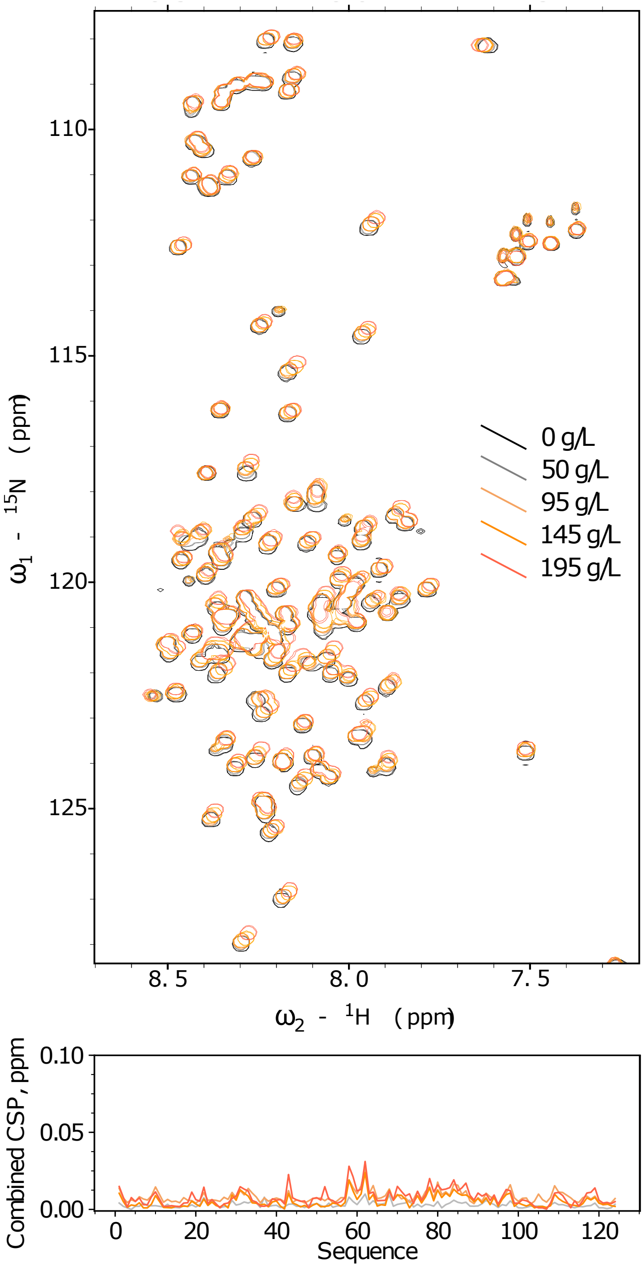

Figure 5.1: ^{15}N-^{1}H *HSQC spectra of NTail in admixtures with Dextran40 (0 - 195g/L) measured at 298K and 20T.*

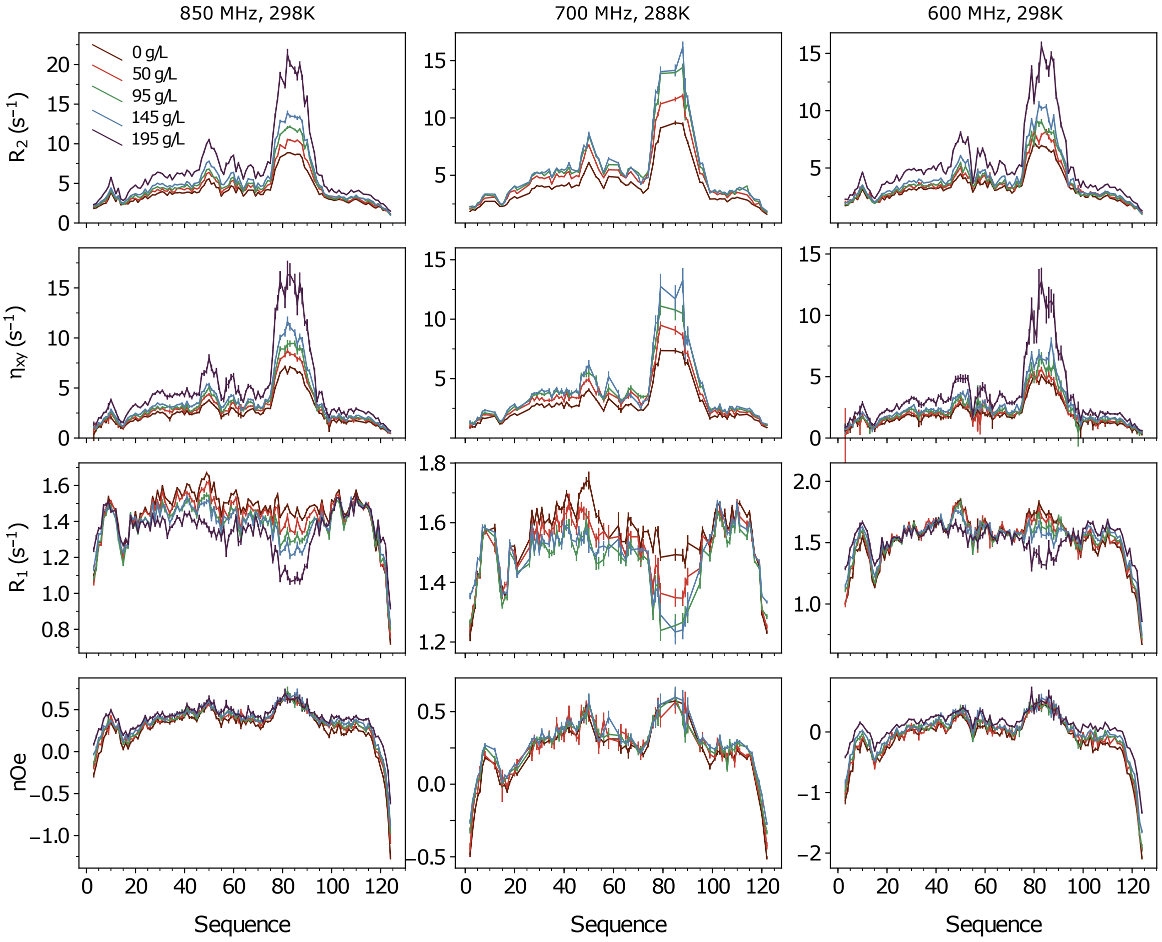

Figure 5.2: ^{15}N backbone relaxation rates of NTail measured in range of admixtures with Dextran40 (0 - 195g/L) at three magnetic fields and two temperatures.

Since friction opposing stochastically modulated conformational fluctuations leads to response of the system dependent on size of reorienting probe (in the studied case represented by polypeptide unit), one needs to find a suitable measure representative of viscogen-associated modifications to the medium. Following derivations of Abragam, longitudinal relaxation rate of water protons is proportional to reorientation rate of water molecules. At certain conditions $R_{1,\text{water}}$ may become biased by contributions from chemical exchange of bound and unbound to macromolecules water molecules, as well as by contributions from translational diffusion which are expected to become more significant at lower magnetic fields with rotational diffusion dominating at higher magnetic fields (77, 273). Therefore, changes in nanoscale viscosity with crowder concentration were evaluated by inspection of $R_{1,\mathrm{water}}$ using a saturation-recovery experiment (figure 5.3).

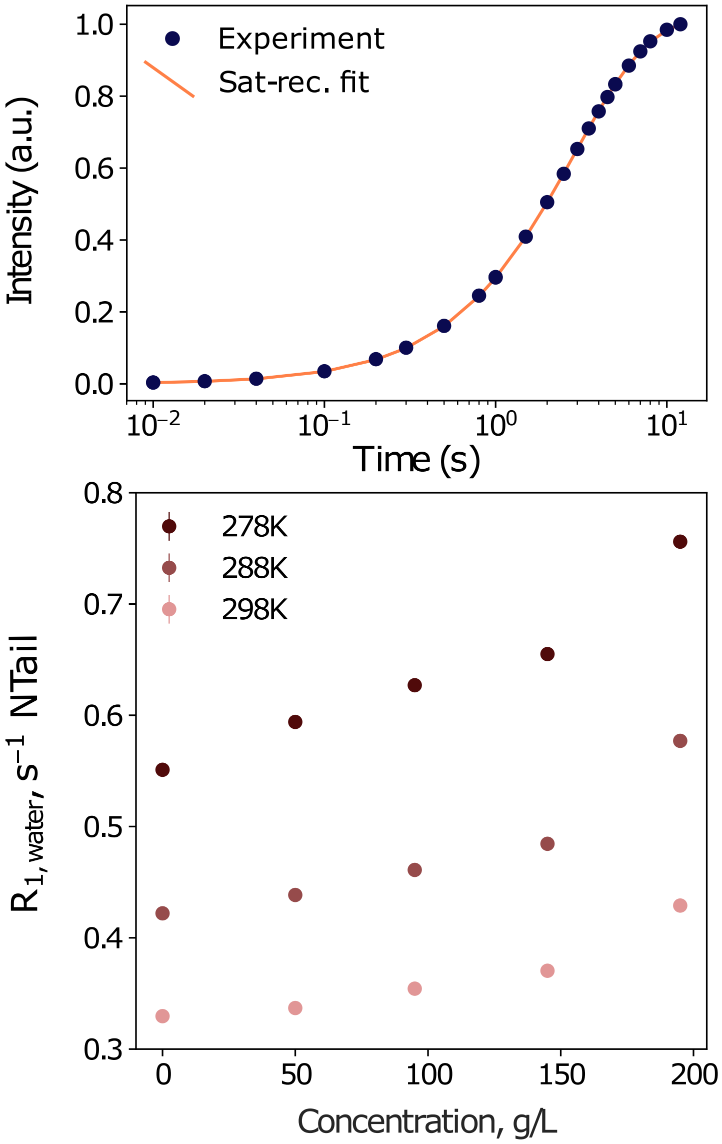

Figure 5.3: Upper panel: exemplary evolution of water proton intensity during saturation recovery experiment. Lower panel: changes of $R_{1,water}$ as a function of concentration of Dextran40 at three temperatures (278, 288 and 298K) recorded at 600 MHz.

The changes of $R_{1,\mathrm{water}}$ with viscogen concentration may be gauged using expression for relative increment of viscosity with crowder concentration ($\rho(C)$), which is approximately proportional to relative deceleration of water molecules reorientation rate (equation 5.2):

$$\rho(\mathrm{C}) = \frac{\eta(\mathrm{C}) - \eta_0}{\eta_0} = \frac{R_{1,\mathrm{water}}(\mathrm{C}) - R_{1,\mathrm{water},\infty}}{R_{1,\mathrm{water},\infty}} \tag{5.2}$$

where $\mathrm{R}_{1,\mathrm{water}}(\mathrm{C})$ and $\mathrm{R}_{1,\mathrm{water},\infty}$ identify longitudinal relaxation of water protons at given concentration and infinite dilution respectively. Admixtures containing up to 145 g/L of Dextran40 reveal linear concentration dependence with similar slope at any studied temperature (278-298K). Above 145 g/L, $\mathrm{R}_{1,\mathrm{water}}$ reflecting relative changes to nanoscale viscosity of the system increases signifying transition to semi-dilute limit at concentration of approximately 180 g/L of Dextran40.

The initial evaluation of viscosity-induced modifications of backbone amide dynamics has been performed with Model-Free analysis of ^{15}N backbone relaxation rates optimizing four parameters for each backbone amide site at each condition separately ($\tau_{i,k}$, $\mathrm{A}_{i,k}$, where i identifies the dynamic modes (1 and 2) and k identifies crowding conditions (0, 50, 95 and 145 g/L) using data recorded at two magnetic fields (600 and 850 MHz) at 298K (eight relaxation rates per condition). Values of estimated amplitudes of motions in different admixtures are very close to the ones estimated for NH-backbone dynamics in conditions approaching infinite dilution (figure 5.4) and vary within range not exceeding 0.1 between different admixtures. The amplitudes describe relative number of correlation times within sensitivity window represented by L(ω). As reorientation decelerates, degree of overlap of distributions of timescales of motions exhibited by polypeptide fragments in relation to the sensitivity at frequencies associated with spin transitions for ^{15}N-^{1}H spin pair, represented by linear combination of Larmor frequencies of the two nuclei, for fast and slow intermediate motions lead to perturbations in amplitudes of motions (Chapter 3.5). Considering distribution of L(ω) within magnetic field strength ranging from 14.1 to 22.3T, deceleration of slow and slow intermediate motions beyond 2.7 ns should decrease amplitudes of intermediate motions with respective amount of information on the distribution now represented by parameters characterizing slow motions (amplitudes and timescales). Similarly, deceleration of fast and fast intermediate motions (occurring on tens and hundreds of ps) beyond 250-300 ps should lead to decrease in amplitudes of fast motions, and respective increase in amplitudes (and timescales) of intermediate motions. The limited influence of viscosity on the estimated amplitudes of motion suggests concerted deceleration of slow and intermediate motions reflecting proportional loss of overlap of the distributions with L(ω) in a manner that would lead to relatively small perturbation of motional amplitudes as a function of viscosity (202, 208).

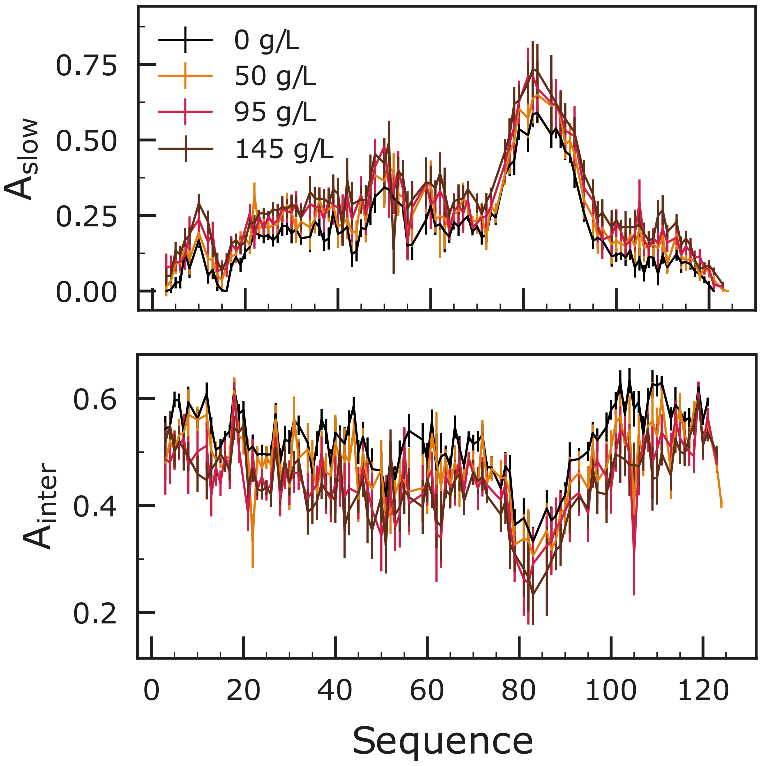

Figure 5.4: Amplitudes of slow and intermediate motions estimated from independent Model-Free analyses for different admixtures with Dextran40 at 298K, each optimizing two timescales, two amplitudes of motions with timescales of fast motions equal to 50 ps and CSA-DD angle θ to 25°.

Therefore, in the subsequent analysis, amplitudes of motions were optimized jointly for all of the admixtures and timescales of motion separately for each of the crowding condition yielding a total of 11 parameters ($\tau_{i,k}$, A_i and θ where i = 1, 2; k = 0, 50, 95, 145 and 195 g/L). As concentration of the crowder increases, so do average timescales of slow and intermediate motions. One could expect that the relative loss of degree of overlap of the distributions with L(ω_N) will be much greater for intermediate correlation times, due to vicinity of the latter to the sensitivity window maximum. Maintaining the relative amplitudes of motions invariant between the admixtures allows to represent relative loss of degree of overlap of the intermediate distributions with $L(\omega_N)$ in a form of average (apparent) intermediate dynamic mode and its rate of change with respect to viscosity of the system. Additionally, it allows to minimize number of parameters allowing for more precise estimation of the apparent dynamic modes. The relative influence of viscosity on timescales of slow and intermediate motions suggests that correlation times representing slow motions (occurring on timescales longer than ω_N^{-1}) are perturbed much more significantly than the intermediate ones. The overall deceleration of slower motions leads to greater population-weighted average slower correlation time and increase in the limiting value of τ_{slow} set by J(0) and statistical independence of motions on estimated apparent intermediate dynamic mode. The phenomenon, in turn gives an impression of increase in separation between the modes of reorientation rates distributions (figure 5.5).

Assuming similar similar loss of sensitivity of $J(\omega_N)$ to slow and intermediate timescales of motions due to increase in viscosity of the medium, rate of change of apparent slow and intermediate dynamic modes with respect to concentration of viscogen is in distinct range with intermediate motions mirroring behavior of the solvent, and crowding-associated changes to correlation times represented by apparent slow dynamic mode are in range three times greater (figure 5.5). Beyond 180 g/L ρ becomes non-linear with concentration and the phenomenon is mirrored by both apparent amide backbone dynamic modes. Considering observations and derivations of Holyst, Cukier and Zwanzig, one may relate relative changes in viscosity of the medium with a site and mode-specific friction coefficient (ϵ_i), which relates reorientation rates of polypeptide fragments with relative deceleration of water molecules (equations 5.2 and 5.3):

$$\tau_i(\mathrm{C}) = \tau'_{i,\infty}(\rho(\mathrm{C})\epsilon_i + 1) \text{ with i} = 1,\ 2 \tag{5.3}$$

where $\tau'_{i,\infty}$ is amide backbone rotational correlation time at infinite dilution. Relating the backbone dynamics to changes in dynamics of the solvent rather than to concentration of the viscogen allows to model relative changes in backbone amide correlation times with viscosity beyond the linear concentration regime (174, 272, 274). To evaluate validity of two independent friction coefficients in representation of crowder-associated modifications to the apparent dynamic modes, Model-Free analysis of the relaxation rate dataset optimizing 6 or 7 parameters ($\tau'_{i,\infty}$, A_i, ϵ_j and θ where i = 1, 2 and j = 1 or j = 1 and 2) has been performed.

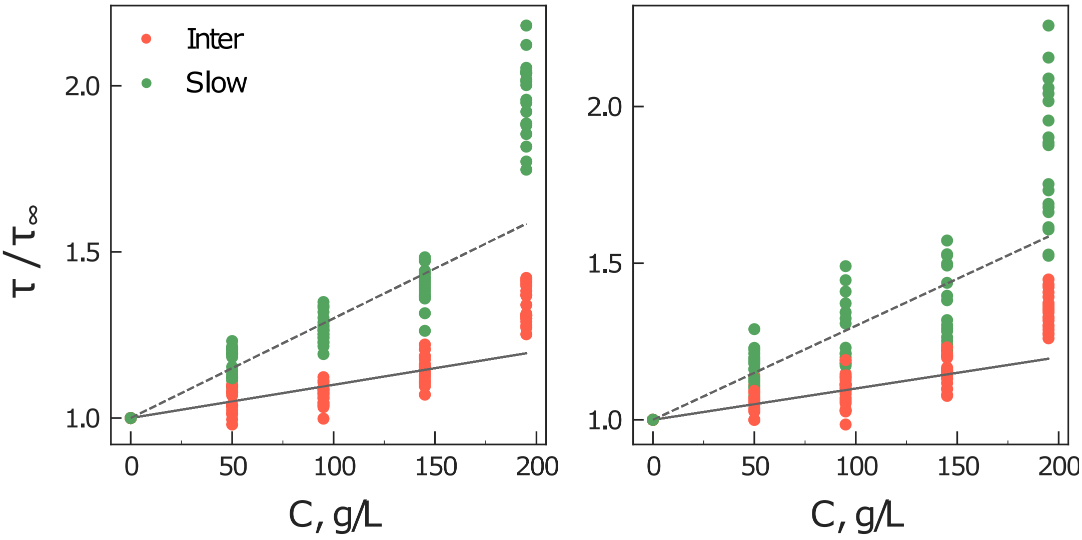

Figure 5.5: Comparison of relative changes of backbone amide correlation times as a function of viscogen concentration with changes in solvent reorientation rate for a few chosen residues within the range 20-40 and 100-120: solid line ρ+1, dotted line: 3ρ(C)+1.

Comparison of estimated χ^2 suggests that two different friction coefficients for apparent dynamic modes (representing intermediate and slow motions) are required to best explain the experimental data (figure 5.6). The assumption of the same friction coefficient for the two dynamic modes leads to priority of optimization of J(0)-dominated relaxation rates, leading to similar data reproduction to the one featuring single friction coefficient (only for τ_{slow}). The values of ϵ_{slow} and ϵ_{inter} oscillate around 2.5 ± 0.8 and 1.0 ± 0.3 respectively. Model-Free analysis optimizing additionally ϵ_{fast} yields values lower than the ones observed for ϵ_{inter} and does not significantly influence quality of the data reproduction. Additional comparison of heteronuclear nOes between the admixtures suggests that changes with viscosity in the apparent dynamic modes occurring on fast tens to hundred ps timescales in dilute to semi-dilute regime are negligible. Therefore, the impact of viscosity on fast motions with average correlation times of approximately 50 ps has been ignored. Encouragingly, in spite of simplicity of the Einstein model (Model 3 with eq. 5.2 and 5.3), the comparison of dynamic parameters at 298K and 0 g/L with results of Arrhenius Model-Free analysis at equivalent conditions yields close comparison with small discrepancies in values of τ_{slow} and A_{inter} (figure 5.7).

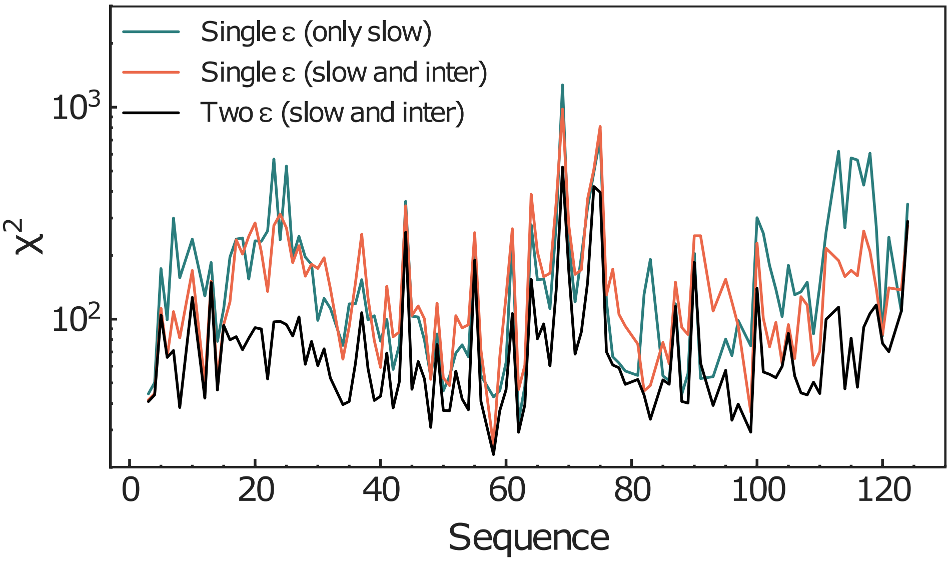

Figure 5.6: Comparison of χ^2 for Einstein Model-Free analysis of relaxation rates of NTail at 298K, 600 and 850 MHz in admixtures with Dextran40 ranging from 0 to 195 g/L, optimizing two amplitudes, two timescales of motions at infinite dilution, CSA-DD angle θ and one or two ϵ for slow and intermediate motions.

As the conformational fluctuations slow down, the apparent dynamic modes representing slow intermediate motions (occurring on single ns) are expected to contribute less to $J(\omega_N)$ resulting in greater amplitudes or timescales of slow motions. In the same time, a portion of correlation times within sensitivity range of $J(\omega_N)$ seems to be perturbed to a smaller extent with apparent dynamic mode remaining within range of single ns, which is contrary to the expectation of greater average timescale allowed by the lower limit imposed by $J(0)$ and adiabatic approximation. The changes are accompanied by deceleration of the fast intermediate motions (occurring on hundreds of ps), as observed for the residues in termini of NTail, likely influencing average value of timescales of intermediate motions leading to smaller degree of perturbation of apparent intermediate dynamic mode and modification to its amplitude. The differences between amplitudes of motions derived from independent Einstein and Arrhenius analysis stem from the accuracy of determination of τ_{slow} and τ_{inter} from the approaches leading to a relative difference in A_{inter} of 0.05 to 0.1. The latter is presumably associated with difference in number of slow correlation times within sensitivity window of $J(\omega_N)$ for conditions featuring different stimuli - temperature or viscosity (figure 5.7).

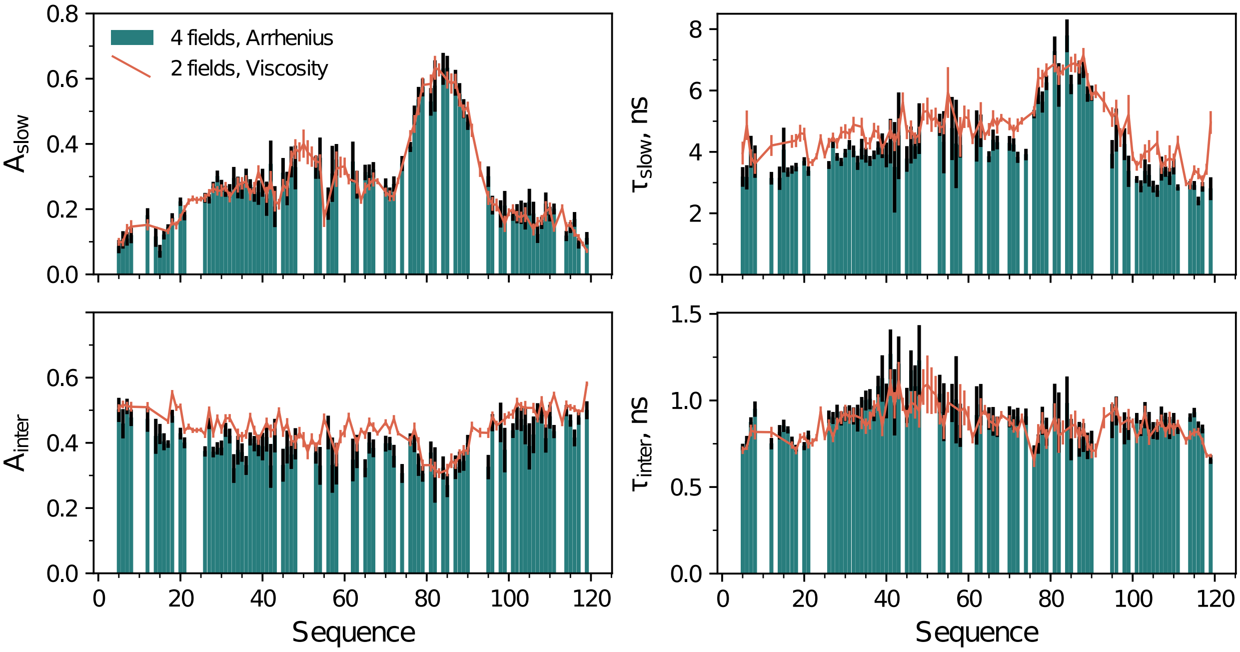

Figure 5.7: Comparison of apparent dynamic modes and associated amplitudes obtained from Einstein (viscosity-dependent) and Arrhenius (temperature-dependent) analysis at 298K and 0 g/L.

The presence of two different friction coefficients may be associated with differences in characteristic length-scale exhibited by reorienting polypeptide fragments, or rather their relative size, opposing the frictional forces acting on them. A body of experimental evidence corroborates relationship between dimensions of the probe and effective viscosity effects, which the probe experiences (219, 275). In the limit, where size of the probe is much smaller than dimensions of viscogen, effective viscous drag experienced by the probe depends linearly on its size. The relative deceleration of water molecules in vicinity of an unfolded chain arising due to hydrogen bonding between polypeptide side-chains and backbone amide protons, as well as surrounding water molecules directly affects reorientation of peptide planes (141, 276). If the intermediate motions are dominated by conformational fluctuations of single peptide planes (represented by transitions within and between Ramachandran minima), the characteristic length-scale associated with their reorientation is comparable to that of water molecules or water clusters, which may explain observation of similar rate of deceleration of polypeptide fragments reorientation with regard to concentration of viscogen to the one observed for water molecules, which corresponds to friction coefficient close to unity. As friction between polypeptide units and the solvent should depend on spatial extent of the motion, which may be represented by moments of inertia of the reorienting polypeptide fragment, the slowest motions presumably dominated by reorientation of several residues in a row (segmental motions) experience greater frictional forces opposing its reorientation (272). The picture in turn, suggests that the conformational fluctuations of polypeptide fragments and viscosity or temperature induced effects influencing its rate are associated with external and internal limitations on reorientation of peptide plane(s).

5.4 Einsten-Arrhenius Model-Free analysis

The established viscosity dependence of the apparent dynamic modes representing conformational polypeptide dynamics probed by the ^{15}N backbone relaxation rates has been combined with previously established pseudo-Arrhenius relationship for timescales of motion. As estimated from longitudinal relaxation of water protons (figure 5.3), the relative changes of solvent viscosity with crowder concentration to a good extent do not significantly depend on temperature within the studied range (278 - 298K). The contribution from heat-activated processes to viscosity is expected to be small, in range of 3-4 kJ/mole, depending on concentration of macromolecules in the medium (277). Together with observation of very similar values of apparent dynamic modes estimated from independent Einstein or Arrhenius Model-Free analysis within a small range of temperatures, one may make an assumption of orthogonality of changes to the modes with respect to viscosity and temperature. The latter allows for combination of viscosity and temperature-dependent datasets to model apparent dynamic modes using Model-Free analysis from ^{15}N-backbone relaxation rates and their viscosity-temperature dependence jointly using Einstein-Arrhenius relationship (equation 5.4):

$$\tau_i(C,T) = \tau''_{i,\infty}(\rho(C)\epsilon_i + 1)\exp\left(\frac{E_{a,i}}{RT}\right) \text{with i} = 1,\, 2 \tag{5.4}$$

where $\tau''_{i,\infty}$ identifies the i-th apparent dynamic mode at infinite dilution and infinite temperature.

^{15}N-backbone relaxation rates dataset encompassing all of the admixtures at 288 and 298K was combined with previously characterized temperature-dependent dataset of ^{15}N-backbone relaxation rates acquired for NTail. Equation 5.4 has been implemented into Model 3 to optimize 15 parameters for a total of 105 relaxation rates ($\mathrm{A}_{\mathrm{i,k}}$, $\tau''_{\mathrm{i},\infty}$, ϵ_i, $\mathrm{E}_{a,i}$ and θ, where i identifies i-th mode and k identifies different temperatures). The calculated dynamic parameters from joint viscosity and temperature-dependent analysis (figure 5.8) are very similar to the ones obtained from separate Arrhenius or Einstein analysis. The energies of activation of the apparent dynamic modes representing slowest motions oscillate within range of 20-30 kJ/mole and its sequence dependence reminds of bell-shape, while the intermediate $\mathrm{E_a}$ reaches local maxima in the transiently folded helical element and in the protein termini, phenomena previously observed in the temperature-dependent analysis of NTail (235). The values of ϵ_i oscillate within previously determined range in separate viscosity-dependent analysis with intermediate mode deceleration rate mirroring increment of water molecules reorientation rate with concentration of the viscogen, while for the slower mode, friction coefficients exhibit marked sequence dependence with values in a range three times greater.

The distribution of ϵ_{slow} across the polypeptide sequence reminds of bulkiness profile of the protein with local minimum for N-terminal fragment of the protein constituting three glycine residues in a row (residues 13 - 15), while increasing in the central and α-helical part of the protein featuring residues of greater moments of inertia or contributing to transient formation of so-called segments as a result of frequent intramolecular interactions (125, 126, 136). The remaining parameters (amplitudes and timescales of motion) are almost indistinguishable from the previous independent Einstein or Arrhenius Model-Free analyses. The overlay of timescales and amplitudes of motions at 288K, 0 g/L and 298K, 95 g/L suggests small-range equivalence of temperature and viscosity-induced modifications to the reorientation rates on both intermediate and slow timescales. The corresponding ρ values for the two comparisons with regard to infinite dilution conditions at 298K yield 1.3 and 1.1 for 288K and 298K, 95 g/L respectively, suggesting greater influence of temperature on water reorientation dynamics than viscogen macromolecules that would lead to similar perturbation to reorientation dynamics of unfolded polypeptide chain on many length-scales.

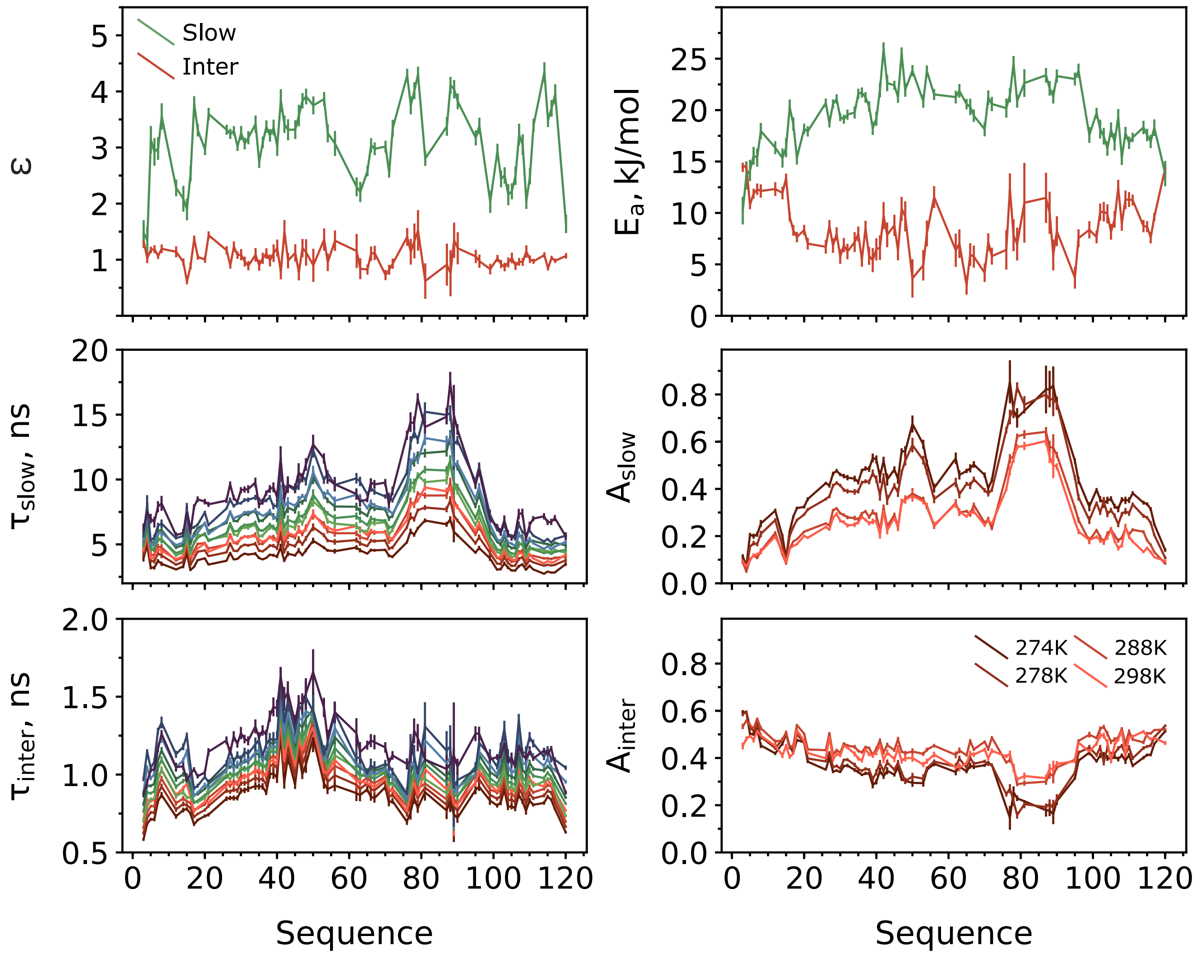

Figure 5.8: Result of the combined viscosity and temperature-dependent analysis: friction coefficients (ϵ), energies of activation (E_a), amplitudes of motion at different temperatures (A_i) and timescales of motion (τ_i). Timescales of motion are presented in the ascending order: 298K, 0 g/L; 298K, 50 g/L; 298K, 95 g/L; 288K, 0 g/L; 298K, 145 g/L; 288K, 50 g/L; 288K, 95 g/L; 278K, 0 g/L; 274K, 0 g/L; 298K, 195 g/L

If ρ represents change to the overall reorientation rate of water molecules proportional to its rotation rate and exchange between free and bound states with the macromolecules and inverse of dwell time spent in the cavities formed by the macromolecule, the temperature-dependent deceleration of water molecules cannot be the only factor determining the timescale of motion of reorienting peptide units. In that case, one could expect that water molecules slow-down in a manner that depends on relative reorientation rate of peptide units proportionally to the dwell time in the bound state. In other words, it is rather unlikely that water molecules would influence reorientation rate of polypeptide units to a major extent, which is direct consequence of the difference in mass between water molecules maintaining contact with polar groups of the macromolecules and mass of the latter. The temperature associated slow-down of peptide units reorientation is associated with increase in frequency or decrease in exchange rate describing intramolecular interactions between neighboring side-chains, due to differences in energetic balance between intramolecular interactions of unfolded polypeptide chain fragments and hydrogen bonding between the former and water molecules in hydration monolayer allowing to observe more abundant motions of correlated nature (segmental motions). The greater response of ρ to temperature than to the concentration of macromolecules implies modifications to the hydrogen bonding propensity between water molecules, i.e. visible in the form of temperature dependence of their clustering propensity, influencing hydrogen bonding with polar fragments of macromolecules (278). The correspondence of the temperature-viscosity trends suggests influence of crowder on the overall unfolded chain dimensions, and the resultant deceleration of conformational fluctuations is correlated with intermolecular interactions between macromolecules of the crowder and the protein of interest, as well as, compaction of the polypeptide. The extent of the modifications is reflected in reorientation rate of water molecules, as has been observed previously (178, 243, 251). At greater crowder concentrations and lower temperatures (i.e. 298K and 195 g/L and 274K, infinite dilution), the relative populations of correlation times representing slow and intermediate motions are expected to be perturbed to a greater extent due to increase in collectivity of the polypeptide fragment reorientation, which may be equally represented by length-scale of the motion (140, 245).

In order to verify the assumptions on the tight coupling between the apparent dynamic modes and reorientation rates of water molecules, one may modify composition of the medium by incorporation of viscogen featuring different size, molecular weight and relative polarity of polypeptide side-chains. To pursue this endeavor, eight more relaxation rates were recorded for 2 different admixtures with 90 and 135 g/L of PEG 10 kDa, 298K at 600 and 850 MHz respectively. Having optimized the 15 parameters per residue representing dynamic modes and their changes associated with modulation to viscosity or temperature encapsulated within the ^{15}N-backbone relaxation rates of NTail in admixtures with Dextran40 (figure 5.8), and knowing values of ρ for the additional admixtures with PEG (yielding 0.16 and 0.25 for 90 and 135 g/L respectively), the relaxation rates were predicted and compared with values at infinite dilution (figure 5.9). If the estimated friction coefficients from Model-Free analysis of ^{15}N-backbone amide relaxation rates of NTail in admixtures with Dextran40 represent relative slow-down of peptide units associated with binding of the unfolded chain to macromolecules of the crowder, one should expect modifications to the extent of modulation of polypeptide reorientation rates associated with intermolecular interactions with crowder of different size influencing rotational correlation time of the latter. Given the four times smaller molecular weight of PEG in comparison to Dextran40, and considering partially branched architecture of the latter, allowing for smaller perturbation in degrees of freedom associated with potential binding to the protein, assuming equal binding affinity between protein and crowder (due to relatively similar polarity of the macromolecules) with binding equilibrium remaining in fast exchange, one could expect up to four times smaller extent in modulation of J(0)-dominated relaxation rates with respect to relative concentration of the viscogen, or, viscosity represented by ρ. The predicted values at any given concentration are almost indistinguishable from the experimental ones, but significantly different to the ones at zero concentration of viscogen. The comparison suggests that deceleration of water molecules is influenced by their interaction with viscogen macromolecules to a similar extent as the relative slow-down in reorientation of polypeptide units on multiple length-scales, and the influence of potential non-specific interactions between macromolecules of unfolded polypeptide chain and the crowder is negligible in the studied range of admixtures. At greater concentration of viscogen, one could expect additional contribution to J(0)-dominated relaxation rates associated with changes in the relative population of slow motions and overall slow-down of reorienting peptide units stemming from their non-specific or specific binding to the macromolecules of the crowder, due to greater rotational correlation time of the latter one associated with small spatial separation (and more frequent collisions) between the macromolecules in semi-dilute or concentrated limit.

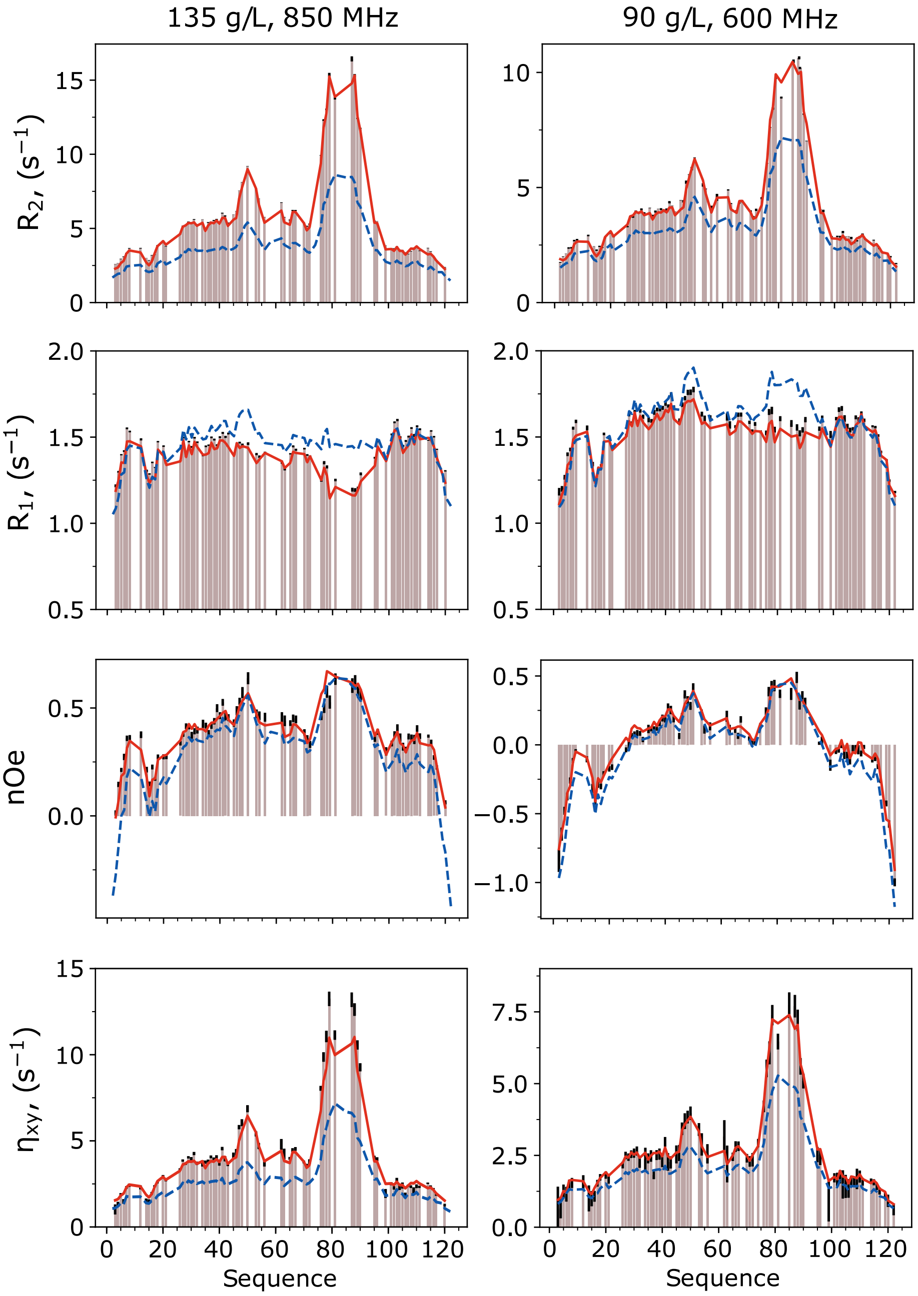

Figure 5.9: Comparison of relaxation rates recorded at two different admixtures of NTail with PEG 10 kDa at 298K (90 g/L at 600 MHz and 135 g/L at 850 MHz; brown and black bars represent relaxation rate values and errors estimated from noise based Monte Carlo simulations within relative uncertainty of peak intensity) with prediction at infinitely dilute condition $\rho = 0$ (blue dotted line) or at ρ estimated from $R_{1,water}$ (0.16 and 0.25 for 90 and 135 g/L of PEG respectively; orange line).

To evaluate robustness of the Einstein-Arrhenius Model-Free analysis, 30% of the relaxation rates have been randomly removed and back-calculated using result of the analysis with remaining 70% in the fitting routine (data not shown). Additionally, the model is capable of estimating dynamic modes as a function of viscosity, at different temperature at which none of the remaining relaxation rates from the crowded dataset are present (figure 5.10). The quality of reproduction of data in the passive fit (figure 5.10B) is virtually identical to the one in the active one (figure 5.10A). In order to determine the minimum dataset required for estimation of all 15 parameters, a significant part of the dataset was selectively removed from the minimization protocol and the resultant dynamic parameters were compared with parameters obtained using full dataset (105 relaxation rates per peptide unit). The reduced datasets featured relaxation rates recorded only at two magnetic fields (600 and 850 MHz) with two crowding conditions (50 and 95 g/L or 145 and 195 g/L) at single temperature (298K) and four or three temperatures at infinite dilution (figures 5.11 - 5.13). Removal of relaxation rates for admixtures at lower concentrations of viscogen does not affect the estimated parameters significantly (figure 5.11) even without half of the temperature-dependent dataset (maximum of 50 relaxation rates per residue). Additional removal of relaxation rates recorded at 274K and infinite dilution (maximum of 38 rates per residue, figure 5.13) does not influence the accuracy of estimation of site and mode-specific friction coefficients, however, significantly overestimates values of the energies of activation of the intermediate mode. The latter is associated with smaller separation between the apparent modes representing intermediate and slow motions in temperature range exceeding 20K, leading to modification in accuracy and precision of determination of intermediate timescales of motion at infinite dilution and temperature. Thus, for an accurate description of the dynamic modes as a function of temperature and viscosity approximately 50 relaxation rates are required. However, reproduction of the dynamic parameters seems to be poor for the dataset containing only admixtures with small amounts of viscogen (figure 5.12). The observation should be rather intuitive and is most likely associated with quality of determination of a slope for linear trend, which is expected to be more accurate on bigger variable range. Application-wise, it is advisable to construct the system with as high concentration of viscogen as it is experimentally possible, that still produces resolvable peak line shapes. In other words, as the apparent dynamic modes approach the values at infinite dilution, the accuracy of estimation of friction coefficients ϵ_i deteriorates. As a rule of thumb, one should aim to characterize polypeptide dynamics as a function of viscosity represented by relative deceleration of water molecules using admixtures with ρ exceeding 0.10.

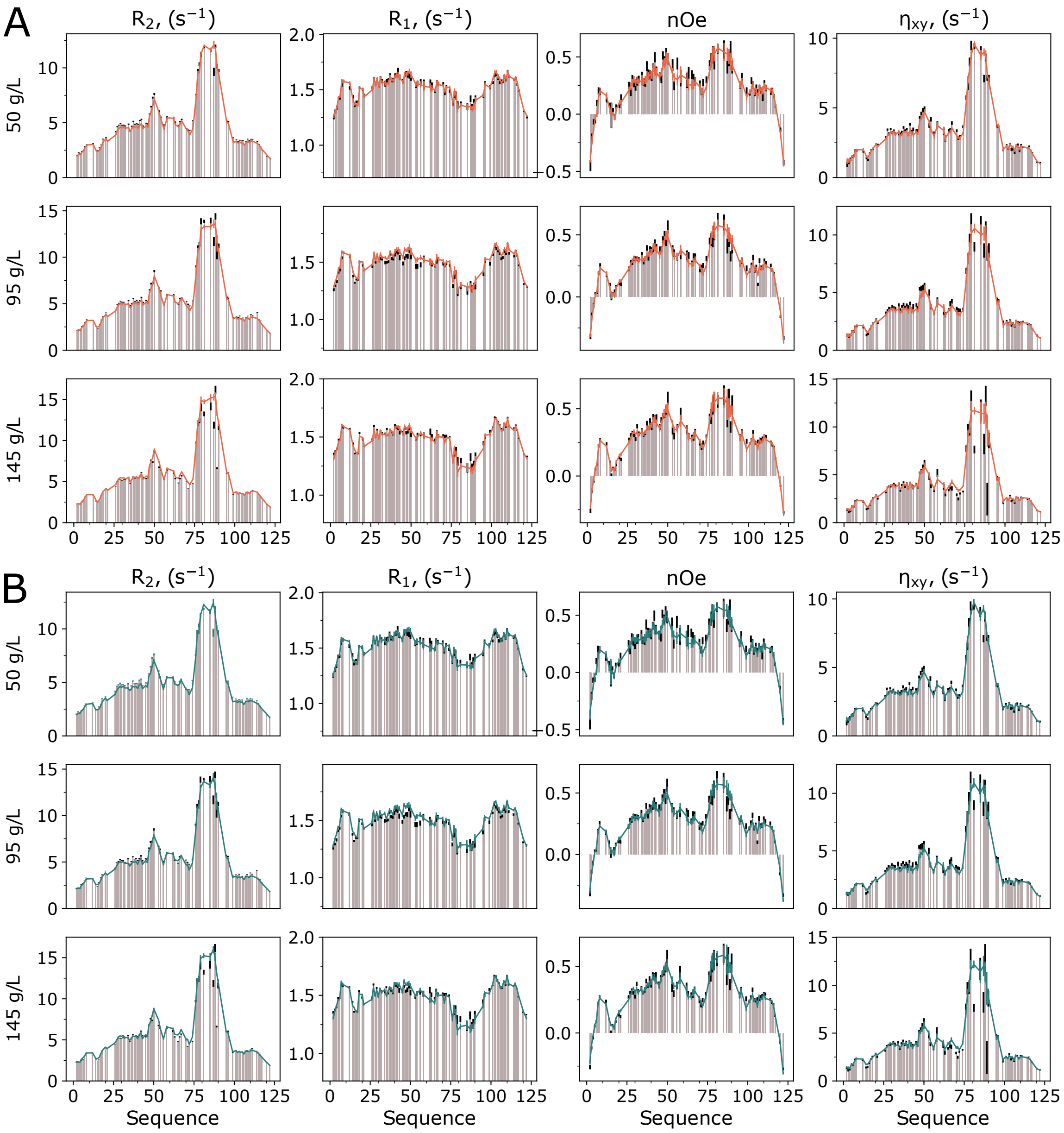

Figure 5.10: Comparison of active fit of ^{15}N backbone relaxation rates of NTail in admixtures with Dextran40 with concentrations ranging from 50 to 145 g/L at 288K (A) with prediction of the same relaxation rates from analysis of data composed of relaxation rates at dilute conditions (0 g/L) at four temperatures, and in admixtures with Dextran40 concentrations ranging from 50 to 195 g/L at 298K (B). Bars represent experimental values, red and green lines mark calculated values from active fit and prediction respectively.

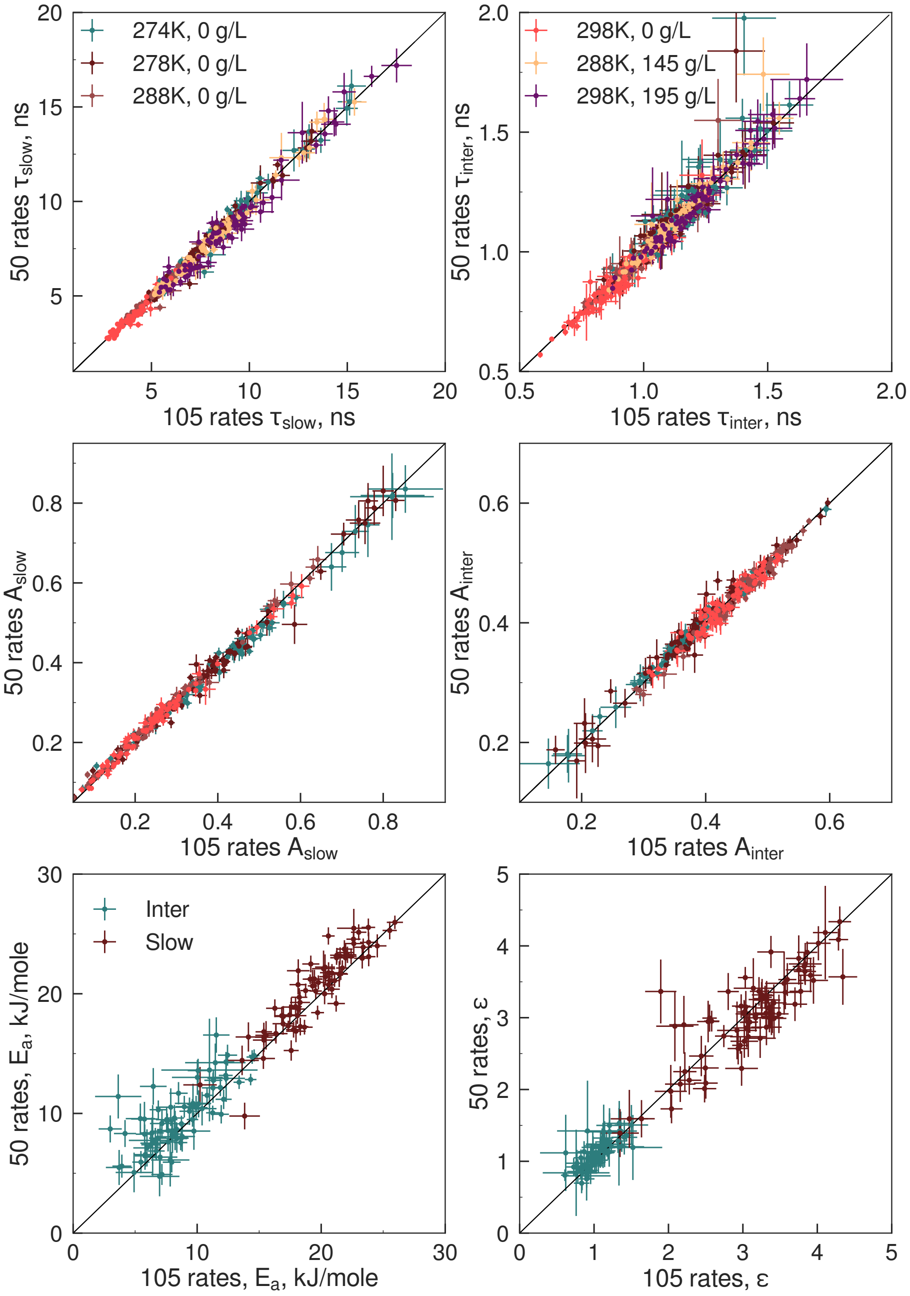

Figure 5.11: Accuracy of determination of apparent dynamic modes and site and mode-specific energies of activation (E_a) and friction coefficients (ϵ_i) with less than half of the dataset: relaxation rates at infinite dilution at four temperatures, relaxation rates at 145 and 195 g/L of Dextran40 at 298K. In all cases only data at two magnetic fields was used (600 and 850 MHz).

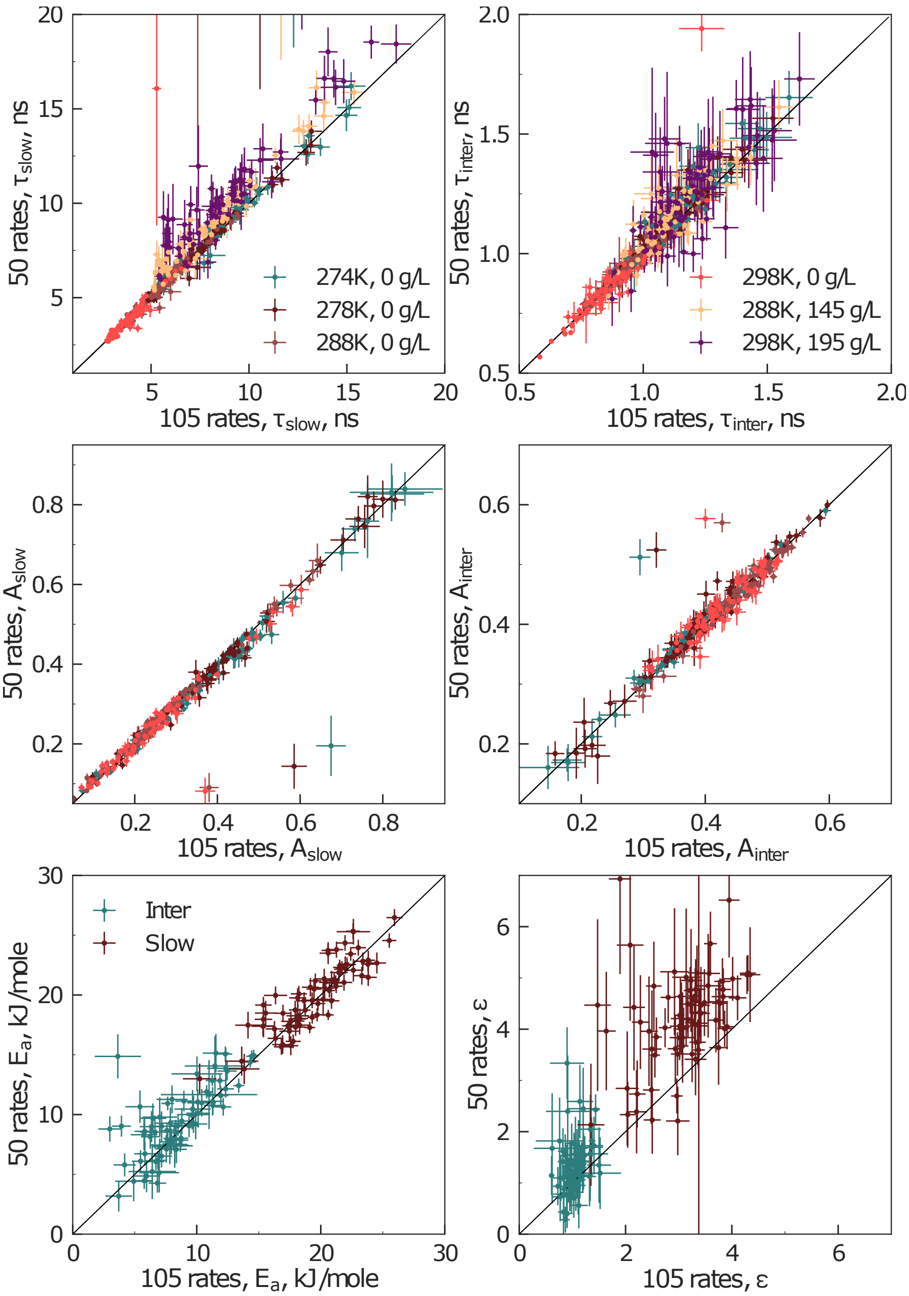

Figure 5.12: Accuracy of determination of apparent dynamic modes and site and mode-specific energies of activation (E_a) and friction coefficients (ϵ_i) with less than half of the dataset: relaxation rates at infinite dilution at four temperatures, relaxation rates at 50 and 95 g/L of Dextran40 at 298K. In all cases only data at two magnetic fields was used (600 and 850 MHz).

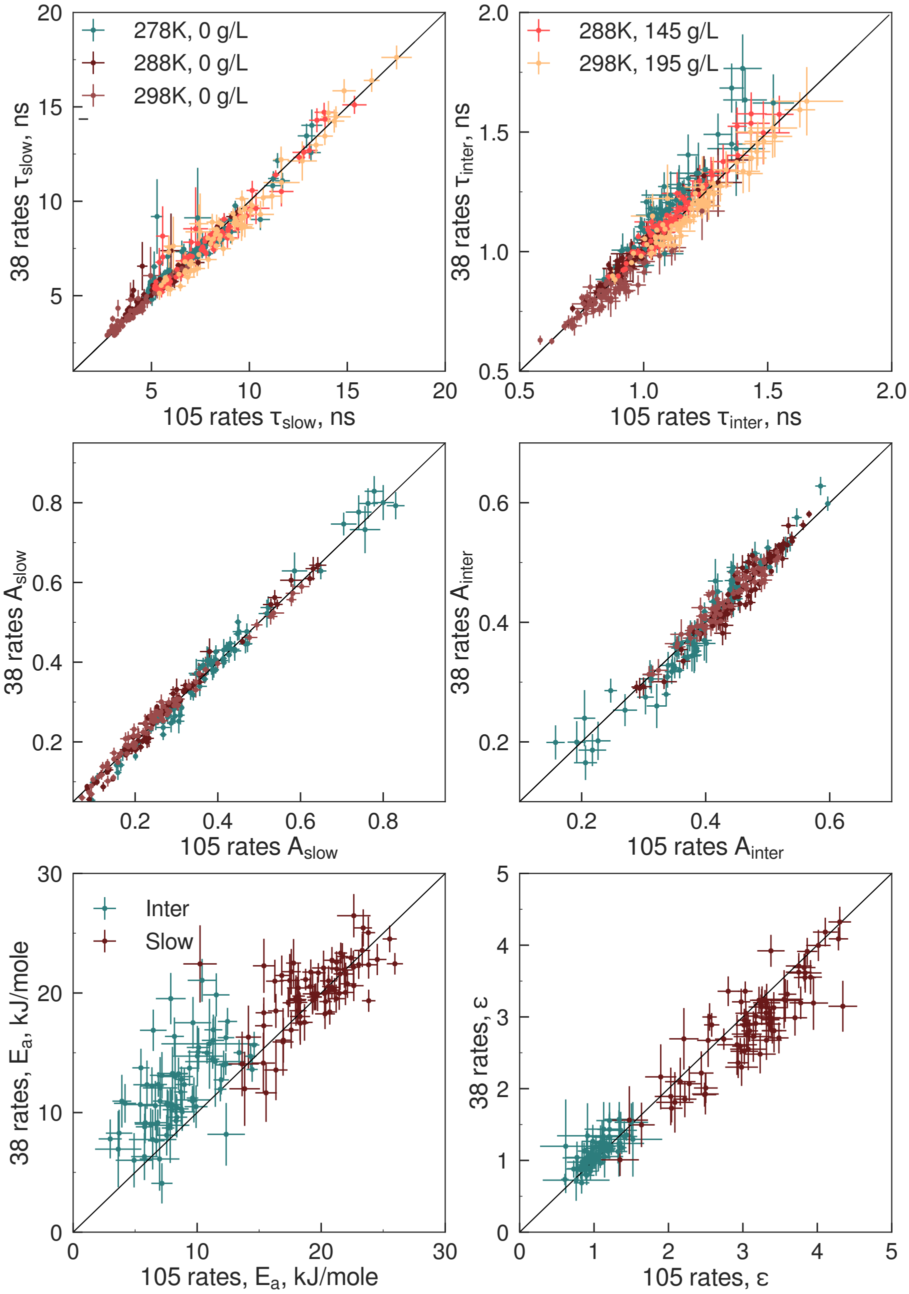

Figure 5.13: Accuracy of determination of apparent dynamic modes and site and mode-specific energies of activation (E_a) and friction coefficients (ϵ_i) with less than half of the dataset: relaxation rates at infinite dilution at three temperatures (278-298K), relaxation rates at 145 and 195 g/L of Dextran40 at 298K. In all cases only data at two magnetic fields was used (600 and 850 MHz).

5.5 Discussion

The analysis of viscosity mediated deceleration of the apparent dynamic modes reveals pronounced changes to correlation times representing slow and intermediate motions exhibited by polypeptide units of unfolded proteins. The overall rate of slow-down appears to decrease with the rate at which the motions occur with greatest deceleration rates observed for the slowest motions. The values of the rate of deceleration of polypeptide unit reorientation are encoded in the so-called friction coefficients, which relate the relative slow-down of reorientation rates associated with conformational fluctuations of polypeptide fragments with relative average deceleration of water molecules. The observed minimal changes in the amplitudes of motions as a function of viscogen concentration suggest limited impact of viscosity on distributions of fast and fast intermediate timescales of motions, or rather, that the overall deceleration of motions occurs in a manner that would compensate the degree of saturation of sensitivity windows (defined by normalized SDFs) of relaxation rate experiments. The importance of dynamic coupling between the solvent and unfolded protein dynamics has been observed in multiple experimental and MD-simulation studies (279–283). The relative proportionality of deceleration of the slow dynamic modes exhibited by polypeptide units of unfolded proteins with respect to deceleration of water molecules re-emphasizes previous observations of high degree of correlation between solvent and unfolded protein segmental dynamics (140, 284–286). The determined friction coefficients sample range of approximately 1.0 for intermediate motions almost without any specific sequence dependence directly mirroring the average slow-down of water molecules, while marked sequence dependence is observed for slower motion friction coefficients. It should be expected that motions exhibited locally, which spatial reorientation does not cause significant displacement of water molecules will experience viscosity-associated slow-down to a lesser extent than more spacious correlated motions of several consecutive residues (125, 136, 218). Indeed, probes of smaller dimensions when compared with size of the viscogen, experience smaller magnitude deceleration associated with friction opposing the thermally modulated stochastic motions, leading to separation of viscosity-concentration dependence into nano, micro and macro regimes (174, 271, 272).

As has been observed for folded proteins, the average slow-down of water molecules in vicinity of the macromolecule is highly dependent on polarity of the polypeptide fragment, and degree of confinement of water molecules within range of its hydration defined by the first hydration shell (143, 144, 287–289). Water molecules in vicinity of a hydrophobic patch will behave differently from those surrounding extended conformations of unfolded chains or protein termini. More precisely, in the first case scenario, one could expect lower probability of finding hydrogen bonding partner with a protein, in turn leading to modifications in hydrogen bond network of the water cluster and slow-down of its rotation and translation (276, 290, 291). In the systems composed of unfolded proteins, reorientation of water molecules in vicinity of polypeptide fragment depends on molecular architecture of unfolded chain, including polarity of atoms belonging to neighboring side-chain groups and size of the latter influencing local topology of the unfolded chain, which is dependent on relative number of degrees of freedom of polypeptide units. The reorientation of water molecules associated with their rotation (without exchange associated with binding to macromolecules) occurs on timescales two to three orders of magnitude faster than reorientation of single peptide unit. Looking through the perspective of water molecules, unfolded chain does not change its topology at timescales associated with hydrogen bonding and reorientation of water molecule (1-100 ps) (142, 145, 278, 292). The reorientation of peptide plane strictly depends on its neighborhood. Thermodynamically, one can think about the stochastically modulated motions as reorientation processes restricted by steric clashes, electrostatic interactions and hydrogen bonds, the latter of which contribute to the tendency of peptide units to form segments. The limitations of peptide segments to execute motion due to frequent intramolecular contacts occurring between and within the polypeptide segment(s) lead to slow-down of motions occurring on different decades of timescales. Limitations to fluctuations over short distances represented by a relative degree of mobility of neighboring peptide units side-chains are encapsulated within friction coefficients (236, 293). The rate of length-scale modulation, which concomitantly increases with number of atoms taking part in the fluctuation seems to be associated with probability of forming hydrogen bonds between polar atoms of side-chains and surrounding water molecules yielding sequence-dependent profile of slower motion friction coefficients similar to the average bulkiness of polypeptide fragments (246, 294). The interpretation reinforces previous assignment of intermediate motions to dihedral angle backbone fluctuations, and slower to segmental motions.

Evaluation of reproduction of the data across different conditions appeals to origin of the orthogonality of the Eisntein and Arrhenius Model-Free analysis of ^{15}N relaxation rates. It would seem intuitive to assume that the frictional effects would be more pronounced at lower temperatures, while the activation energies of the NH-backbone motions should be perturbed by the viscous medium (248, 295). The incorporation of equation 5.4 to Eyrings rate theory predicts that viscosity scaling of reorientation rates of peptide units is temperature dependent with low activation energy for viscous flow over the one for infinitely diluted conditions, on the order of 3-4 kJ/mole for PEG aqueous solutions (277). The temperature dependence of water molecules deceleration rate with respect to the concentration of viscogen (representing so-called nanoscale viscosity), which is effectively probed by water ^{1}H longitudinal relaxation rates is barely visible and does not allow for any quantitative interpretation of the phenomenon. For instance, comparing activation energies of the bulk part of solvent estimated from $R_{1,\ \mathrm{water}}$ across different admixtures one may find that they are relatively similar, ranging from 19.1 to 20.5 kJ/mole in crowding regime from 0 to 195 g/L. Similarly, the rate of change of water retardation with crowder gains relative values not exceeding 1.05 in the temperature range of 278 to 298K. Therefore, in the applied temperature and crowding range, from the point of view of water molecules, which reorientation rates seem to be a reference of choice (*vide infra*), the relationship connecting the rate of NH backbone reorientation across different temperatures and viscosities is relatively orthogonal. In the methodology one is not obliged with choice of a single (inert) crowding agent and routine of accurate concentration measurements, which precision of estimation and preparation thereof is troublesome for highly viscous samples. Instead, one can rely on a short, 20 minute measurement of $R_{1,\ \mathrm{water}}$, which effectively acts as a ruler for nanoscale viscosity of the system, and allows to directly asses relative average slow-down of the solvent. This, in turn allows to change the crowding agent depending on its solubility in the medium of interest or in the presence of undesirable specific interactions at higher concentrations (e.g. above C*) to another system, while still being able to describe the entire dataset with the same formalism. Consequences of the tight coupling between solvent and protein backbone motions are potentially far-reaching. Knowledge of average perturbation of water molecules in a medium of interest and the dynamic parameters characterizing reorientation rates of polypeptide fragments in a distinct, more measurement-friendly conditions allows to obtain a direct information on the unfolded chain dynamics in more complex scenarios.

5.6 Conclusions

The presented results show first, direct experimental evidence of the tight coupling between the backbone motions of unfolded protein and that of the solvent. The developed methodology seems to be highly consistent and orthogonal to the existing ones and allows to couple multiple timescales of motion for dynamic parameters estimated at different temperatures. The addition of viscogen does not specifically modify the reorientation dynamics of polypeptide units that could not be incorporated into the description of relative deceleration of motions with friction coefficient formalism and relative deceleration rate of water molecules. The incorporation of relative changes in number of correlation times with respect to sensitivity of experimental measurement to timescale of motion allows to gather initial insight on distributions of correlation times representing intermediate, slow intermediate and slow polypeptide fragment reorientation rates. Viscosity mediated deceleration of polypeptide units reorientation allows to separate the distributions providing additional insight into partially overlapped distributions of rates of conformational transitions. The relative proportionality of water retardation to average slow-down of reorienting polypeptide units represented by the apparent dynamic modes allows to incorporate perturbation to conformational landscape representing temporal modifications to degrees of freedom of reorienting polypeptide units. The range of values of estimated friction coefficients is related to spatial extent of motion exhibited by the probe, or in other words, conformational restrictions acting on the polypeptide units associated with steric hindrance, hydrogen bonding and electrostatic interactions between neighboring peptide units, in turn influencing the dynamic interplay between polar groups of atoms of the unfolded chain and water molecules within the hydration layer. The possibility of prediction of NH-backbone dynamics occurring on ps to ns timescales as a function of viscosity knowing only one additional parameter, namely the relative increment of solvent deceleration (ρ), presents promising application for quantitative interpretation of protein dynamics in more complex systems, such as reconstituted complexes, membraneless organelles or live cells.

CHAPTER 6

Applications of the unified description

Complex macromolecular environment found in cells, as well as in concentrated phases representing separate macromolecular compartments, significantly modifies dynamics of proteins on multiple length and timescales. The high concentration of macromolecules found therein leads to significant compaction of macromolecular chains of unfolded proteins and greater probability of collisions between them resulting in overall deceleration of their dynamics. The observed from Model-Free analysis of ^{15}N-backbone relaxation rates pico to nanosecond timescales associated with reorientation of polypeptide fragments are therefore expected to be significantly perturbed in the concentrated phases composed of proteins, sugars, as well as fatty and nucleic acids. The developed methodology allows to represent viscosity and temperature-associated changes in conformational fluctuations of unfolded proteins, and has been applied to model systems featuring distinct dynamic fingerprints. The coefficients describing relative deceleration of polypeptide fragments with regard to relative slow-down of water molecules obtained for the model systems corroborate previous observations of the tight coupling between reorientation rates of polypeptide fragments and dynamics of surrounding it solvent. The apparent dynamic modes estimated on the basis of combined Einstein-Arrhenius Model-Free analysis of relaxation rates in dilute phase and overall deceleration of water molecules allow to predict changes to reorientation rates of polypeptide fragments of Map kinase kinase 4 in live *Xenopus oocytes* and Measles Virus NTail in concentrated phase.

6.1 Introduction

In the native environment, any protein is constantly engaged in a myriad of interactions effectively modifying their size, shape, structure and functional dynamics. Advancements in NMR allowed for application of more sophisticated techniques to observe the dynamic processes, functional states and protein binding equilibria in their native environment. Proteins over-expressed in bacterial *E. coli*, micro-injected to *Xenopus oocytes*, electroporated or fusion constructs with cell penetrating peptides are evaluated with regard to their functional state and dynamics in cellulo (296–298). Secondary structures of folded proteins, in general, are not perturbed by the concentrated macromolecular environment found in cell, and apart from minor case-specific observations of their stabilization, their rate of conformational fluctuations in-cell does not deviate significantly from the Stokes-Einstein predictions (250, 270). Disordered proteins, however, stemming from their inherent flexibility and lack of stable secondary structure are expected to experience more pronounced changes to their time-dependent topology represented by conformational sampling on multiple length- and timescales in response to the complex macromolecular environment (299–301). Additionally, the presence of multiple binding partners allows unfolded proteins to engage in a variety of binding equilibria, in turn affecting their dynamics on multiple timescales and regulating their function (302). For instance, Parkinson's related α-Synuclein retains its disordered nature in *E. coli* and mammalian cells, yet, the intracellular environment renders changes in conformational sampling of the proteins leading to restriction of their reorientation potential on multiple length and timescales, represented by slow-down of ps to ns backbone motions reflected in changes to ^{15}N relaxation rates, as well as, significant reduction in their size observed via paramagnetic relaxation enhancement. The differences between relative slow-down of local backbone motions in the cellular environment in comparison to macromolecular settings *in vitro* with equivalent concentration of viscogen suggest involvement of not only weak non-specific binding effects to other biomacromolecules located within crowded cellular interior (quinary interactions), but also potential binding effects with other cellular proteins or membranes as has been observed in a myriad of studies *in vitro* (243, 303). Similarly, tau protein retains its disordered nature in oocytes, where severe line broadening effects testify to the complexity of binding equilibria exhibited by the protein native environment (260, 304). Pronounced changes in ps to ns dynamics are observed for MutX and MutY mutants of GB3 overexpressed in *E. coli*, distinct from those observed in Dextran-buffered solution or cell lysate with up to five fold slow down of the slowest reorientation rates with respect to viscogen-free buffered solution. The deceleration seems to go on a par with significant changes in the folding-unfolding equilibrium rendering minute increase in population of the protein's unfolded state and slow-down of the exchange rate associated with intermolecular interaction (305). In-cell environment is not only crowded, but also rich in other protein binding partners, specific ions and compartments allowing for proper functioning of biomolecular machineries found therein. The functional state of a protein is thus highly impacted by the native environment, as has been shown in a number of studies. For instance, superoxide dysmutase 1 (SOD1), mitochondrial intermembrane space import and assembly protein 40 (Mia40), as well as copper chaperone for cytochrome c oxidase (Cox17), proteins involved in regulation of the redox potential of mitochondria, crucially show diminished capabilities of binding ions necessary for their function in-cell as compared to *in vitro* studies. The differences seem to originate either from misfolded state of the protein *in vivo* or lack of regulator proteins, which shift the dynamic ion binding equilibrium towards free states (306–308). Analogically, some post-translational modifications require binding partners (chaperones) and co-factors found in very small concentrations in cell in order to allow their proper functioning (309–311). The complexity of binding equilibria in cell and the extent to which functional dynamics of proteins is perturbed by heterogeneous intracellular environment still remains unknown (304, 312).

An important aspect of proper functioning of biomolecular machineries is liquid-liquid phase separation of proteins (LLPS), which allows for diverse inter and intramolecular interaction equilibria, whereby physicochemical properties, dynamics and conformational state of a protein is effectively modulated by local concentration of binding partners (313, 314). These spontaneously formed concentrated phases allow for membraneless compartmentalization of proteins allowing for optimal control of physicochemical environment and subsequently stimulate complex biological processes (314–316). An instance, where high concentration of proteins plays a role in cell functioning is nuclear pore complex machinery, where almost gel-like architecture of FG-rich nucleoporins (Nups) allows for fast, coordinated transport through nuclear envelope in spite of remarkably weak binding affinities (K_d on the order of mM) to importins which mediate the transport (72, 317). The binding between the Nups and karyopherins is remarkably specific and undergoes without major conformational changes to the Nups with highly localized perturbations in ps to ns dynamics of residues belonging to and surrounding the FG repeats as revealed by ^{15}N relaxation experiments. Many viral proteins phase separate together with nucleic acids allowing for formation of viral factories, which accelerate nucleocapsid formation

(318, 319). In more specific cases, binding of proteins or nucleic acids may trigger or block phase separation of proteins. For instance, binding of Kapb2 (nuclear transport protein) to tyrosine-repeats, RGG motifs of FUS PY-NLS abrogates its phase separation allowing for its efficient transport to cellular nucleus. The inter and intramolecular interactions of FUS family proteins may further affect dynamic equilibrium of miRNA silencing complex (miRISC) allowing for efficient sequestration of miRNA targets and acceleration of their deadenylation (187). High local protein concentration allows for formation of glass-like states of tardigrade-associated proteins allowing to increase their survival probability throughout hydration-dehydration cycles, assembly of functional biopolymers (e.g. elastin) and mitigates formation of neurotoxic proteinaceous species (e.g. tau and α-synuclein) (320–328). Further, RG-rich Ddx4 (DEAD-box helicase) engaged in RNA-processing and regulation undergoes LLPS with significant perturbation of translational diffusion with up to one hundred fold slow-down as compared to the one observed in dilute phase, and modification to local backbone dynamics with average slow-down of tumbling rates of up to five times as estimated from integral correlation time of the slowest motions (S^2 τ_C) (329). This assembly possesses significant population of excited state (up to 30%), formation of which occurs without major changes to conformational space spanned by the protein in either of states, yet with additional significant deceleration of excited state dynamics as observed in modifications to ^{15}N R_2 likely arising due to intra and intermolecular interactions within the concentrated phase (330). Formation of these assemblies, seems to be amino acid sequence dependent with major contributors to onset of phase separation and integrity of the phase originating from cation-π and π-π interactions, where appropriate spacing of aromatic and positively charged residues seems to be one of the selection rules for observation of the phenomenon for unfolded proteins (183–185, 331).

The observation of dynamics for proteins involved in formation of the complex environments by NMR is challenging for a number of reasons. The concentration of protein inside droplet reaches astounding numbers of 300 to 400 g/L, which necessitates use of considerable amount of material for its production, while formation of a liquid droplet itself may occur in conditions for which protein aggregation may be thermodynamically favorable due to sufficiently high (local) protein concentration. Similarly, viability of cells in standard NMR tubes is limited by availability of nutrients and hampered gas exchange due to high local concentrations of cells per unit volume (332–334). NMR spectra of phase separated proteins with mixtures of concentrated proteins interspersed in surrounding dilute medium, as well as collapsed, concentrated phases or those coming from in-cell exhibit pronounced inhomogeneous broadening due to presence of paramagnetic ions or magnetic susceptibility differences. The (macro)molecular environment in concentrated phases and in-cell, featuring variety of binding partners allows proteins to engage in multitude binding events causing deceleration of their reorientational motions and interaction-associated chemical exchange. The latter occur on multiple timescales and contribute to transverse relaxation rates, in turn leading to severe line broadening and diminished peak intensities. Combined with finite concentration of protein of interest in the cellular environment and finite packing volume of cells in NMR tubes, estimation of the apparent dynamic parameters describing unfolded protein conformational fluctuations in-cell becomes a challenge (270, 316). Ability of prediction of intrinsic dynamics of proteins in the highly complex environments should in principle allow for fast and efficient interpretation of more complex protein binding equilibria in cell. Herein, the proposed in the previous chapter unified description has been applied to quantify and predict the local dynamics of unfolded proteins both in cellulo (Map kinase kinase 4) and in concentrated phase (Measles Virus NTail).

6.2 Methods

Unfolded protein dynamics in cellulo

The intrinsically disordered regulatory domain of MKK4 (residues 1-86, UniProtKB accession number P45985) was expressed and purified as described previously (70). C-terminal fragment of Measles Virus Nucleoprotein (residues 400-520, UniProtKB accession number Q89933) was expressed and purified as described previously (335). *Xenopus oocytes* were prepared as reported previously. 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 $MgSO_4$, 88 mM NaCl, 2.4 mM $NaHCO_3$, 16 mM Hepes and pH 7.4). Cells were then washed several times and for at least 20 minutes with the same buffer without collagenase and stored in the modified Barth's solution with calcium (0.41 mM $CaCl_2$ and 0.3 mM $Ca(NO_3)_2$). Oocytes in stage V and VI were manually selected and stored overnight at 292 K in the modified Barth's solution with calcium supplemented with 100 U mL^{-1} of penicillin, streptomycin and 100 mg/L of gentamycin. After overnight incubation at 292K, 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 labeled proteins in this pole. Microinjection was performed manually using a Nanoject II (Drummond), and 250 oocytes were each injected with 50 nL of protein solution. Concentrated solution of MKK4 (4 mM) has been micro-injected to the animal part of oocytes leading to a final concentration of approximately 50 - 80 μM in the observation volume. The injections have been performed by Justine Magnat (336). *in vivo* measurements were performed in 5 mm Shigemi tubes on spectrometers with 600, 700 and 850 MHz operating ^1H frequency equipped in TCI or QCI Bruker cryoprobes. Prior to each experiment, supernatant of the oocyte suspension was collected and single-increment ^{15}N HSQC and $R_{1,\ \mathrm{water}}$ 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% D_2O in the final volume. Static field homogeneity was optimized on buffer in the same Shigemi tube. The NMR experiments started four to five hours post-injection at 292K to ensure optimal conditions for live oocytes. Typical set of experiments included single increment ^{15}N HSQC for detection of overall protein signal, water proton saturation recovery (30 min), ^{15}N-^1H CSA-DD transverse cross-correlated η_{xy} (16 scans, 1.5 s interscan delay, 128 increments per plane, acquisition time 9h), and ^1H-^{15}N heteronuclear NOE (8 scans, 8 s saturation and recovery delays, 128 increments per plane, acquisition time 9h). The length of dephasing gradients was shortened from 1 ms to 500 μs to decrease losses due to static field inhomogeneity. Experiments were only analyzed from the time window where the cells were viable - within the first fifty hours post-injection. MKK4 assignments were transferred from η_{xy} spectra of *in vitro* samples recorded at 288K. Upon calculation of relaxation rates, peak intensities from interleaved experiments were boot-strap averaged between three subsequent experiments from the same preparation to yield η_{xy} of a single biological replica. The obtained values of η_{xy} were additionally averaged over two biological replicas giving the final in-cell relaxation rates.

Effect of viscosity and temperature on the apparent unfolded chain dynamics estimated from Model-Free analysis of NMR relaxation rates

$R_{1,\text{ water}}$ was measured using a saturation recovery sequence with 1000 Sinc1.1000 1 ms saturating pulses and 20 relaxation delays in range of 1 ms to 12 s. Data was analyzed using NMRPipe, and integrals fitted to exponential recovery using in-house software. In-cell $R_{1,\text{ water}}$ 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 rate as buffer measured *a posteriori* from the same preparation. MKK4 sample for dilute phase measurements has been prepared in Hepes buffer equilibrated at pH 7.0 with 150 mM NaCl and 10% D_2O. Data for the joint viscosity-temperature-dependent relaxation modeling was recorded on four spectrometers with magnetic field strength ranging from 14.1 to 22.4T and four temperatures covering range of 273 to 288K. ^{15}N auto-relaxation experiments have been performed with HSQC-type pulse sequences with water suppression achieved using Echo-Antiecho detection scheme (210). Delays for R_1 and $R_1\rho$ experiments ranged from 40 ms to 1.4 s and 1 to 300 ms respectively. Spin-lock field strength in ^{15}N $R_1\rho$ experiment has been adjusted to 1.5-2.0 kHz with adiabatic half-passage hyperbolic tangent pulses with 3 ms ramps and temperature compensation of up to 300 ms. Amide proton decoupling during relaxation delay of R_1 experiment has been achieved using iBurp pulse (2 ms) centered at 8.5 ppm. Heteronuclear nOe measurements were performed with saturation and recovery periods of 6 to 8 s. ^{15}N CSA-dipole-dipole interference (η_{xy} and η_z) have been performed with Echo-Antiecho detection scheme with delays of 60 and 100 ms for transverse and longitudinal cross-correlated cross-relaxation respectively (195, 211, 212). In all of the relaxation experiments interscan delay was set to value oscillating within range of 1.0 to 2.5 s. ^{15}N backbone relaxation rates of Measles Virus (MeV) NTail acquired under crowded conditions with 0, 37.5 or 75 g/L of Dextran40 equilibrated in

50 mM phosphate buffer at pH 6.0 with 150 mM NaCl and 10% D_2O at 278 and 298K at 600 and 850 MHz. MeV NTail concentrated phase has been obtained by concentration of MeV NTail solution in 3 mm NMR tube. ^{15}N relaxation rates and $R_{1,\mathrm{water}}$ in droplet have been measured on continuous concentrated phase at 298K. The samples for estimation of dynamics in dilute and concentrated solutions of MeV NTail, as well as the relaxation experiments have been prepared and performed by Dr. Serafima Guseva.

Model-Free analysis of ^{15}N relaxation rates performed as a function of temperature and viscosity in dilute regime for MKK4 and MeV NTail has been performed with three dominant contributions to SDFs, optimizing two amplitudes of motions per temperature, two timescales of motions at infinite dilution and temperature, two activation energies and two proportionality constants relating viscosity associated deceleration of polypeptide fragments with relative slow-down of water molecules in presence of crowder (friction coefficients). In Model-Free analysis of MKK4 CSA-DD angle has been set to constant value of 25° and fast timescales of motions to 45 ps. Model-Free analysis of MeV NTail (in dilute phase) additionally included optimization of fast timescales of motions and CSA-DD angle. In all cases, the amide nitrogen-proton vector length has been assumed to be equal to 1.015Å and ^{15}N CSA magnitude to -172 ppm. For MKK4, the EA-Model-Free analyses have been conducted for a total of 80-83 relaxation rates optimizing 14 parameters - four temperatures ranging from 273 to 288K and two or three admixtures with Dextran40 reaching concentrations of 50, 95 and 180 g/L. For MeV NTail, the EA-Model-Free analyses in dilute phase have been conducted for a total of 32 relaxation rates optimizing 12 parameters - two temperatures (278 and 298K) and two admixtures with Dextran40 (37.5 and 75 g/L). Additionally, independent Model-Free analysis of MeV NTail in concentrated phase has been performed, optimizing 5-6 parameters (two amplitudes, two timescales of motions and CSA-DD angle, or magnitude of CSA tensor) with fast component constrained to a constant value of 50, 75 or 100 ps, for 6 ^{15}N relaxation rates. The prediction of in-cell dynamic modes of MKK4 has been performed with linear extrapolation of amplitudes of motions to 292K and combined extrapolation of reorientation rates using Einstein-Arrhenius formalism. Optimization of the parameters has been performed with Levenberg-Marquardt algorithm.

6.3 Protein dynamics in cellulo

The successful application of the model across different macromolecular settings inspired examination of its applicability to complex macromolecular settings. Application of Model-Free analysis with the EA formalism has been performed following addition of 12 ^{15}N backbone relaxation rates of MKK4 in admixtures with Dextran40 (50 and 95 g/L) at 273 or 278K (600 or 700 MHz, figure 6.1) to the temperature dependent relaxation rate dataset (273 - 288K, 600 - 950MHz) acquired for samples in dilute conditions. The observed increment of solvent viscosity (ρ) is about 50% greater at equivalent concentrations of Dextran40 in comparison to the one observed for admixtures with NTail and might be associated with different buffer composition (pH 7.0, 50 mM Hepes for MKK4 admixtures, whilst pH 6.0, 500 mM NaCl, 50 mM phosphate buffer for NTail) affecting effective size of the crowder, its degree of chain expansion, thickness of hydration layer and hydrogen bonding equilibrium between water molecules and the viscogen macromolecules. Similarly to the analysis of NTail, the temperature-dependent dataset of MKK4 (figure 6.1) was combined with viscosity-dependent dataset and modeled jointly using Model 3 with Einstein-Arrhenius relationship (equation 5.4), optimizing 14 parameters for a total of 68 relaxation rates per residue. The results of the fit are shown in figure 6.2. The site and mode-specific values of activation energies, apparent dynamic modes (timescales) and associated with them amplitudes of motion are very similar to the ones obtained for independent Arrhenius analysis (figure 4.4, Chapter 4 section 3). Remarkably, the optimized friction coefficients (ϵ_i) oscillate around 0.9-1.2 for intermediate mode mirroring relative slow-down of water molecules, and around 3.0-3.5 for slow mode for majority of NH-backbone sites of the protein, in spite of overall greater ρ to the one observed for viscogen admixtures with NTail, reinforcing the assumption of tight coupling between disordered protein backbone motions and reorientation rate of the solvent. With deceleration of slow intermediate correlation times, degree of overlap of their distribution with $\mathrm{L}(\omega_{\mathrm{N}})$ decreases. The observed values of slow apparent dynamic mode and its viscogen-induced perturbation suggests its broad distribution of correlation times on timescales longer than ω_{N}^{-1} (at magnetic field strengths exceeding 14T). The viscosity-associated slow-down of intermediate motions leads to decrease in population of correlation times within timescale window for which $\mathrm{L}(\omega_{\mathrm{N}})$ is the most sensitive to (its value is greater than 0.8). The latter is compensated by deceleration of fast intermediate motions. The determined timescales of apparent intermediate dynamic mode representing population-weighted average correlation times within sensitivity window of $\mathrm{L}(\omega_{\mathrm{N}})$, in turn set the limit for timescale of apparent slow dynamic mode and accuracy of estimation of the modes due to statistical independence of motions implicit in the adiabatic approximation, leading to possibility of estimation of small values of A_{slow} that coupled with greater values of τ_{slow} would reproduce J(0). The latter might be associated with relatively broad distribution of correlation times on timescales longer than ω_{N}^{-1} and leads to lower accuracy of estimation of ϵ_{slow}, which high values might be associated with overlap of many distributions of correlation times representing motions with varying degree of complexity. The accuracy and precision of estimation of friction coefficients additionally depends on the range of viscogen concentrations used, or more precisely, on SDF sampling as a function of viscosity. As the use of larger concentrations of viscogen for estimation of friction coefficients increases precision of estimation of the latter (Chapter 5, section 4), one could expect more precisely defined values of ϵ_{slow} for relaxation rate dataset encompassing greater viscogen concentration range.

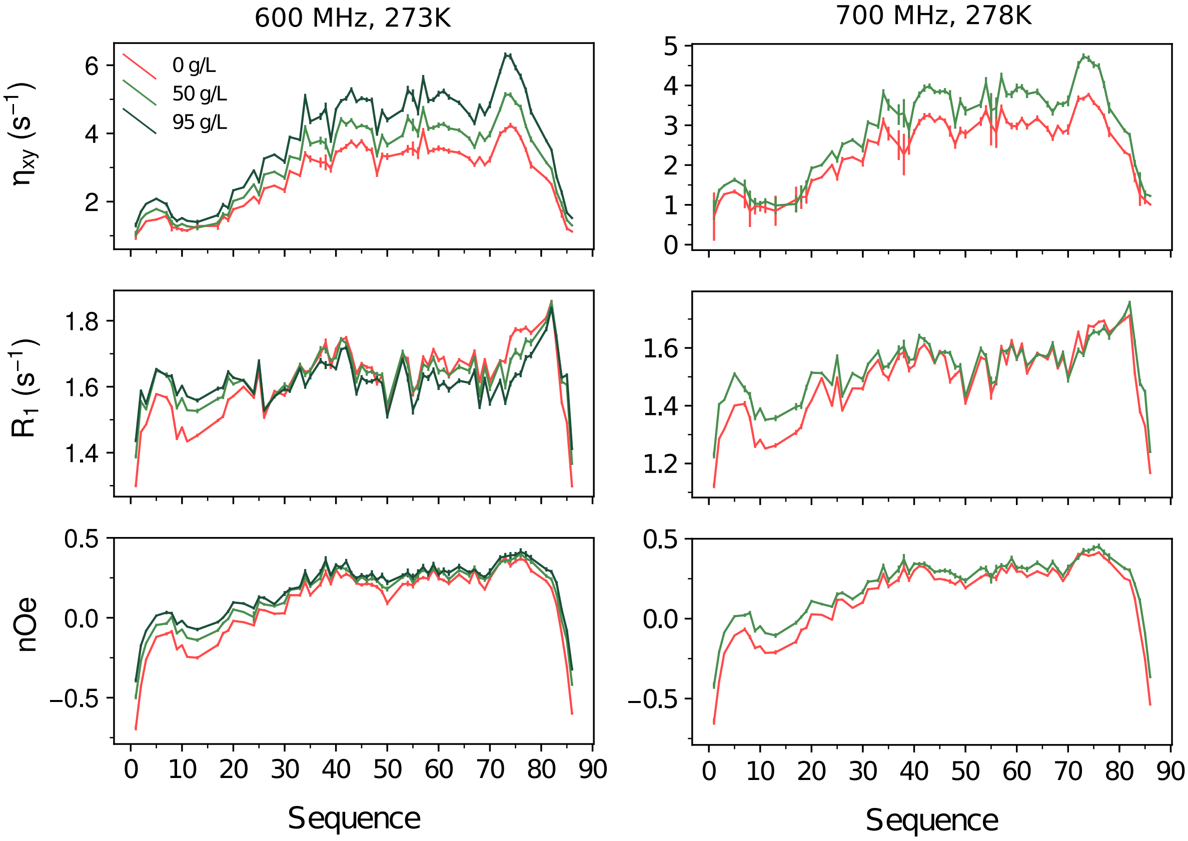

Figure 6.1: ^{15}N backbone relaxation rates of MKK4 for different admixtures with Dextran40, at magnetic field strengths of 14.1 and 16.4T at 273 and 278K.

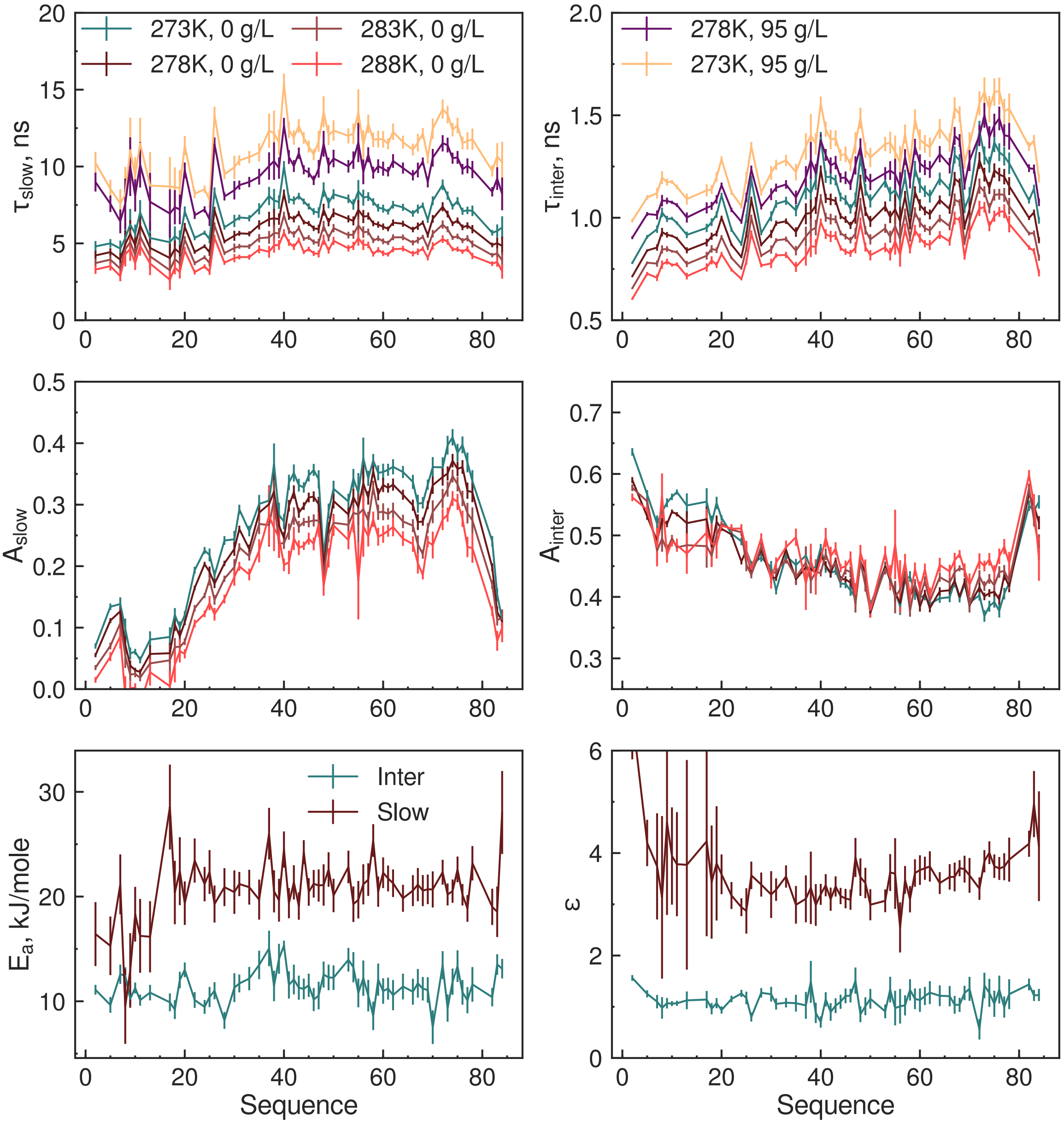

Figure 6.2: Result of the combined viscosity and temperature-dependent Model-Free analysis: friction coefficients (ϵ), activation energies (E_a), amplitudes and timescales of motion at different temperatures and infinite dilution (A_i, τ_i) for 80 relaxation rates with admixtures of Dextran40 with concentrations ranging from 50 to 95 g/L.

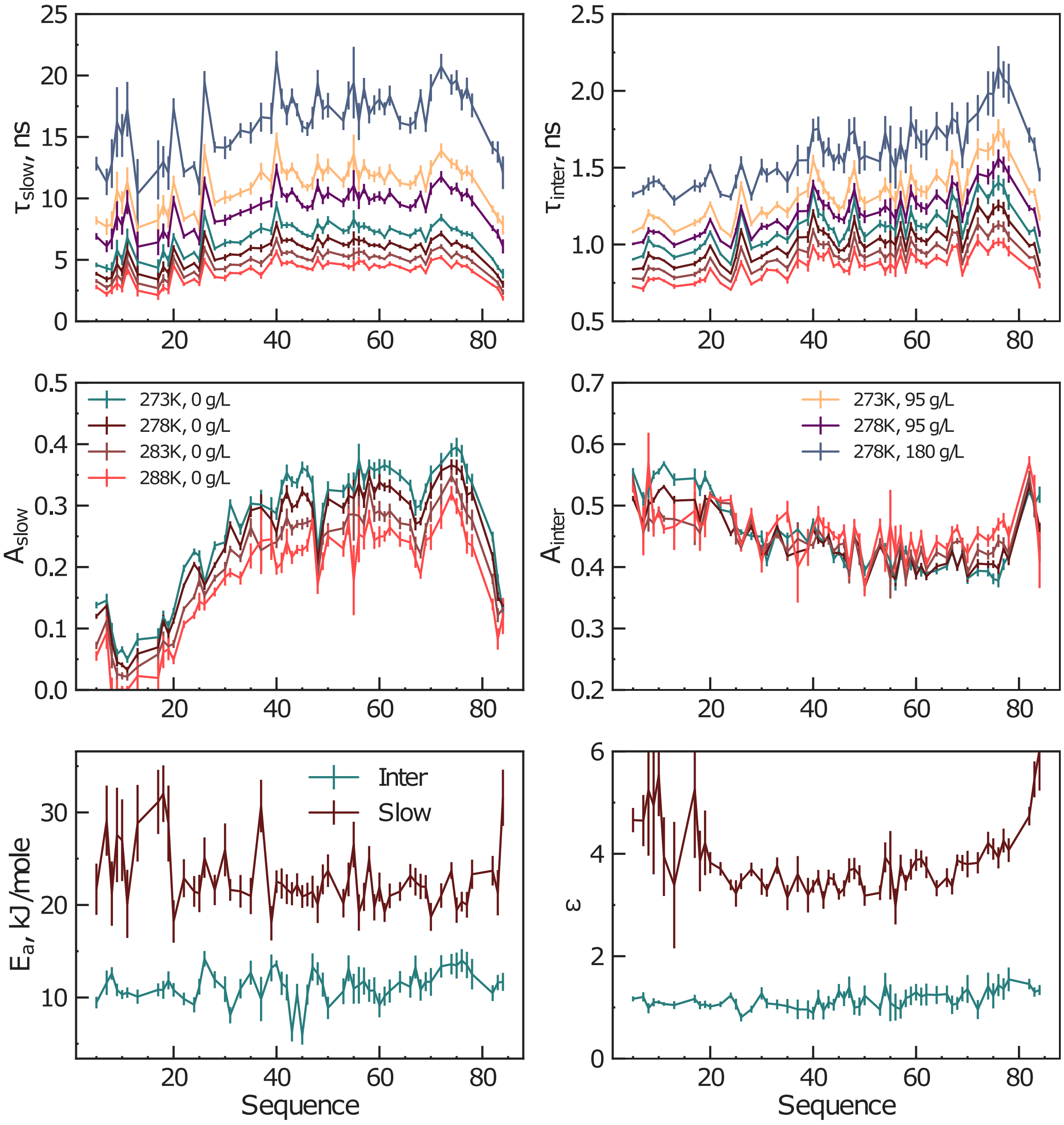

Figure 6.3: Result of the combined viscosity and temperature-dependent Model-Free analysis: friction coefficients (ϵ), activation energies (E_a), amplitudes and timescales of motion at different temperatures and infinite dilution (A_i, τ_i) for 83 relaxation rates with admixtures of Dextran40 with concentrations ranging from 50 to 180 g/L.

Inclusion of three additional relaxation rates measured at 278K, 180 g/L of Dextran40 (representing semi-dilute regime with ρ of approx. 0.75) and 950 MHz does not significantly modify values of estimated parameters of underlying distributions of timescales of motions as a function of temperature and viscosity (figure 6.3). The minor modifications to amplitudes and timescales of motions in-between the expanded dataset with admixtures in semi-dilute and dilute regimes has its origin in significant amount of data representing dynamics in the dilute regime. The apparent minor perturbations to values of ϵ reaffirms the fact, that the deceleration of motions in the range of single to several ns might be represented by relative deceleration of water molecules in semi-dilute regime including flexible parts of MKK4.

In order to estimate the apparent dynamic modes representing reorientation rates of unfolded polypeptide fragments in-cell, one has to optimize experimental conditions associated with acquisition of NMR spectra and relaxation rates. Figure 6.4A shows an overlay of HSQC of MKK4 *in vitro*, and microinjected to oocytes. Although, the viscosity-associated slow-down of protein backbone dynamics combined with magnetic field in-homogeneity of the sample (caused in particular by complex (macro)molecular settings) leads to significant broadening of peak line shapes of more than half of resonances belonging to more bulky part of the protein, the remaining residues retain significant intensity and show very little perturbation to their chemical shifts over time. The comparison of chemical shifts (figure 6.4A) implies significant perturbation to local electronic density surrounding ^{15}N nuclei in central and more bulky part of the protein (residues 25-50) presumably originating from specific binding to partners of the protein in cellular environment remaining in fast to fast intermediate exchange. Additional broadening effects may be associated either with presence of paramagnetic ions or magnetic field inhomogeneity due to complex macromolecular composition of cells, i.e. associated with high concentration of lipids and sugars in the cell walls. To ensure that the observed signal comes from in-cell, the total amount of signal was monitored using ^{15}N-edited ^{1}H spectra throughout sets of experiments. As natural abundance of ^{15}N nuclei is below 0.4%, the observed signal was originating mainly from the injected protein. Indeed, the spectra acquired for oocyte supernatant directly before and after the set of live cell experiments do not reveal significant amount of protein (figure 6.4B). The most intense spectra of protein in-cell were observed starting 24 hours post-injection with initial rise of the signal likely associated with redistribution of injected material throughout the cell, otherwise remaining concentrated in one compartment possibly leading to greater viscosity and in turn contributing to homogeneous line broadening. On the other hand, around 50 hours post-injection ATP signal was almost invisible, suggesting that oocyte metabolism may be stopped at that point (figure 6.4C). Therefore, only relaxation rates from time window between 24 and 50 hours were analyzed. A typical set of measurements on micro-injected MKK4 comprised ^{15}N-edited ^{1}H spectrum (2 minutes) followed by longitudinal relaxation of water protons (30 mins) and a minimal relaxation experiment (η_{xy} or heteronuclear nOe, 8h). The set of measurements was repeated over time period of up to 50 hours at which point the oocytes have probably started to die (*vide supra*).

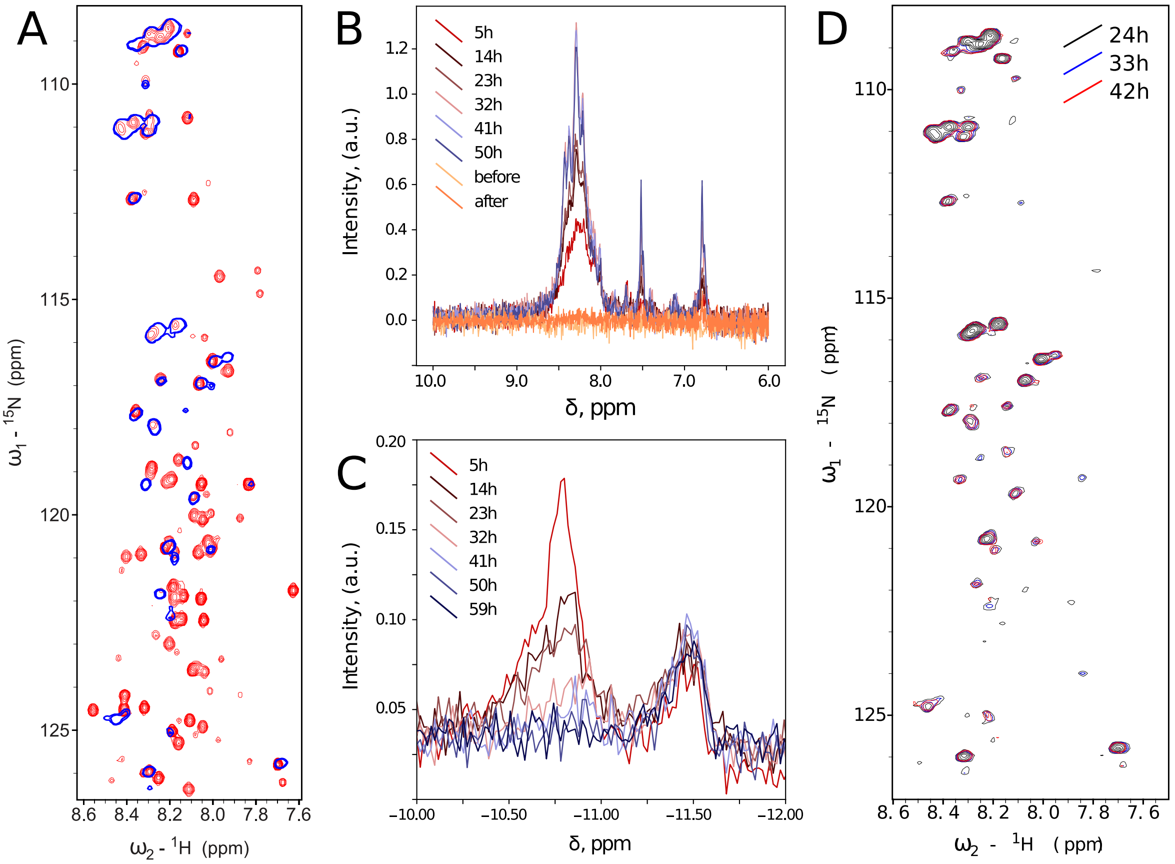

Figure 6.4: Panel A: comparison of ^{15}N-^{1}H HSQC recorded at 292K, 850 MHz in vitro (red) and in live cells (blue, accumulation of 192 scans over time period of 24 hours, starting 24 hours post-injection); panel B: ^{15}N-edited ^{1}H spectra of injected MKK4 at different time points post-injection; panel C: overlay of ^{31}P spectra for chemical shift range characteristic for α-ATP region at different time points post-injection; panel D: overlay of the first planes of the η_{xy} experiment started at 24, 33 and 42 hours post-injection.

After ensuring that intensities (and chemical shifts) of analysed resonances in relaxation experiments do not change significantly over time in time frame from 24 to 42 hours (figure 6.4D), they were boot-strap averaged upon calculation of relaxation rates to yield single biological replica. η_{xy} experiments were performed in up to three independent sets of experiments each with different batch of cells (at 700 MHz), whilst heteronuclear nOes with a single batch of cells (at 600 MHz). Remarkably, the experimental ^{15}N relaxation rates and ^{1}H $\mathrm{R}_{1,\mathrm{water}}$ in-between biological replica were quite similar and subject to additional averaging. The obtained experimental results, together with predicted values at 292K *in vitro* (ρ of 0) and *in vivo* (ρ of 1.00 ± 0.08) are shown in figure 6.5. The $\mathrm{R}_{1,\mathrm{water}}$ in-cell represents an average over multiple compartments containing different amount of material, hence differing in viscosity. Within oocyte interior, one may distinguish animal and vegetal poles, as well as, nuclei, each representing markedly different $\mathrm{R}_{1,\mathrm{water}}$ profiles as determined from MRI leading to assumption that amount of material decreases in the order of nuclei, animal pole and vegetal pole (337). Therefore, the relative increase in $\mathrm{R}_{1,\mathrm{water}}$ should be representative of average nanoscale viscosity increment (ρ), which allows to predict the protein backbone dynamics in cell. The model successfully predicts the temperature and viscosity effects for both slow, and intermediate to fast dynamic modes - J(0)-dominated η_{xy} and heteronuclear nOe dominated by a ratio of $\mathrm{J}(\omega_{\mathrm{H}} \pm \omega_{\mathrm{N}})$ and $\mathrm{J}(\omega_{\mathrm{N}})$. The comparison suggests, that the strong coupling between the solvent and protein backbone dynamics observed *in vitro* allows to represent average perturbation of the protein dynamics arising due to friction opposing reorientation of the polypeptide fragments in complex cellular environment. Indeed, the prediction requires previous knowledge of unfolded protein dynamics as a function of viscosity *in vitro*, but once these are known, only one additional parameter (ρ) needs to be estimated. It is important to note, that the measurement of $\mathrm{R}_{1,\mathrm{water}}$ is an order of magnitude faster than recording ^{15}N relaxation rates at conditions with significantly high value of signal-to-noise ratio (*in vitro*), which is not the case in-cell. This is important, since the viability of cells usually determines the overall acquisition time of NMR experiments *in vivo*. The measurement of $\mathrm{R}_{1,\mathrm{water}}$ may allow initial (coarse) evaluation of the relative macromolecular composition, as well as, prediction of the crowding-induced changes to ps-ns unfolded chain dynamics. Additionally, $\mathrm{R}_{1,\mathrm{water}}$ may supplement any relaxation measurement routine providing an internal way of consistency validation of the dataset in terms of sample (nanoscale) viscosity without significant increase in precious NMR time occupancy.

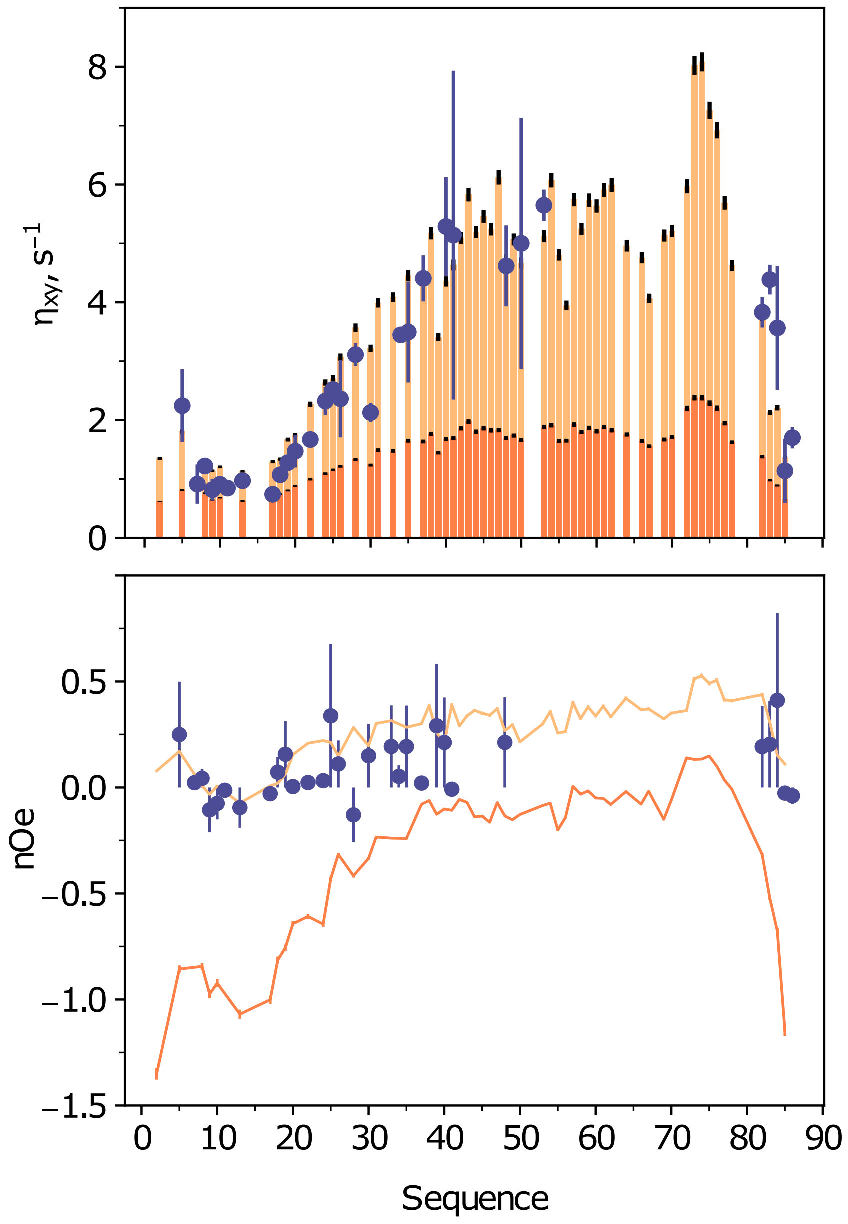

Figure 6.5: Comparison of experimental relaxation rates (η_{xy} and heteronuclear nOe at 700 and 600 MHz respectively) at 292K acquired in oocytes (blue) with prediction at infinite dilution (ρ = 0.0, dark orange) and nanoscale viscosity according to the values of $R_{1,water}$ (ρ = 1.00, pale orange). Experimental values of η_{xy} show average and standard deviation of two biological replica, heteronuclear nOe shows boot-strap average and standard deviation for single biological replica. For prediction of relaxation rates from simultaneous viscosity and temperature-dependent Model-Free analysis, amplitudes were obtained by linear extrapolation to 292K, while timescales of apparent dynamic modes were obtained using Einstein-Arrhenius relationship and value of relative solvent viscosity (defined by ρ) at given temperature.

6.4 Protein dynamics in droplets

Comparably challenging, albeit difficult to estimate are conformational dynamics of polypeptide fragments in concentrated phases. Figure 6.6 shows ^{15}N relaxation rates of MeV NTail measured as a function of crowding at different temperatures ranging from 278 to 298K at magnetic field strengths of 14.1 and 20.0T. Pronounced changes to relaxation rates in transverse plane (R_2, η_{xy}) are observed for polypeptide fragments undergoing fast exchange between folded and unfolded species (spanning residue range of 480-500 MoRE). Together with small differences in R_1 and heteronuclear nOes, the relaxation rates suggest overlapped distributions of correlation times representing intermediate motions spanning range of hundreds of ps to several ns within the protein fragment. The remaining part of the protein possesses several minor features in its dynamic profile, which exhibit high degree of similarity to SeV NTail - relatively faster relaxation rates in the residue range of 435-445 and 510-520 in comparison with the remaining part of the amino acid sequence. As expected, increase in macromolecular concentration or decrease in temperature slows down reorientation rate of peptides planes leading to increased values of R_2. The relaxation rates have been analyzed using Model-Free analysis with Model 3 combined with the Einstein-Arrhenius description for timescales of motion - the quality of fit may be evaluated in figure App.13 (appendix).

Figure 6.7 presents the apparent dynamic modes of MeV NTail in crowded environment *in vitro*. As expected from the previous analysis, the apparent dynamic modes representing population-weighted average of correlation times with respect to sensitivity windows defined by $L(\omega_N)$ and $L(\omega_H)$ oscillate within range of several ns for the apparent slow and intermediate mode respectively with significantly greater values in MoRE, than in other parts of the unfolded chain. The apparent fast dynamic mode shows minor sequence dependence with average timescale of 50 ps and details on the distribution of correlation times representing fast motions (faster than several hundreds of ps) encapsulated in the amplitude of motions. The estimated amplitudes of motions exhibit similar profiles to the ones obtained for SeV NTail with marked differences across the sequence for slow and fast motions. The values of slow amplitudes of motion are overall smaller in comparison to SeV NTail by approximately 0.1 suggesting smaller contribution of slower motions to the overall dynamics of the protein. The activation energies oscillate within the range of values observed previously - around 10-12 and 20-25 kJ/mole for intermediate and slow mode respectively.

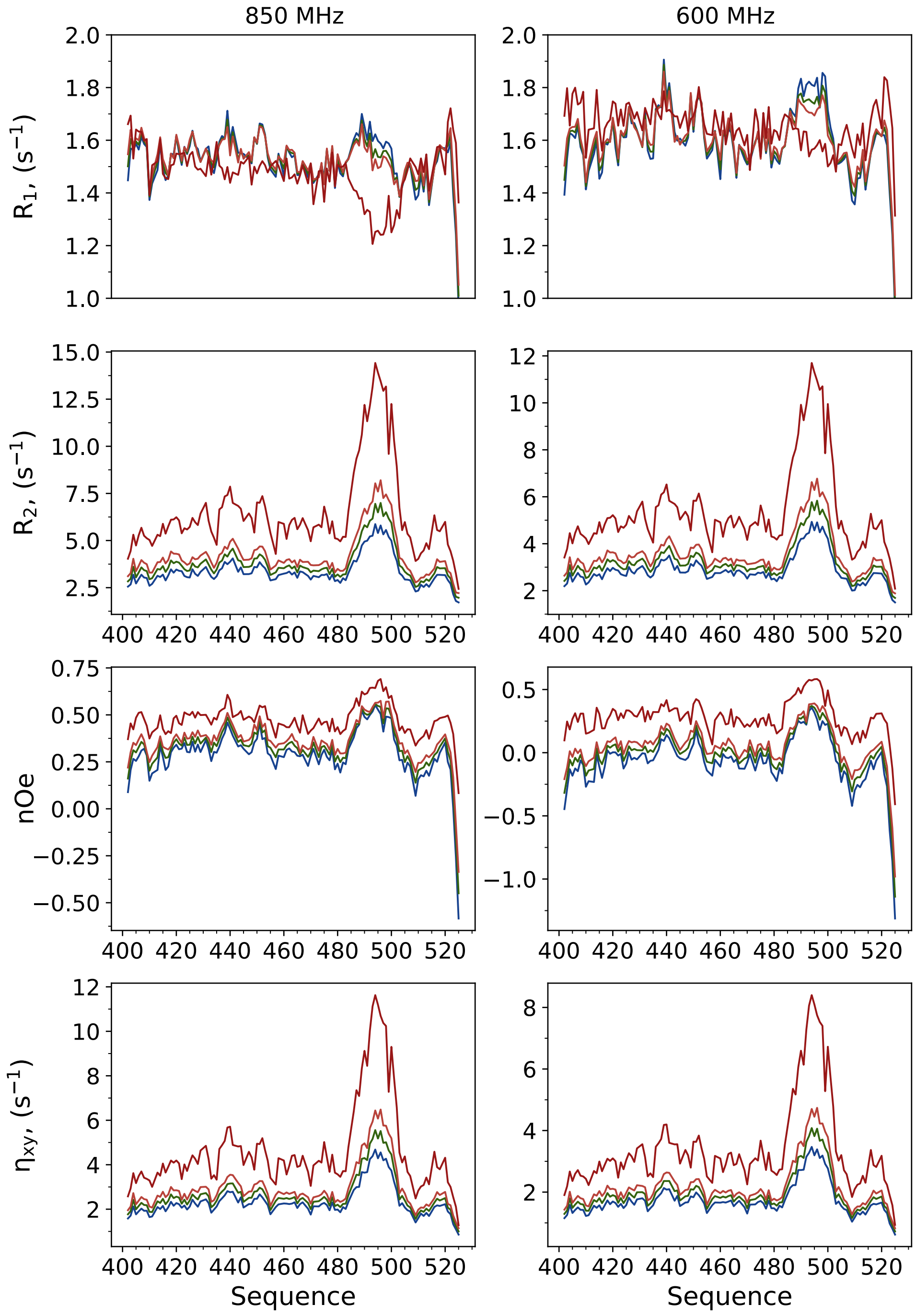

Figure 6.6: ^{15}N backbone relaxation rates of NTail measured in the range of admixtures with Dextran40 0 (blue), 37.5 (green), 75 g/L (orange) at 298K and 0 g/L (red) at 278K and two magnetic fields strengths - 14.1 and 20.0T.

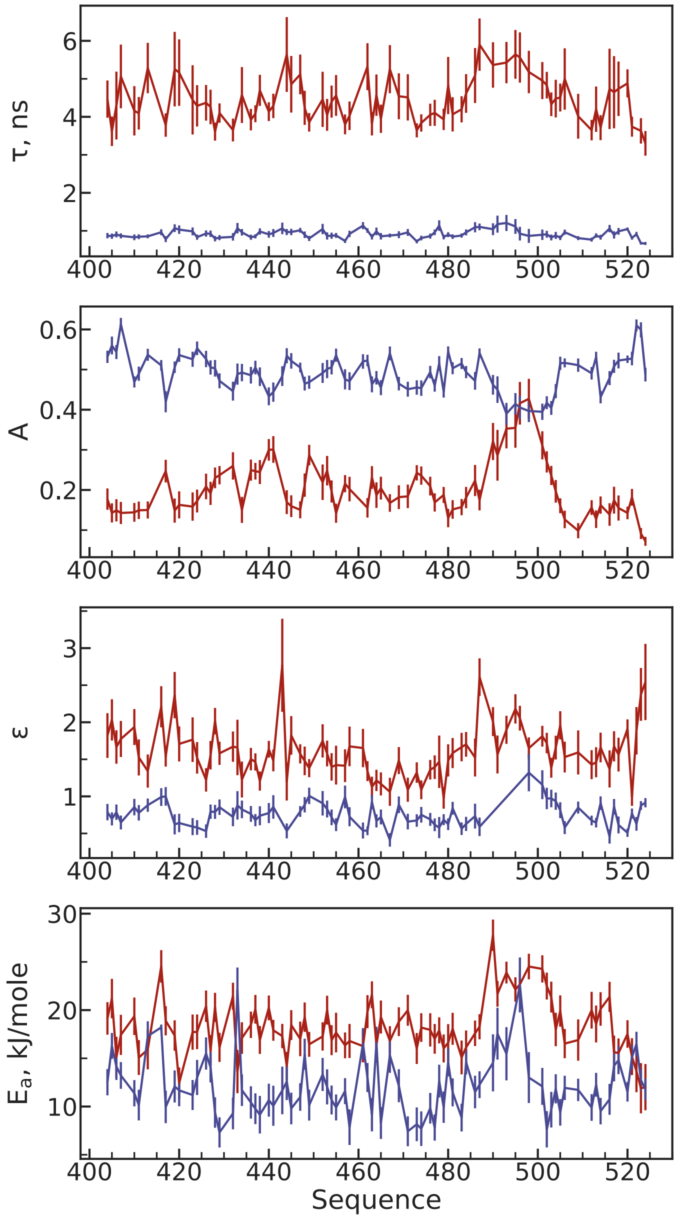

Figure 6.7: Result of the combined Einstein-Arrhenius type Model Free analysis of ^{15}N *backbone relaxation rates of Measles NTail using Model 3 with equation 5.4 as described previously. The spin relaxation rates as a function of temperature and viscosity are represented by site-specific activation energies* ($E_{a,i}$) *and friction coefficients* (ϵ_i)*, as well as, amplitudes and timescales of motions* (A_i *and* τ_i) *at 298K and at infinite temperature and dilution respectively for slow and intermediate motions. Values of the apparent fast dynamic mode have been constrained to 50 ps, while the amplitudes of fast motions have been estimated from the normalization condition of amplitudes of motions. Red and blue mark values of slow and intermediate modes respectively.*

The relatively small difference between energies of activation of slow and intermediate motions might be associated with overlap of distributions of correlation times representing slow and intermediate motions, as has been observed for residues in N-terminal part of MKK4 and the ones in termini of NTail. More pronounced differences in E_a may be observed around the α-helical MoRE. The estimated values of friction coefficients are smaller in comparison to the values estimated for MKK4 or SeV Ntail by a factor of 0.75 to 0.50 for intermediate and slow motions respectively. The discrepancy may be explained on the basis of overlapping distributions of correlation times representing fast and slow intermediate motions occurring on hundreds of ps and single ns respectively. Since ϵ depends on spatial extent of the reorientational motion (i.e. size of the probe opposing friction), as observed in the Model-Free analysis of ^{15}N backbone relaxation rates of SeV NTail, motions occurring on faster timescales should exhibit smaller values of ϵ. If distributions of correlation times representing slower intermediate and slower motions (with apparent modes in range of timescales below and above 3 ns respectively) are overlapped, the resultant ϵ values should correspond to average over rates of change of correlation times with regard to relative deceleration of the solvent for distributions of fast intermediate and intermediate, as well as, slow intermediate and slow motions. The small gap between the friction coefficients (and the energies of activation) suggests greater proximity of distributions representing intermediate and slow motions in central part of the unfolded chain with regard to the ones exhibited by SeV NTail. The relatively greater ϵ_{slow} values for residues located in the protein termini with regard to the ones in central part of the polypeptide fragment is presumably associated with relatively broad distribution of correlation times of slow motions representing stochastic fluctuations of polypeptide units with a number of degrees of freedom. Additionally, the distribution of correlation times spanning range of single ns might be partially overlapped leading to inaccurate estimation of the dynamic modes, as has been shown in Chapter 3 (section 3). Similarly, the estimated timescales of intermediate motions originating from an overlapped distribution of motions occurring on hundreds of ps and single ns representing limitations to reorientation of polypeptide units on smaller length-scales should yield smaller dependence on viscosity.

Assessment of protein dynamics in the concentrated phase have been performed in two ways. First, an independent Model-Free analysis has been performed on the ^{15}N-backbone relaxation rates originating from protein immersed in the concentrated phase (R_2 and nOe at 600 and 850 MHz and R_1 and η_{xy} at 850 MHz) assuming three independent dynamic modes with one of the timescales constrained to a constant value of 50 ps (Model 3, figure 6.8). The quality of fit may be evaluated by inspection of figure App.14. The assumption of greater timescale of fast motions does not improve the quality of the fit, yet, due to the normalization conditions of amplitudes, the use of greater τ_{fast} yields smaller amplitudes of fast motions, in turn influencing values of amplitudes and timescales of slow and intermediate motions (figure 6.9). The obtained apparent dynamic modes oscillate within 20-40 ns and 1-2 ns for slow and intermediate timescales of motions, while amplitudes of motions featuring rather flat profiles with markedly different values to the ones estimated from Model-Free analysis of relaxation rates in the dilute phase. The estimated dynamics of MeV NTail in the dilute phase suggest relatively overlapped distributions of correlation times ranging from hundreds of ps to single ns, as well as, single to several ns. The viscosity-associated deceleration of the apparent dynamic modes leads to their separation in a manner that should be dependent on spatial extent of the motion. It is expected that the distributions of correlation times representing intermediate and slow intermediate motions with modes at hundreds of ps and single ns slow-down to single ns and several ns respectively, while the slowest motions decelerate to tens of ns, leading to greater separation between the distributions.

The estimated from Model-Free analysis in the concentrated phase values of τ_{inter} are presumably associated with sensitivity of ^{15}N R_1 at 20.0T to correlation times in range of 1.50-2.30 ns, peaking at 1.85 ns. Within the assumption of three statistically independent motions represented by apparent dynamic modes in Model-Free analysis, the remaining (non-overlapping with $L(\omega_N)$ maximum) part of distribution of correlation times spanning single ns in the dilute phase may be represented by amplitudes of slow motions, in turn affecting values of amplitudes of intermediate and timescales of slow motions, the latter representing population-weighted average over correlation times spanning range of single to tens of ns. The estimated from the combined EA-Model-Free analysis apparent dynamic modes representative of average timescales of slow and intermediate motions in the concentrated phase (on the basis of ρ) remain in similar range for τ_{slow} and twice greater for τ_{inter}. The intermediate timescales of motions represent an average over distribution of correlation times spanning range of hundreds of ps to several ns.

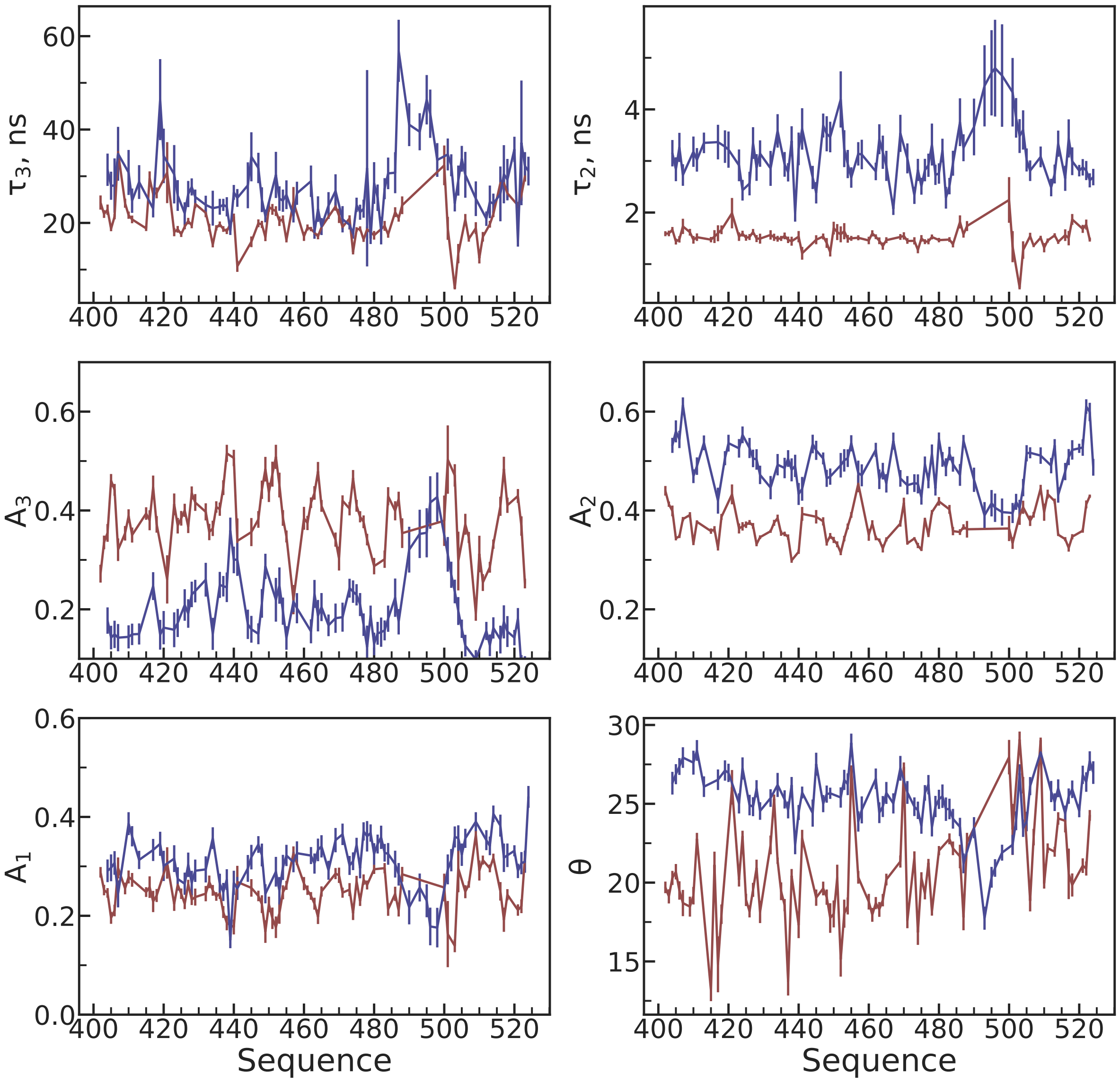

Figure 6.8: Comparison of the dynamic modes estimated using 5 parameter Model-Free analysis of relaxation rates in concentrated phase assuming constant value of fast timescales of motions equal to 50 ps (two amplitudes, two timescales and CSA-DD angle, red) with prediction from the combined Einstein-Arrhenius Model Free analysis and $\rho = 3.44$ from longitudinal relaxation rate of water molecules in the concentrated phase (blue).

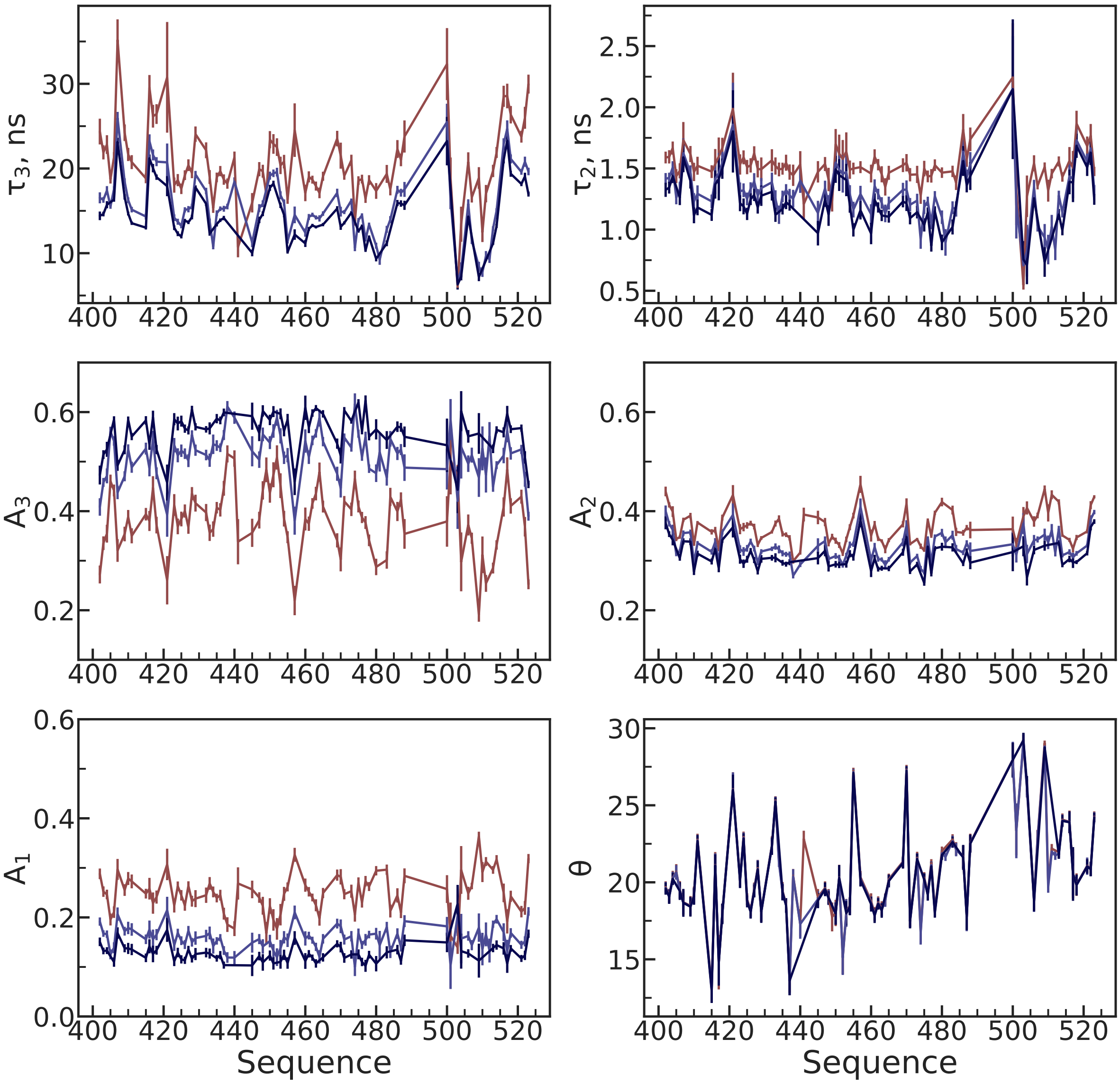

Figure 6.9: Comparison of the dynamic modes estimated using 5 parameter Model-Free analysis of relaxation rates in concentrated phase optimizing two amplitudes, two timescales and CSA-DD angle, assuming constant value of fast timescales of motions equal to 50, 75 or 100 ps (red, light-blue and blue respectively).

If the single ns timescales of motions determined from the independent Model-Free analysis in the concentrated phase are better separated from the slower intermediate distribution, the estimated apparent dynamic modes might represent reorientation of polypeptide units associated with changes in the number of internal degrees of freedom available to the peptide plane, which are expected to decelerate with viscosity with a smaller rate. The respective distribution of correlation times in the dilute phase should oscillate within the range of hundreds of ps. The estimated from ^{15}N-backbone relaxation data analysis in the concentrated phase CSA-DD angle differs from values estimated from independent Model-Free analysis in dilute phase by few degrees with magnitude of difference (of up to 10° increasing towards the protein termini - θ equal to 25° and 15-25° for dilute and concentrated phases respectively. Even though on average chemical shifts observed in the concentrated phase are very similar to the ones observed in the dilute phase, ^{15}N chemical shift tensor might posses a varying degree of anisotropy, which magnitude and relative orientation with respect to the NH-backbone amide vector depends on local electronic environment of ^{15}N nucleus and to an extent is influenced by polypeptide fragment conformation and hydrogen bonding between neighboring molecules and fragments of macromolecules of varying polarity. The magnitude of difference in CSA-DD angle between unfolded and folded helical elements with values of 25-26° and 21-22°has been observed previously (235, 338). The range of perturbations to CSA due to static hydrogen bonding based on *ab initio* calculations of crystalline and isolated N-acetyl-Valine-Leucine is expected to yield $\delta\Delta\sigma$ of 5 ppm with minor modifications to CSA-DD angle within range of 1-4°(339). The remaining part of CSA-DD angle modifications might be resultant from changes in local geometry of peptide unit(s) associated with population of strained conformations for extended time intervals leading to modifications in relative orientation or magnitude of the CSA tensor (194, 340–342). Of special interest are segments of polypeptides rich in charged or bulky residues - charge attraction and repulsion, as well as stacking of aromatic rings between neighboring side-chains may lead to more compact conformations, in turn modifying the CSA tensor of subsequent residues leading to oscillatory patterns observed for CSA-DD angle sequence dependence, i.e. 415-418 (AQVS), 436-441 (EDRRVK), 450-454 (SYRET), 471-475 (TPLDI), 485-489 (PQDSR), 504-508 (SEEQ) with magnitude of residue to residue perturbation in the range of 3-5°. The extent of the CSA and CSA-DD angle modifications may originate from more frequent inter or intramolecular interactions of protein macromolecules in the concentrated phase associated with higher concentration of the protein or local compaction favoring anisotropic perturbations to the electronic density surrounding the backbone ^{15}N nuclei.

To evaluate the quality of prediction of relaxation rates in the concentrated phase based on the combined EA-Model-Free analysis (in the dilute phase) and average value of relative deceleration of water molecules (ρ of 3.44), assuming that the amplitudes of motions do not change in the concentrated phase with respect to the dilute one, or in other words, that the viscosity scaling of the apparent dynamic modes representing population weighted average of correlation times takes into an account changes to their populations associated with crowding, relaxation rates in the concentrated phase have been calculated and compared with values observed experimentally. Three different scenarios have been considered: with viscosity-independent fast component (A), with fast component scaled by a factor of 1.6 (ϵ_{fast} equal to 0.18, B) and one with modification to amplitudes of motions obtained from Model-Free analysis in the dilute phase with increased and decreased value of the amplitude of slow and intermediate motions by 0.05 and scaling of the fast apparent dynamic mode (A_{slow}, $A_{\text{inter}} \pm 0.05$, ϵ_{fast} equal to 0.18, C). In the first two cases, the estimated R_1 and R_2 values yield close comparison, while the predicted heteronuclear nOes significantly diverge from experimental ones in the first case (without scaling of the fast motion timescales with viscosity). Addition of friction-dependent scaling of the fast component allows to minimize the disagreement between predicted and experimental heteronuclear nOe (figure 6.10), suggesting that the fast modes are subject to minor viscosity-dependent deceleration, phenomenon otherwise unobserved in the relaxation rate analysis of dilute to semi-dilute preparations of SeV NTail or MKK4 due to relatively small differences in viscosity between the admixtures with ρ not exceeding 0.4.

Further discrepancies between the predicted and experimental data might stem from transient intra or intermolecular interactions between neighboring side-chains or with other (macro)molecules in the concentrated phase modifying distribution of reorientation rates representing slow motions. The modification of the amplitudes of motions on the order of 0.05 with increased values of the amplitudes of slow motions and decreased values of the apparent intermediate dynamic mode of correlation times distribution yields almost exact agreement between experimental and predicted transverse and longitudinal relaxation rates with small but systematic differences in heteronuclear nOe remaining in range not exceeding 0.1 (figure 6.10C). The estimated from the Model-Free analysis timescales of motions ranging from tens of ps to hundreds of ps average to value of 50 ps representative of distribution of fast motions and are subject to viscosity-dependent deceleration.

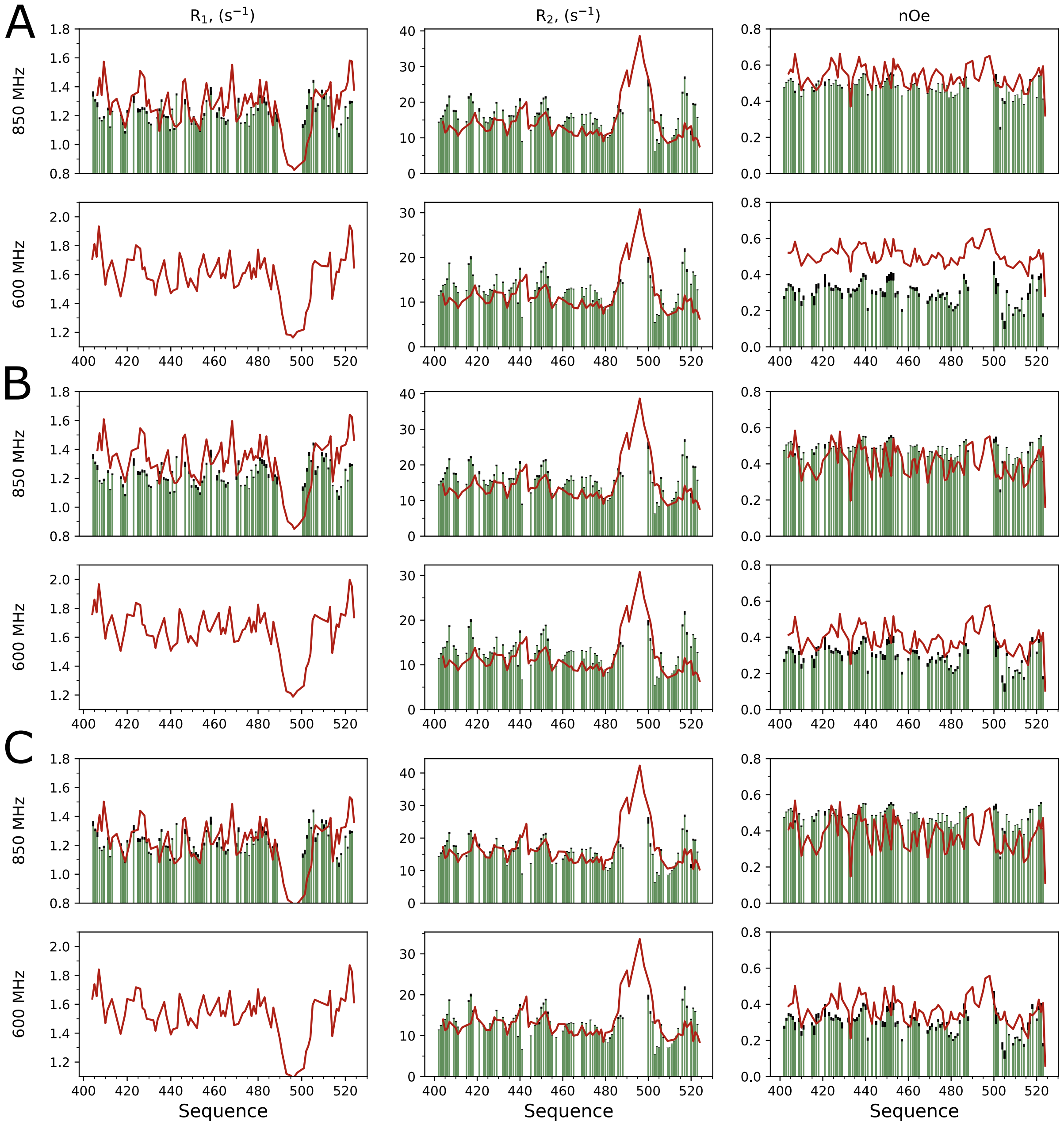

Figure 6.10: Comparison of experimental relaxation rates observed at 298K with Einstein-Arrhenius Model-Free prediction with ρ value equal to 3.44 for three different sets of analyses: one without viscosity-dependent scaling of fast dynamic mode (A, ϵ_{fast} equal to 0), one with scaling of fast dynamic mode (B, ϵ_{fast} equal to 0.18) and with increased and decreased value of the amplitude of slow and intermediate motions by 0.05 respectively and scaling of fast dynamic mode (C, A_{slow} + 0.05, A_{inter} - 0.05, ϵ_{fast} equal to 0.18).

Without scaling of average value of fast correlation times by a factor 1.6 (τ_{fast} modification from 50 to 75 ps) leads to under and overestimation of heteronuclear nOe values at 850 and 600 MHz respectively, suggesting that τ_{fast} of 50 ps might over and underrepresent the distribution of fast motions by tens and hundreds of ps reorientation rates respectively, in turn implying narrower, or rather more populated distribution of correlation times representing fast intermediate motions occurring on hundreds of ps. The model predicts the overall changes to transverse and longitudinal relaxation rates probing motions occurring on single and tens of ns timescales with exception of several residues for which differences within range of 0.1 and 2-5 s^{-1} for R_1 and R_2 respectively are observed. The discrepancies may originate from non-specific binding events between macromolecules in the concentrated phase, or from accuracy of determination of friction coefficients and approximation of underlying distributions of correlation times from Model-Free analysis in the dilute phase. Further inaccuracies in prediction of unfolded protein dynamics in complex macromolecular environment are associated with estimation of relative deceleration of water molecules from the longitudinal relaxation rates of water protons. For instance, an inaccuracy in determination of $R_{1,\text{water}}$ on the level of 5-10% in more complex environment, assuming that $R_{1,\text{water}}$ in the dilute phase has been determined with 1% accuracy leads to overestimation of τ_{slow} on the order of 5-10% for ρ of 3.0 leading to relative difference in R_2 on the order of 1 to 2 s^{-1}, which is half smaller to the one observed between predicted and experimental data.

6.5 Discussion

The successful application of the combined EA expression coupling timescales of motion with model 3 for three different proteins: C-terminal parts of Sendai and Measles Virus Nucleoprotein, as well as Map kinase kinase 4 allows to draw some general observation with regard to the generality of the phenomena associated with crowding at the nanoscale level. In all of the studied cases the slow and intermediate friction coefficients oscillate around the values of 2.0-3.5 and 0.8 to 1.1 respectively. The differences in the estimated slow mode friction coefficients between the studied cases may originate from differences in molecular architecture of the unfolded chain - side-chain size and its polarity, in turn influencing its intramolecular interactions between neighboring residues, polypeptide chain time-dependent topology or specificity of the dynamical interplay between protein residues and water molecules. The frictional effects acting on the reorienting polypeptide units associated with interactions between neighboring peptide units are coupled to the solvent deceleration by hydrogen bonding to the peptide chain fragments. The incorporation of electrostatic and van der Waals interactions between side-chains of neighboring peptide units, which remain hydrogen bonded with surrounding water molecules may lead to greater frictional effects acting on the reorienting polypeptide units (343, 344). The observation of separate distributions of timescales of motions occurring on single ns in the concentrated phase, which in dilute phase should span range of hundreds of ps, suggest that the latter inform on the internal degrees of freedom of peptide planes, while the motions with mode values equal to several ns (single ns in dilute phase) inform on external degrees of freedom, which deceleration is associated with hydrogen-bonding dependent deceleration of the peptide plane reorientation.

The observed field dependence of $\mathrm{R}_{1,\mathrm{water}}$ in the concentrated phase originates presumably from exchange of water molecules between water molecules in the first hydration monolayer and the ones found in the cavities transiently formed by fragments of the unfolded chain. The exchange rate between the two states would be limited by lifetime of the cavity depending on reorientation rate of polypeptide units (and their side-chains), in other words, the dwell time of water molecules within the cavity should not exceed single to several ns. In that case, one should be able to observe the exchange contribution at lower magnetic fields, which would be dependent on correlation times and populations of transiently immobilized water molecules (345). In that case, ρ should inform on rotational perturbation factor and potential contributions from exchange associated with modulation to degree of confinement of water molecules on ns timescales. Presumably, water molecules in the cavities would have greater probability of formation and longer lifetime of hydrogen bonds with amide protons and other polar atoms of polypeptide side-chains. The exchange rate between water molecule in the cavities and the ones in first hydration monolayer should depend on lifetime of the cavity, which is proportional to the timescale of conformational rearrangement of unfolded chain fragment in the concentrated phase, and reorientation of water molecules within and in vicinity of the cavity, both of which depend on concentration of macromolecules in the medium effectively decelerating water molecules and contributing to modulation of conformational fluctuations of polypeptide fragments (141, 143, 276, 290).

The intramolecular interactions between neighboring peptide units may lead to greater degree of correlation between reorienting polypeptide units and are commonly referred to as segmental motions. Within the framework of length-scale dependent viscosity one should expect greater values of friction coefficients for segments of conformationally restricted bulky residues as compared to more flexible glycine or serine-rich fragments (136, 173, 174). It is difficult to expect that the sequence dependent profile of the apparent slower mode friction coefficients of MKK4 and MeV NTail would originate solely from the phenomenon. The origin of increased values of ϵ_{slow} for residues in the protein termini might stem from decreased capability for reorientation of polypeptide units as a result of inter or intramolecular interactions between neighboring side-chains leading to crowding-induced compaction (178, 243, 301). Alternatively, the formation of cavities trapping an ensemble of water molecules in between the relatively static from the perspective of reorientation rate of water molecules unfolded chain may further enhance coupled deceleration of water molecules and reorienting peptide planes (142, 257). The extent of the latter is expected to be rather small due to difference in molecular weight of unfolded polypeptide fragments, its moments of inertia and that of water molecules and clusters thereof. Fragments of unfolded proteins featuring more bulky side-chains are expected to be less prone to form the cavities to a much lesser extent due to smaller number of degrees of freedom available to the reorienting polypeptide units (258). The frictional effects associated with interactions between neighboring side-chains for the residues complement formation of dynamic loops by unfolded chain(s) capable of constraining water molecules and their clusters (289, 344).

Given the amount of assumptions made on characterization of motional process governing ^{15}N relaxation, from semi-classical Bloch-Wangness-Redfield theory, Model-Free analysis to distinct averaging processes yielding apparent motional modes, and simplicity of the model (equation 5.4), as well as the complexity of molecular environment of live cells or concentrated macromolecular environment of the collapsed liquid-liquid phase separated phase, the agreement between predicted and experimental data is reasonable. Being able to predict the unfolded chain dynamics prior to relaxation rate measurement offers advantages in terms of interpretation of experiments in more complex environments, such as reconstituted complexes, phase-separated proteins or in-cell. The averaging of dynamic modes and relative sensitivity of experimental relaxation rates to correlation times representing motions occurring on tens of ps to tens of ns timescales obscures the information on the unfolded protein dynamics. By relating the timescales of motions to the relative deceleration of water molecules one may recapitulate information on reorientational motions of single peptide planes on internal and external degrees of freedom on different length-scales in more complex macromolecular settings. Considering the comparison of relaxation rates measured *in vitro* and *in vivo* without previous knowledge of how viscosity affects the backbone conformational dynamics on different timescales, one could have an impression that the observed difference in ^{15}N relaxation rates in central and C-terminal part of the protein originates mainly from specific binding to other proteins in-cell. For example, a very recently characterized binding of MKK4 to p38α occurs in a similar region of the protein leading to modification of its dynamics *in vitro* yielding analogous sequence-dependent profile of relaxation rates dominated by slow dynamic modes (70). The predicted relaxation rates in the complex environment provide additional and accurate information on ground state dynamics of the protein in the complex.

The application of EA formalism allows for successful prediction of protein dynamics in concentrated phase of protein, where slow timescales of motion are predicted to be an order magnitude (approximately 9 to 10 times) slower in comparison to the ones in dilute phase. The applied analysis and comparison of predicted and experimental relaxation rates suggests that the timescales of fast motions are subject to minor changes (to within tens of ps) in the extremely concentrated interior of droplets. The effect of viscosity on the fast dynamic modes could not have been observed in the Einstein-Arrhenius Model-Free analysis of ^{15}N-backbone relaxation rates of SeV NTail or MKK4 due to rather small difference in viscosity between dilute phase and crowded environment *in vitro* (ρ values of up to 0.3). At considerably high protein volume fractions, the majority of the solvent is associated with the protein, i.e. the fraction of water molecules in hydration layer of the protein in comparison to bulk water is considerably high. The average slow-down of NH-backbone reorientation rates represented by deceleration of the solvent to a small extent is expected to be influenced by reorientation rate of water molecules within hydration layer. The largest retardation factors representing relative slow-down of rotational correlation times of water molecules, which ρ is significantly influenced by, are observed in the first hydration layer yielding values of three to five for folded proteins in comparison to bulk water (143, 288). Taking into account deceleration of water molecules and the exchange between water molecules remaining within the cavities formed by polypeptide segments, the observed value of ρ in the concentrated phases of approximately 3.4 seems to be in the correct range. Within the droplet interior, apart from solvent mediated slow-down of the polypeptide unit reorientational dynamics influenced by its degrees of freedom at different length-scales, one could expect additional contributors, i.e. frequent specific or non-specific interactions between and within macromolecular chains leading to additional retardation of motions. In that case, one can anticipate potential modification to populations of correlation times with respect to the experimental sensitivity windows reflected in the estimated amplitudes and timescales of slow motions leading to significant deviations between predicted and experimental R_2 values. Given small differences between predicted and experimental relaxation rates in the analyzed example of MeV NTail, the slow-down of NH-backbone reorientational dynamics due to inter and intramolecular interactions both between and within polypeptide fragments, as well as fragments of macromolecules of the crowder, expressed as a function of relative deceleration of water molecules seems to be the main mechanism governing modifications to backbone dynamics of unfolded proteins in those highly concentrated phases.

6.6 Conclusions

The discussed here approach allows to uniformly describe the dynamic modes of an unfolded or transiently folded system across a range of temperatures and crowding conditions for three different proteins. The heterogeneity of estimated friction coefficients remarks on the complexity of motions underlying distributions of correlation times representing reorientation of polypeptide units limited by internal and external degrees of freedom occurring on four decades of timescales. The relatively greater friction coefficients for protein residues in termini suggest importance of inter and intramolecular interactions with other macromolecules in the medium associated with limited volume for the peptide plane reorientation. In limiting cases, values of friction coefficients range from 0.2 for dynamic modes representing fastest of the motions presumably associated with stretching and deformation vibrations of NH-backbone vector to 4.5 for slowest of the motions involving reorientation of peptide units limited by external degrees of freedom at longer length-scales. The application of the model additionally allows to predict changes in dynamics of the proteins in crowded macromolecular environment as a function of relative deceleration of water molecules with complexity ranging from *in vitro*, through concentrated solutions originating from collapse of phase separated protein droplets, ending on complex cellular environment. The tight coupling between reorientation of the solvent molecules and the polypeptide units seems to be of general nature and applicable to many systems of interest. The applied approach allows to accurately predict deceleration of NH-backbone fluctuations associated with crowding for many complex systems and foster further investigation of intra and intermolecular interactions such as reptation or association between macromolecular species in crowded environments.

CHAPTER 7

Summary and future outlook

7.1 Unified description of unfolded protein dynamics under native conditions

Model-Free analysis allows to determine dynamic modes associated with loss of memory function due to stochastic processes associated with reorientation of polypeptide units from the experimentally derived relaxation rates informing representing rate of magnetization decay originating from nuclear spin interactions occurring on ps-ns timescales. Numerical simulations suggest that ratio between the modes of distributions of timescales of motions giving rise to the observed relaxation rates should be at least five, and their distributions should not overlap significantly to accurately and precisely determine timescales and amplitudes of motions. The presented in Chapter 3 set of numerical simulations allows to predict which combinations of amplitudes and timescales (i.e. SDF components) originate from overlapped distributions of timescales, in turn allowing to indicate regions of protein for which determination of dynamic modes from Model-Free analysis might be error prone, e.g. the ultra-flexible polypeptide segments, such as glycine or serine-rich motif of MKK4, or protein residues located at termini of unfolded chain of NTail. When the distributions of timescales begin to overlap, the estimated timescales of slow motions correspond to shoulder or tail of the joint distribution with resultant timescales of motions of proximal modes of at least five, and the corresponding amplitude that is typically smaller than the one expected from narrow separate distributions. Comparison of dynamic modes derived from average correlation functions derived from simulations suggest possibility of estimation of more than three apparent dynamic modes describing reorientation process of polypeptide units, as long as experimental precision and accuracy, as well as degree of overlap of the distributions with respect to sensitivity of experiment allow for it. The reduction in the number of dynamic modes from four to six observed in MD simulations to three typically estimated in Model-Free analysis of experimental relaxation rates leads to apparent timescales and amplitudes of motions representing population-weighted average correlation times. Part of the information on slow and slow intermediate motions may be represented by slow amplitudes of motions, in turn modifying the apparent modes of distributions (timescales) and associated with them amplitudes representing motions in a manner that is defined by sensitivity of the experimentally accessible relaxation rates, which values are proportional to the linear combinations of SDFs at frequencies associated with allowed spin transitions for ^{15}N-^{1}H backbone amide spin pair.

The application of the temperature-dependent Model-Free analysis within the limits of adiabatic approximation to another unfolded system - regulatory domain of MKK4 featuring distinct dynamic profile to the one evaluated previously (NTail) allows to confirm a minimum of three dynamic modes to describe the experimental relaxation data (Chapter 4). The timescales and amplitudes of motions determined from temperature dependent analysis of NTail and MKK4 allow to draw some general observations with regard to rate of temperature dependent acceleration of the apparent dynamic modes. The motions occurring on tens and hundreds of ps dominate the dynamic modes of residues located in the termini of both proteins. The rate of change of amplitudes of fast motions with temperature suggests more pronounced distribution of correlation times representing motions positioned at hundreds of ps for MKK4 featuring greater overlap with $L(\omega_{\mathrm{H}})$. The smaller apparent rate of change of amplitudes of fast motions estimated for residues situated in the termini of NTail suggests significant contribution of fast tens of ps motions, and relatively well defined distribution on hundreds of ps motions with mode positioned further away from the $L(\omega_{\mathrm{H}})$ maximum at 14.1T (295 ps) - smaller number of correlations times in the sensitivity window, to the one observed for the residues in the N-terminus of MKK4. The residues in central part of both proteins feature similar rate of acceleration with temperature represented by a_{fast} leading to assumption of two distributions positioned at tens of ps and hundreds of ps with similar degree of overlap with $L(\omega_{\mathrm{H}})$ (150 - 295 ps), i.e. at 20-30 ps and 300-400 ps, which temperature dependent acceleration will lead to overall decrease and increase in the amplitudes of fast and slow motions respectively. Since amplitudes of slow motions depend on number of correlation times in the window of sensitivity defined by $L(\omega_{\mathrm{N}})$, the apparent rate of slow amplitude decrease with temperature is associated with increase in amplitudes of intermediate motions, which is compensated by loss of correlation times associated with acceleration of fast intermediate motions (occurring on hundreds of ps) towards $L(\omega_{\mathrm{H}})$ maximum. For timescales of motions in the range, the relative sensitivities of the two SDFs in magnetic field strength ranging from 14.1 to 22.3T compensate each other. The rate of this compensation expressed by a_{diff} points to greater rate of acceleration of slower intermediate motions towards the maximum of $L(\omega_{\mathrm{H}})$, decreasing degree of overlap of the distribution with $L(\omega_{\mathrm{N}})$. The greater values of a_{diff} for NTail point to positioning of slower intermediate distribution of timescales of motions further away from $L(\omega_{\mathrm{H}})$ maximum or greater rate of acceleration of the latter ones in comparison to the fast intermediate motions yielding greater values of $L(\omega_{\mathrm{H}})$. The observations are consistent with estimated values of intermediate dynamic modes at values of 800 - 1200 ps with greater values in center of the chain and apparent activation energy approximately in range of 8 to 12 kJ/mole. Values of a_{diff} estimated for residues in C-terminal part of MKK4 (for the last few residues) are comparable to values observed for residues in termini of NTail suggesting that the respective rates of change of slower intermediate timescales with respect to $L(\omega_{\mathrm{N}})$ maximum are comparable for both proteins, in both cases determining values of intermediate dynamic modes for timescales ranging from 1100 to 1200 ps, in a limit where values of τ_{slow} do not force their apparent acceleration.

The viscosity-associated deceleration of dynamic modes determined from the temperature dependent analysis leads to a decrease in number of correlation times within range defined by $L(\omega_{\mathrm{H}})$ and $L(\omega_{\mathrm{N}})$ in a manner that depends on the coefficient relating deceleration of water molecules with rate of slow-down of polypeptide fragments reorientation rates. The relative deceleration of intermediate and slow intermediate modes representing relative decrease in number of correlation times in the sensitivity ranges defined by $L(\omega)$ distributions, which is partially compensated by increase in number of correlation times in sensitivity window defined by $L(\omega_{\mathrm{H}})$ associated with slow-down of fast and fast intermediate motions. In the classical Model-Free analysis, deceleration of correlation times is associated with modifications to population-weighted average representation of correlation times with regard to the sensitivity of the experiment and concomitant changes in values of the amplitudes of motions represented by a relative increase in A_{slow} and decrease in A_{inter} or A_{fast} respectively, that would depend on the relative deceleration of motions. The presented in Chapter 5 approach allows to describe the viscogen-associated deceleration of stochastically modulated motions giving rise to the allowed spin transitions observed in form of ^{15}N-backbone relaxation rates. Modification of medium composition by incorporation of hundreds of g/L of macromolecules leads to increase in viscosity, which in turn slows-down the conformational fluctuations of polypeptide fragments in a manner that is dependent on their relative reorientation rates. The Model-Free analysis of relaxation rates measured systematically as a function of viscogen concentration reveals the remarkably tight coupling of backbone motions to solvent dynamics with the ones occurring on single ns timescales mirroring relative deceleration of water molecules reorientation rates (estimated by the relative increment in longitudinal relaxation rates of water protons), while relative deceleration of slower correlation times is on average three times faster. The origin of the behavior is presumably due to greater length-scale of the motion involving dependence of polypeptide units reorientation rates on greater number of external degrees of freedom. The observed viscosity-dependent scaling (friction) coefficients (ϵ) are not visibly perturbed by temperature in the studied temperature range (10K) suggesting orthogonality of the approach to the temperature-dependent Model-Free analysis. The combined viscosity-temperature Model-Free analysis yields dynamic modes indistinguishable from the ones reported previously estimated from separate temperature or viscosity-dependent Model-Free analysis. The robustness of the model evaluated by thorough cross-validation and the tight coupling between motions of the solvent and the ones exhibited by the polypeptide units yields promising applications in prediction of temperature or viscosity induced modifications to dynamic modes of unfolded proteins (346).

The method has been applied to the previously studied from point of view of temperature dependence of the apparent dynamic modes representing ps-ns conformational fluctuations of C-terminal fragment of Measles Virus Nucleoprotein (NTail), as well as, Map kinase kinase 4 (MKK4), reaffirming the tight coupling of backbone motions to the solvent dynamics with ϵ_{inter} of 0.8 - 1.0, and ϵ_{slow} of 2.0 - 4.0. The origin of differences between friction coefficients for different proteins is presumably connected to extent of limitations imposed on reorientation of peptide plane(s) associated with interactions between neighboring side-chains resultant from imbalanced polarity of macromolecules' fragments, and hydrogen bonding between water molecules, as well as, between amide hydrogen atoms and polar side-chain groups involved in sustenance of hydrogen bonding network in the first hydration layer of an unfolded chain. As spatial extent of the motion influences degree to which hydration network is perturbed, their overall expected slow-down is expected to yield greater values for reorientation processes which are limited by greater number of external degrees of freedom associated with polar interactions between side-chain groups. The timescales of motions influenced by the limitations in reorientation of peptide plane(s) lead to change in number of correlation times in observation windows of respective SDFs (J(ω) representing probability of observing spin transitions at Larmor frequencies of ^{1}H and ^{15}N or combination thereof), arising due to stochastically modulated motions and are represented by friction coefficients. From perspective of water molecules with reorientation rates ranging from single ps to hundreds of ps, the unfolded chain is almost static, featuring loops, kinks and cavities formed by the polypeptide fragments and its side-chains leading to a spectrum of timescales of water molecule reorientation rates associated with their encapsulation in the cavities of different sizes. The hydrogen bonding between polar side-chains and amide hydrogen atoms should influence reorientation capability of water molecules both in terms of its rotation and translation in a a manner that depends on degree of flexibility and polarity of polypeptide fragments, both directly dependent on molecular architecture of unfolded protein side-chains.

The fragments of unfolded proteins of greater flexibility constituting glycines and serines are expected to feature faster reorientation rates associated with smaller limitations on the external degrees of freedom and shorter lifetime of the cavities encapsulating water molecules formed by the reorienting chain on timescales longer than tens of ns. The regions of varying degree of flexibility featuring polar and apolar voluminous side-chains are expected to form the water encapsulating cavities for extended periods of time leading to slower reorientation rates of polypeptide units. The incorporation of viscogen to the medium leads to slow-down of reorientation rate of water molecules associated with their hydrogen bonding to the polar groups of viscogen and encapsulation within the cavities formed by the macromolecular chains of the crowder. At the same time, the limitations imposed on the unfolded chain associated with repulsion between the fragments of unfolded peptide chain and the macromolecule increase in a manner that depends on the available degrees of freedom of the peptide units and chain fragments, leading to a decrease in the radii of gyration of the (bio)macromolecules and extended lifetime of water remaining within the cavities of varying size. The overall slow-down of bulk portion of water molecules associated with hydrogen bonding resultant from crowding associated deceleration of peptide units potentially leading to modulation of cavity lifetimes formed by neighboring peptide units further enhancing deceleration of water molecules and limiting their concerted reorientation. The balance between steric interactions of polypeptide fragments and viscogen macromolecules leads to a modulation of the deceleration rate of peptide unit reorientation that depends on the lifetime and volume of the cavities. More flexible parts of the chain are expected to yield short-living cavities in dilute phases with increase of its formation probability depending on degree of attractive interactions between neighboring peptide units. In that case, one could expect greater friction coefficients for protein fragments capable either of greater intramolecular interactions between neighboring side-chains, or, the ones capable of undergoing greater compaction, featuring relatively low and high number of degrees of freedom. The experimentally observed relative deceleration of reorientation rates of polypeptide units with varying number of both internal and external degrees of freedom yielding greater values in the Molecular Recognition Element of C-terminal fragment of Sendai Virus Nucleoprotein (MoRE NTail) involved in the fast exchange between folded and unfolded conformations, and flexible part of the N-terminal fragment of Map kinase kinase 4 (MKK4) corroborate the proposed description. The comparison of τ_{slow}, τ_{inter} and $\mathrm{A}_{\mathrm{slow}}$ values estimated for polypeptide units of MeV NTail and SeV NTail provides additional information on degree of overlap of distributions of correlation times in the dilute phase representing stochastically modulated motions opposing the limitations imposed by external and internal degrees of freedom on varying length-scales, suggesting more overlapped (compact) distributions of correlation times in the range of $\mathrm{L}(\omega_{\mathrm{N}})$ for the former one. The latter is presumably associated with probability of secondary structure formation, or in other words, potential of intramolecular interactions between neighboring side-chains. The latter implies lower values (greater probability) of hydrogen bonding, electrostatic or van der Waals interaction potential between neighboring side-chains for SeV construct.

The possibility of prediction of ^{15}N-backbone relaxation rates of unfolded protein in a medium containing viscogen of different size and physicochemical properties solely on the basis of relative slow-down of average reorientation rate of water molecules and dynamic modes of polypeptide fragments in the infinitely dilute medium corroborates the observation of high degree of dependence of solvent dynamics on conformational fluctuations of unfolded protein or macromolecules of crowder suggesting general applicability of the model to more complex systems. Indeed, the model is capable of predicting the relative slow-down of polypeptide fragments reorientation rates on ns timescales in cellulo suggesting that the relative slow-down of the NH-backbone dynamics in highly complex cellular environment may be linked to the average deceleration rate of the solvent. Although, the observation might be system dependent and additional contributions associated with specific binding between key proteins in the cellular environment should be considered, the model might shed light on ground state dynamic profile of any protein of interest in complex environment, allowing for prediction of local unfolded protein dynamics, associated purely with non-specific effects arising due to macromolecular crowding, before designing more complex experiment in the native environment. Similarly, the model is capable of predicting the NH-backbone dynamics of unfolded proteins and folded proteins featuring fragments involved in fast exchange between folded and unfolded intermediates in highly concentrated phases, on the basis of backbone dynamics in the dilute phase and relative deceleration of water molecules in the former with respect to the latter, where relative deceleration of the slowest dynamic mode reaches astounding order of magnitude. The crowding-associated slow-down of average timescales of motions may lead to inaccurate estimation of the apparent dynamic modes from transiently folded parts of proteins in more complex environments due to line broadening effects associated with intrinsic dynamics, magnetic field inhomogeneity and local differences in magnetic field susceptibility due to high concentration of macromolecules complemented by presence of paramagnetic ions in the native environment, and sensitivity of spin relaxation experiments to timescales of motions exhibited by reorienting peptide units in the concentrated phase. The ability of prediction of ground state backbone dynamics on multiple timescales facilitates evaluation of conformational fluctuations of polypeptide fragments in highly complex environments found in cell or in concentrated phases. The viscosity-associated deceleration of the apparent dynamic modes allows to separate contributions arising from limitations to reorientation of peptide planes from external and internal degrees of freedom observed in the form of correlation times spanning range of hundreds of ps and single ns in the dilute phase with their apparent deceleration to single and tens of ns observed experimentally in the concentrated phases. The inability to accurately determine the contributions from internal and external degrees of freedom limiting stochastically modulated motions is associated with sensitivity of respective SDFs to the distributions of timescales of motions at frequencies associated with spin transitions describing interactions between nuclear magnetic moments allowed for the ^{15}N-^{1}H backbone spin pair.

7.2 Beyond single ns timescale

Observation of backbone protein dynamics is limited by lifetime of a correlation function of spatial part of spin-lattice interaction tensor describing reorientation process of peptide units occurring concurrently to the allowed spin transitions contributing to the experimental relaxation rates with probability described by the SDFs at frequencies associated with the spin transitions. Therefore, the apparent dynamic modes representing average correlation times arising due to collective motions of polypeptide fragments and long-range dynamics becomes almost impossible with classical relaxation rate measurement. An interesting approach in pursuit of observation of conformational fluctuations of unfolded proteins on longer timescales relies on application of Silica nanoparticles (SNP). As long as the interactions between folded protein and SNPs remain in the fast exchange regime and reorientation rate of SNP is much longer than rotational correlation time of the protein, while the interaction between protein and SNPs does not modify internal dynamics of a polypeptide, difference between ^{15}N-R$_2$ of backbone amide in solution containing SNP and dilute one is directly proportional to the population-weighted average S^2 of the NH-backbone in the free form (347). The deceleration of reorientation rate of the protein in the bound state allows for experimental observation of internal motions slower than those observed in the classical Model-Free analysis with τ_{int} remaining in the limit of 0.1 τ_{NP}. Magnitude of the ΔR_2 and effective time window of τ_{int} observation may be varied by adjusting concentration of NPs and their size respectively. The exact timescales of motions may be derived from RED analysis of MD trajectories with varying integration time window. The method has been successfully applied to colicin E7 immunity protein (Im7) and calcium-binding domain 1 of canine sodium-calcium exchanger NCX1 (CBD1), in both cases revealing fraying of the alpha-helix (Im7) and extended loop dynamics reaching single microseconds. Application of the method to Ubiquitin does not reveal dynamics on timescales longer than single ns, which is presumably associated with modification to the internal protein dynamics in the bound state with SNP (348). The interpretation of ΔR_2 SNP profile for unfolded proteins is slightly more difficult, as modulation of the value, due to their broad conformational sampling and multiple binding modes is sequence dependent. Measurement of relative interaction strength between SNPs and isolated amino acids allows to asses and predict the ΔR_2 profile of SNP-bound unfolded protein from an empirical, free-residue interaction model (FRIM). The range of the unfolded protein dynamics modulation defined as rate of decay of the perturbation with polypeptide sequence from the interaction centers oscillates between four (apolar) and thirteen (polar or charged) residues. The model has been successfully applied to four unfolded proteins (p53TAD, α-synuclein, prokaryotic ubiquitin-like protein and cytosolic loop (Cyto2) of canine sodium-calcium exchanger NCX) and their mutants (349). As the FRIM relies on summation of the binding profiles from interactions between SNPs and single amino acids, it does not take into account cooperativity effects of neighboring residues upon binding and does not allow for prediction of absolute binding affinities (350). An alternative to the FRIM, based on statistical-mechanical description of polypeptide binding to SNP surface, termed SILC (simultaneous linker-ligand cooperative binding) allows to estimate energies of interactions by taking into an account cumulative binding modes associated with neighboring effects (side-chain interactions and backbone tethering) leading to a continuum of transitions describing bound and unbound equilibrium, allowing for better prediction of mutagenesis effects on the observed ΔR_2 profiles of SNP-bound unfolded proteins and thermodynamics of the interaction. The model and its potential extensions may serve as a basis for understanding of dynamically-modulated interactions of unfolded proteins with its partners (351).

Characterization of motions occurring on timescales longer than several ns typically observed in Model-Free analysis of classical relaxation experiments at high magnetic field strengths (above 10T) may be provided from analysis of longitudinal relaxation experiments recorded at magnetic field strengths of lower magnitude. At lower magnetic fields, the values of SDF at Larmor frequencies of ^{15}N and ^{13}C approach SDF at zero frequency, which in turn allows to more accurately estimate dynamic modes of distributions of correlation times representing slowest motions, in turn influencing accuracy of estimation of associated with them amplitude represented by S^2 for the slowest motions (214). At lower magnetic fields, however, the spectral resolution becomes an issue with overlapping spectral peaks, which in turns obscures information content that may be determined from the experiment. A solution to the issue is presented in the form of high-resolution relaxometry, which following application of a shuttle enabling the sample to oscillate within stray field of different magnitude with strength down to 0.5T. In the stray field, however, it is impossible to apply any radio-frequency pulses rendering high influence of cross-relaxation pathways on the evolution of magnetization visible in the form of non-exponential magnetization decay (352). To alleviate the issue, one may apply numerical iterative procedures for estimation of relaxation rates by simulation of the experimental pulse sequence outside of the homogeneous field (353). Alternative solution has been provided with implementation of an additional probe at spectrometer top operating at lower magnetic field. In a so-called two-field NMR approach, the sample oscillates between higher and lower field allowing for excitation of spins and detection at the former, while allowing to preserve control over coherence pathways during relaxation delay at the latter (354). The approach has been successfully applied to ^{13}C (side-chain) relaxation rates of Ubiquitin allowing to better determine the distribution of correlation times representing side-chain motions using two-field NMR and high resolution relaxometry in comparison to the analysis of high-field relaxation data only (151, 355). Application of the methodology to unfolded systems featuring broad distribution of correlation times on timescales longer than tens of ns will be likely a subject of further studies.

7.3 Dynamics in biologically relevant systems

Flexible proteins such as prion-like proteins, containing low-complexity domains composed of aromatic and charged residues tend to phase separate as a result of intra and intermolecular interactions between them. A body of evidence originating from experiments and simulations of unfolded proteins suggests that charge and aromatic residue patterning directly influences capability of the system to phase separate. Current models predicting phase-separation tendency of unfolded proteins take into an account molecular architecture of side-chains dissecting fragments of proteins into so-called stickers and spacers, with the former being responsible for interaction within, and between fragments of macromolecular chains, while the latter allowing for mobility of the chain in search of binding motifs (183, 184, 331, 356, 357). Stickers are typically composed of low complexity repeats equipped with charged or aromatic residues i.e. arginine or tyrosine, surrounded by residues with less bulky side-chains such as glycine or serine. Number of aromatic residues in variants of FUS proteins influence tendency of unfolded chain to form intramolecular contacts observed by differences in relative size distribution and intramolecular NOEs, in turn, affecting propensity of macromolecules to phase separate - the more aromatic residues, the lower protein concentration is required to observe phase separation. Consistently with the observation, mutations of tyrosines and arginines to serines and glycines respectively deteriorates capacity of proteins to phase separate. Similar observations have been made for Ddx4 lysine variants, where 24 arginines have been substituted with lysines or phenylalanines with alanines suppressing their ability to form condensates further substantiating the importance of cation-π interactions in phase separation of proteins (329). The substitutions of glycines or serines to alanines or glutamines does not significantly modify the onset of phase separation, yet may influence fluidity of the droplet as observed by measurements of translational diffusion coefficients using Fluorescence recovery after photobleaching (314). The coarse grained simulation of FUS proteins using stickers and spacers model, which takes into account aromatic (stickers) and non-aromatic residues represented by single-amino acid beads successfully predicts the experimentally observed phase separation diagrams (187, 313, 358). Patterning of charged residues also seems to play a role in determining LLPS preferences. For instance, for Ddx4 and NICD the patterning directly influences capacity of the proteins to phase separate with highest propensity observed for variant with oppositely charged N and C-terminus, while the lowest for charged scrambled variant. The remaining mutants featuring patches of charged residues equally spaced throughout the sequence with non-zero local charge per residue exhibited intermediate behavior in terms of condensate formation (326).

Post-translational modifications, such as phosphorylation or arginine methylation further modulate capability of proteins to from liquid-like organelles by modifying polarity of macromolecular fragments, increasing volume of side-chains and modifying their hydrogen bonding propensity. The phosphorylated chain fragments will locally loose part of their flexibility by virtue of increasing probability of steric clashes and electrostatic interactions, further modifying intrachain dynamics and potential capability of interaction with other partner proteins, leading to modifications in functioning of biologically important assemblies. Crucially, phosphorylation of FMRP or Caprin1, proteins participating in regulation of mRNA stability and translational repression leads to modification of their phase-separation capability with successful formation of condensates under conditions where only one of the partner proteins is phosphorylated, and subsequently modifying deadenylation and translation rates of mRNA in a phosphorylation-dependent manner, while methylation of FMRP leads to granule disassembly and translational activation (359–361). Similarly, phosphorylation of proteins belonging to FUS family typically interferes with their phase-separation propensity (362). On the other hand, arginine methylation within RGG motifs of FUS or Nuage protein Ddx4 reduces its potential to form phase-separated organelles and influences translational diffusion of the proteins in both dilute and concentrated phases (363, 364).

Crowding agents, such as PEG or Dextran influence physicochemical properties of the medium leading to modifications in critical concentration required for phase separation of a number of proteins (365, 366). For instance, PEG induced phase separation of tau protein upon addition of RNA via induction of phase-separation leads to formation of assemblies which do not evolve into neurotoxic oligomeric or fibrillar species (323, 367). Similarly, PEG facilitates formation of 40-amino acid RG-rich peptides bearing primary sequence similarity to FUS, hnRNPA1, and FMRP (328, 361). Attempts in explanation of the phenomenon on basis of scaled-particle theory and excluded volume interactions have been made by Benedek and co-workers (171, 304). Addition of viscogen molecules limits volume available to macromolecules, in turn limiting mean free path of the protein under investigation leading to more frequent collisions, and consequently influencing probability of successful complex formation between phase separating macromolecules. Additionally, if the crowder molecule preferentially interacts with phase separating protein with affinity greater than the one determined by its self-interaction, the viscogen may create an interaction hub allowing for formation of (secondary) nuclei leading to a decrease in saturation concentration. Supplementary contributions may originate from crowder mediated modification to viscosity of the medium, which in turn slows down reorientation rate of protein fragments, and in turn facilitates droplet formation process. The role of the unfolded protein dynamics in initiation, growth and sustenance of liquid-like organelles and modifications of the reorientation processes due to the post-translational modifications and crowding, is still under debate opening potential avenues for further studies.

APPENDIX

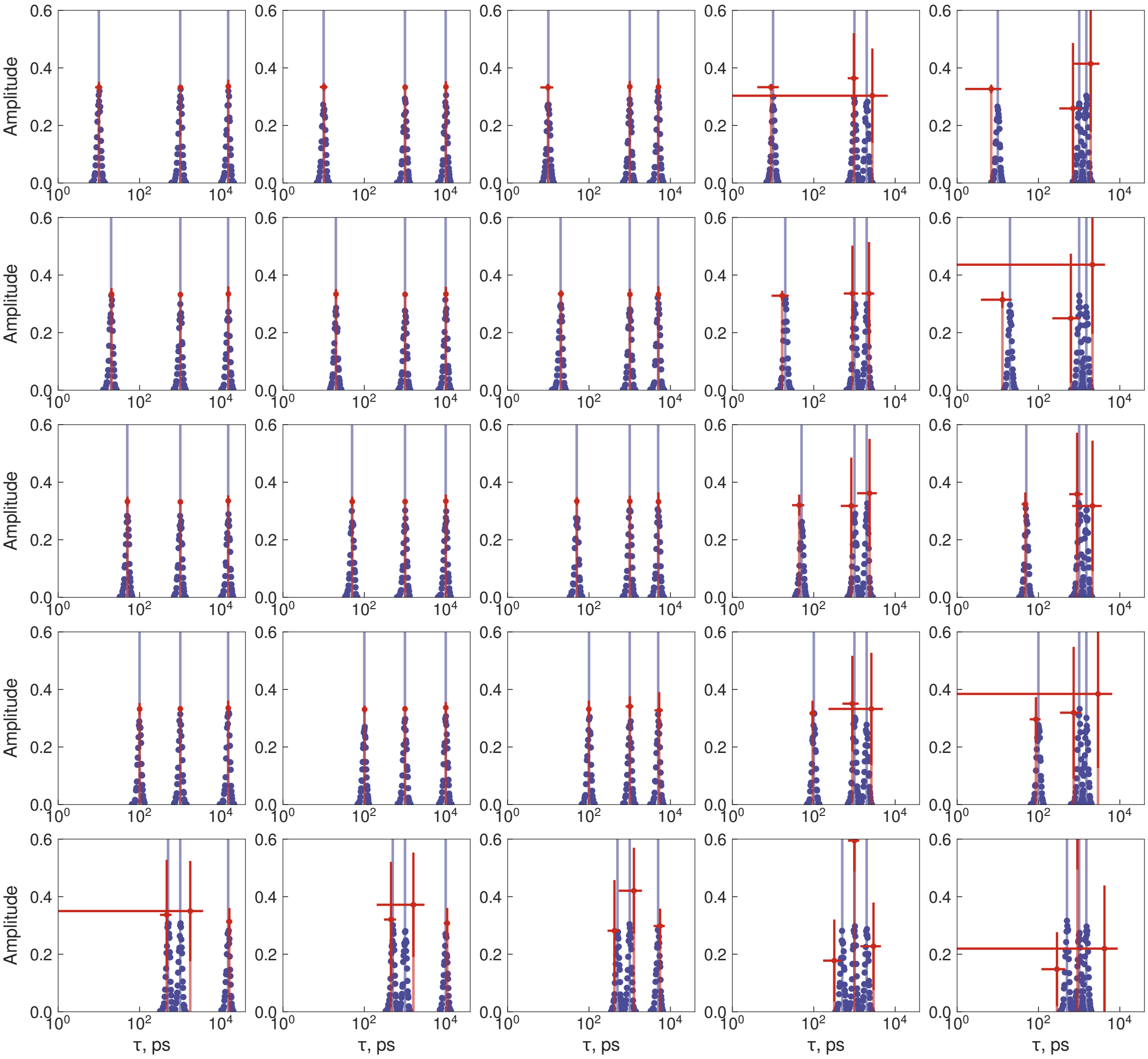

Figure App.1: Comparison of simulated distributions of correlation times used for calculation of relaxation rates with apparent dynamic modes estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated correlation times (Population ratio: 1:1:1, Ω = 0.1). Blue points represent distribution of simulated correlation times for each motion with average represented by blue bar. Orange points and bars represent average and standard deviation of estimated apparent dynamic modes on the basis of Model-Free analysis of relaxation rates calculated from simulated distributions of correlation times.

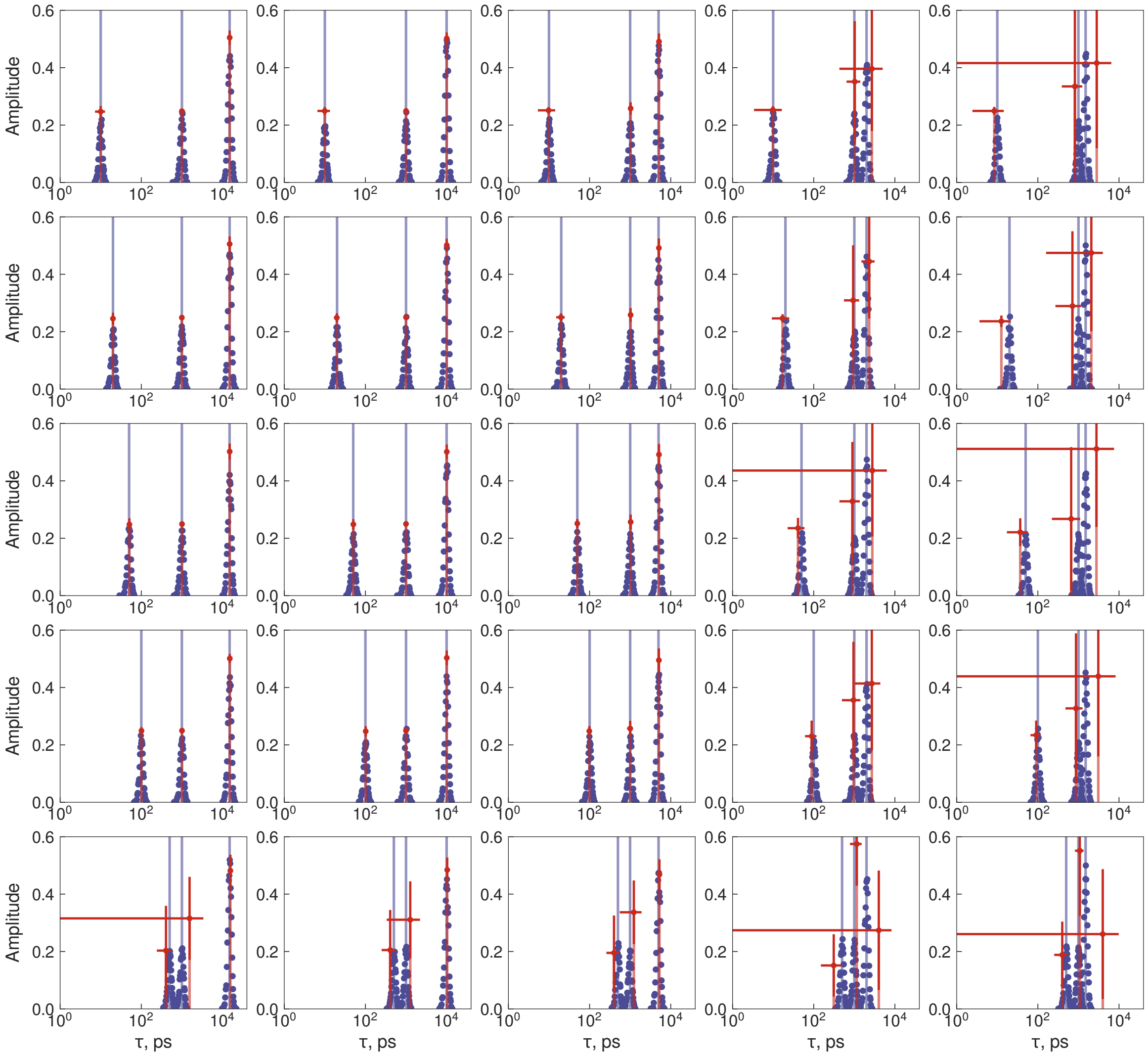

Figure App.2: Comparison of simulated distributions of correlation times used for calculation of relaxation rates with apparent dynamic modes estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated correlation times (Population ratio: 1:1:2, Ω = 0.1). Blue points represent distribution of simulated correlation times for each motion with average represented by blue bar. Orange points and bars represent average and standard deviation of estimated apparent dynamic modes on the basis of Model-Free analysis of relaxation rates calculated from simulated distributions of correlation times.

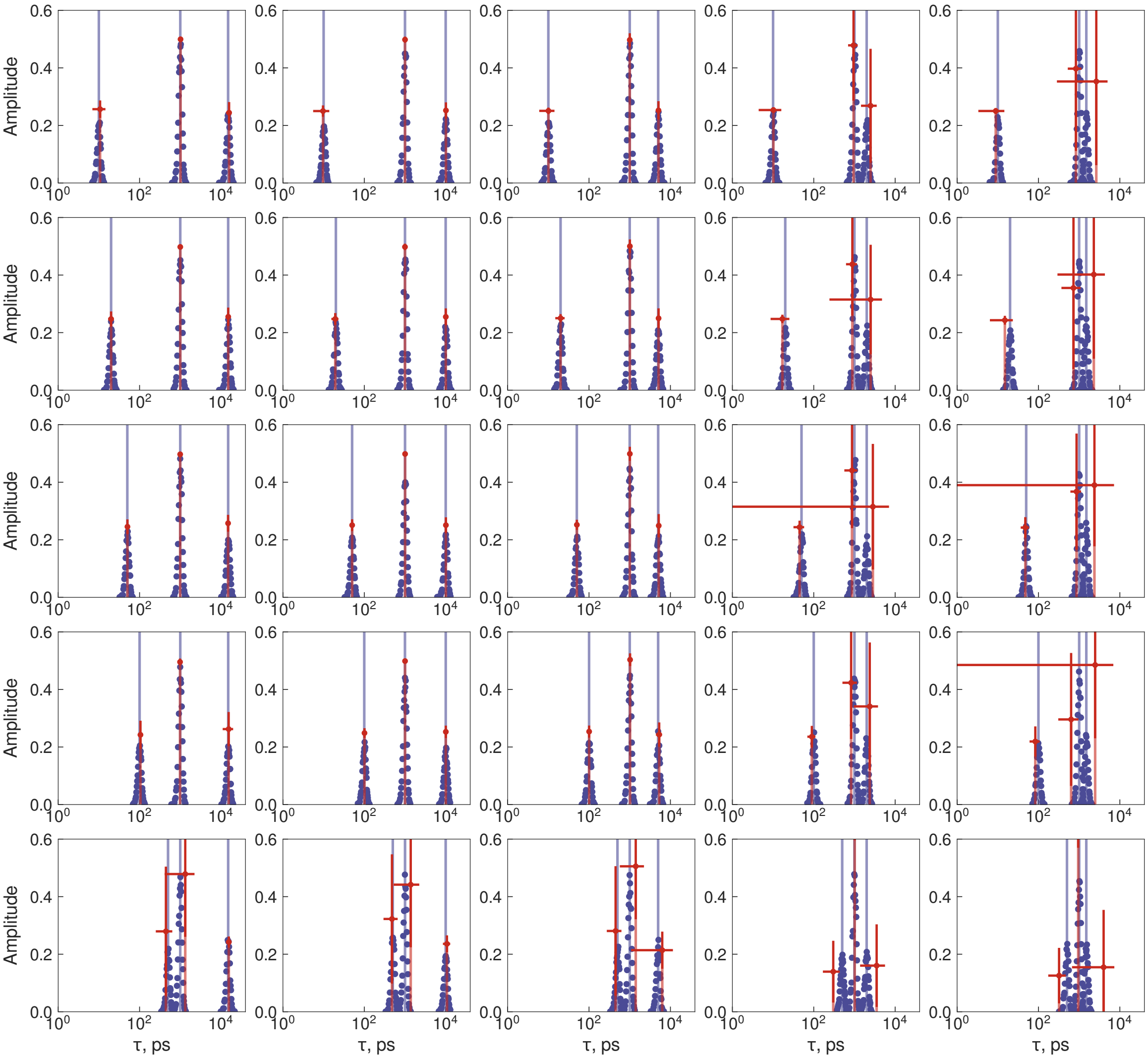

Figure App.3: Comparison of simulated distributions of correlation times used for calculation of relaxation rates with apparent dynamic modes estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated correlation times (Population ratio: 1:2:1, Ω = 0.1). Blue points represent distribution of simulated correlation times for each motion with average represented by blue bar. Orange points and bars represent average and standard deviation of estimated apparent dynamic modes on the basis of Model-Free analysis of relaxation rates calculated from simulated distributions of correlation times.

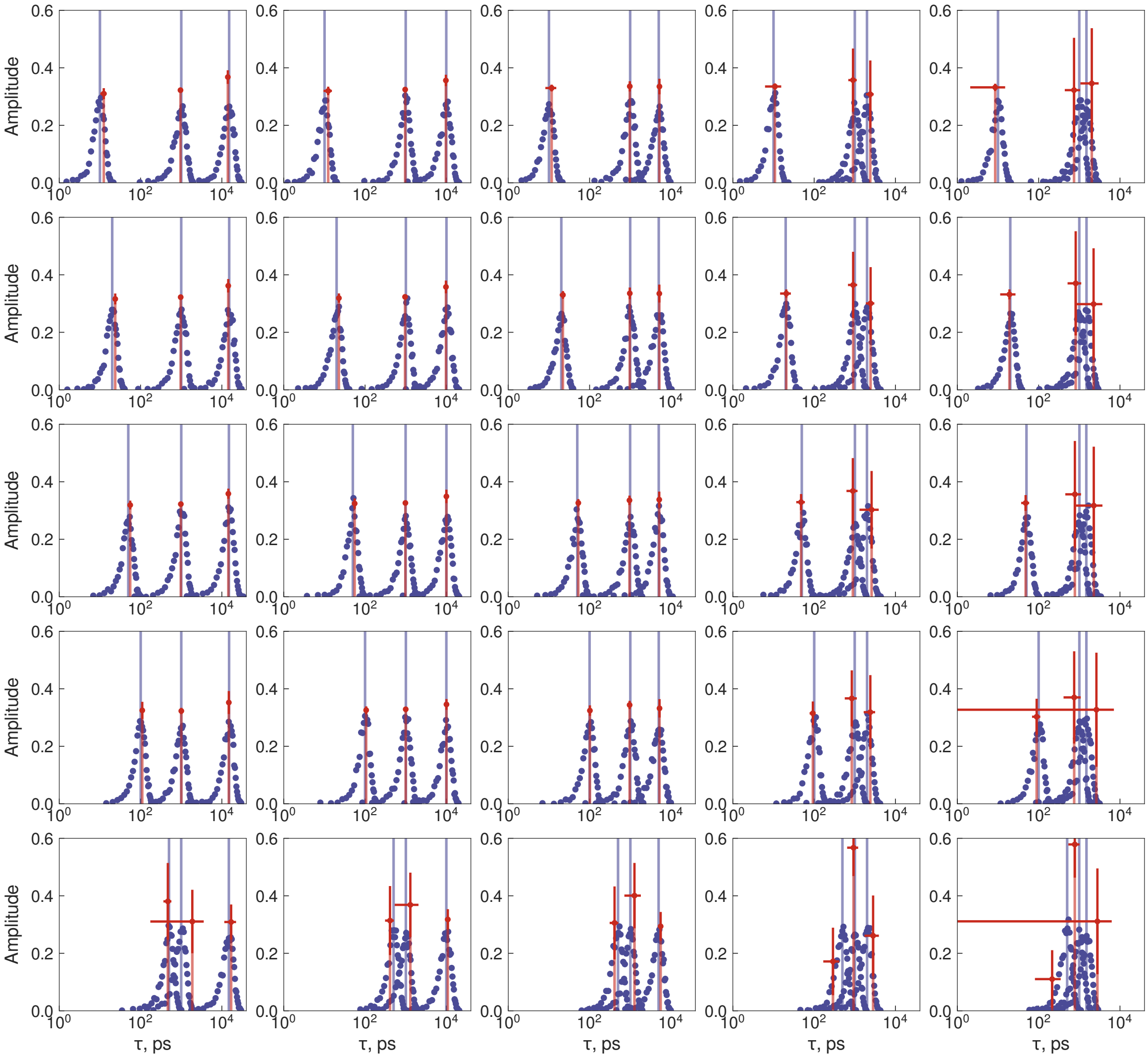

Figure App.4: Comparison of simulated distributions of correlation times used for calculation of relaxation rates with apparent dynamic modes estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated correlation times (Population ratio: 1:1:1, $\Omega = 0.3$). Blue points represent distribution of simulated correlation times for each motion with average represented by blue bar. Orange points and bars represent average and standard deviation of estimated apparent dynamic modes on the basis of Model-Free analysis of relaxation rates calculated from simulated distributions of correlation times.

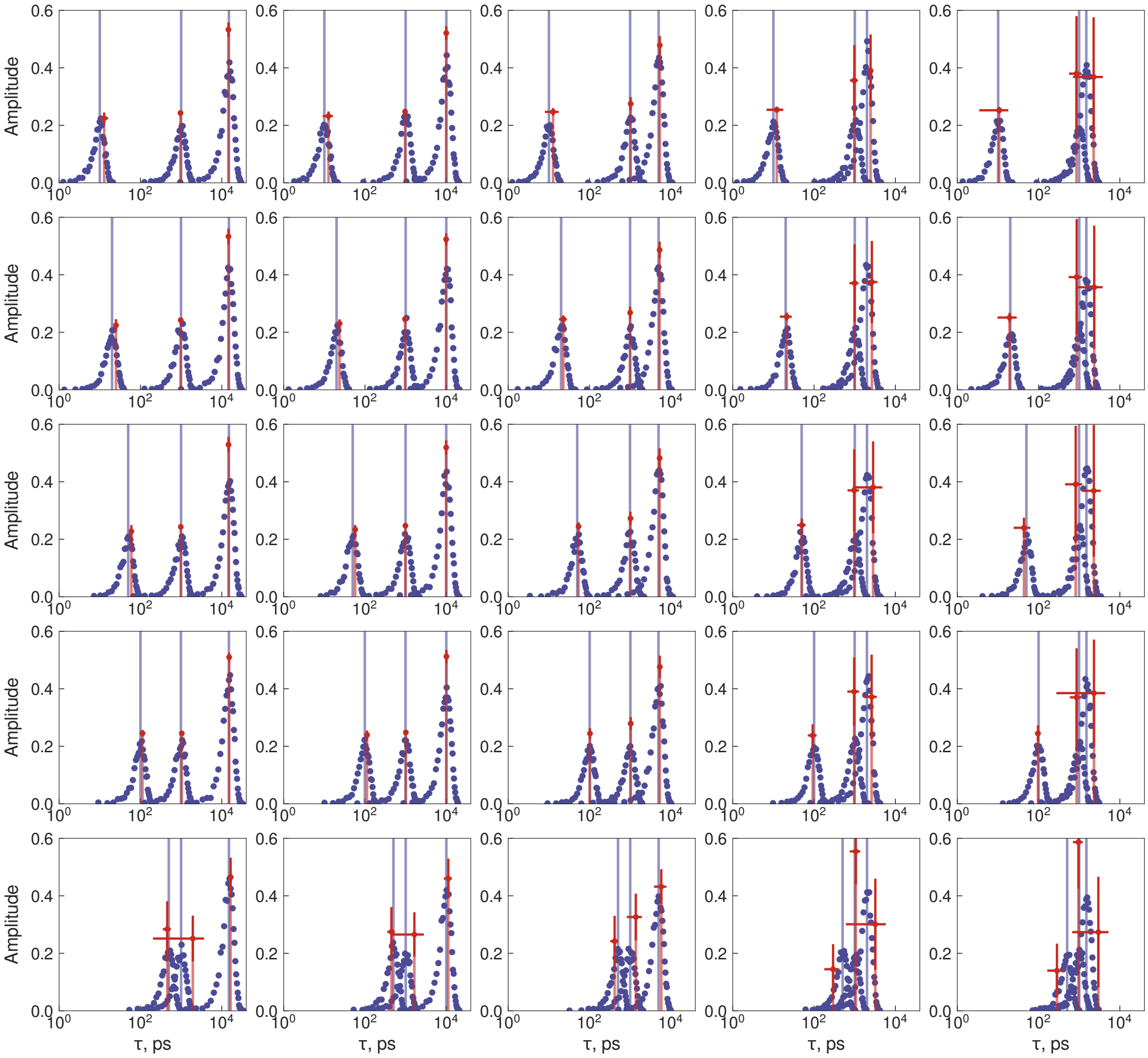

Figure App.5: Comparison of simulated distributions of correlation times used for calculation of relaxation rates with apparent dynamic modes estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated correlation times (Population ratio: 1:1:2, Ω = 0.3). Blue points represent distribution of simulated correlation times for each motion with average represented by blue bar. Orange points and bars represent average and standard deviation of estimated apparent dynamic modes on the basis of Model-Free analysis of relaxation rates calculated from simulated distributions of correlation times.

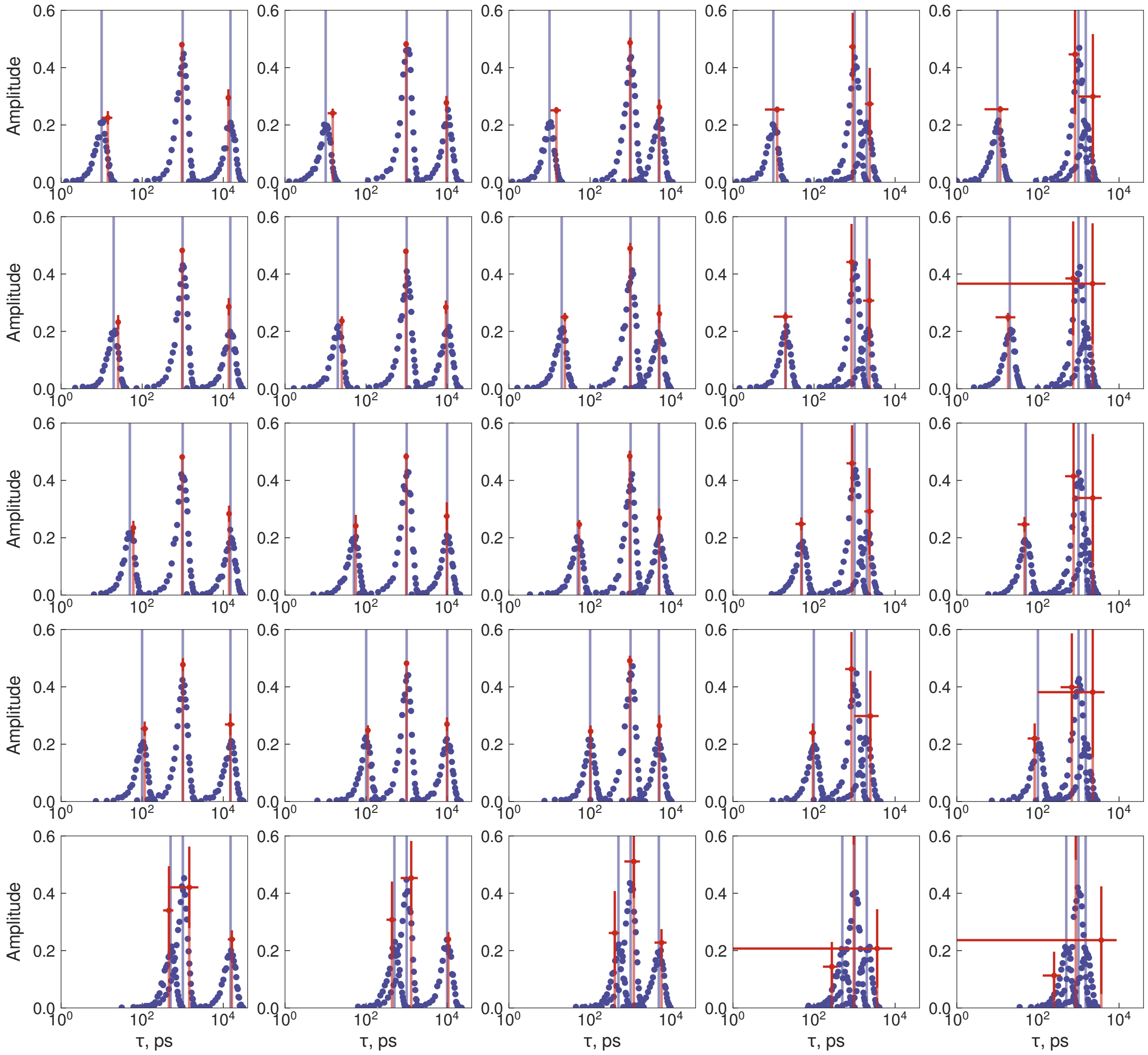

Figure App.6: Comparison of simulated distributions of correlation times used for calculation of relaxation rates with apparent dynamic modes estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated correlation times (Population ratio: 1:2:1, Ω = 0.3). Blue points represent distribution of simulated correlation times for each motion with average represented by blue bar. Orange points and bars represent average and standard deviation of estimated apparent dynamic modes on the basis of Model-Free analysis of relaxation rates calculated from simulated distributions of correlation times.

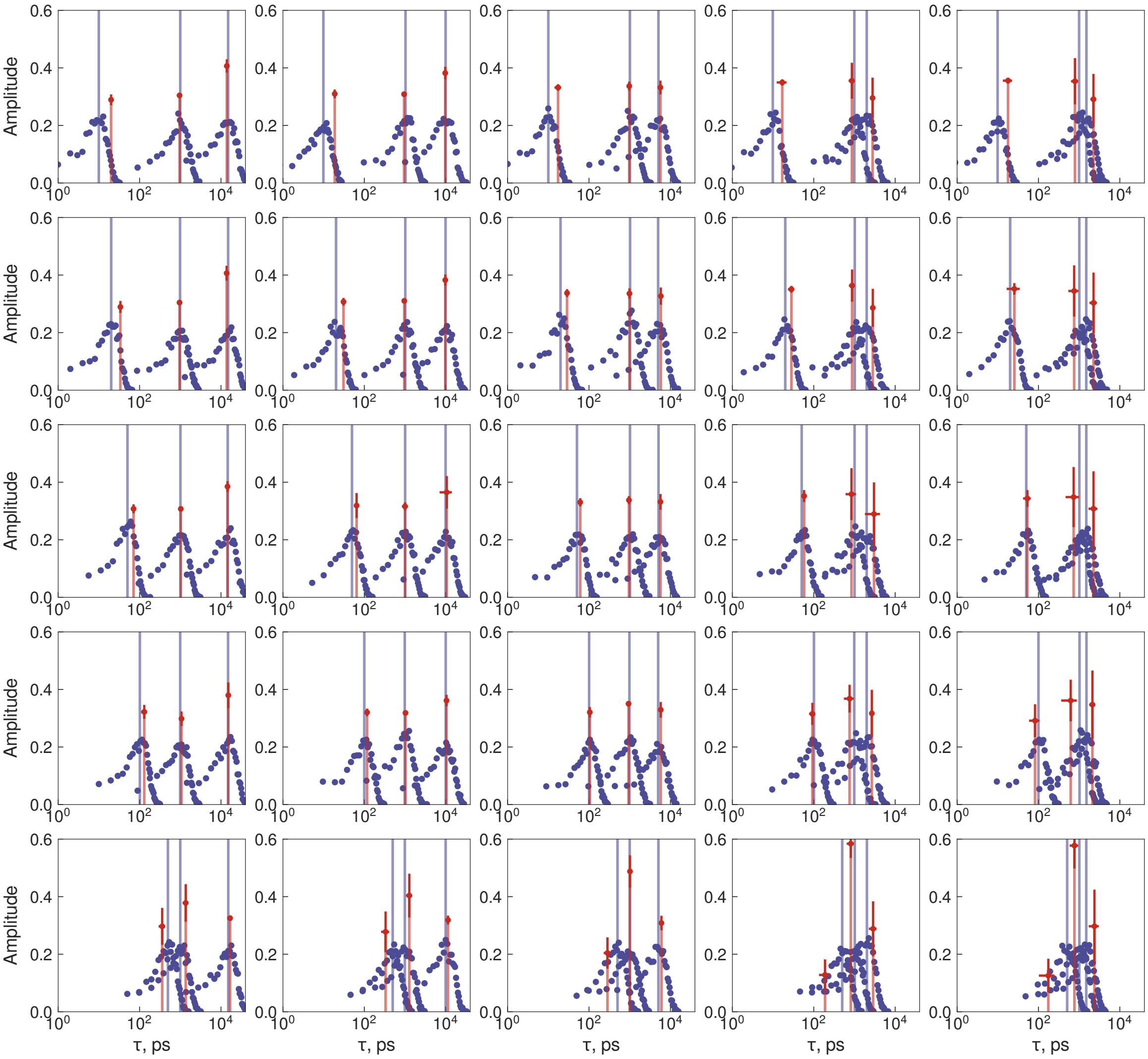

Figure App.7: Comparison of simulated distributions of correlation times used for calculation of relaxation rates with apparent dynamic modes estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated correlation times (Population ratio: 1:1:1, Ω = 0.6). Blue points represent distribution of simulated correlation times for each motion with average represented by blue bar. Orange points and bars represent average and standard deviation of estimated apparent dynamic modes on the basis of Model-Free analysis of relaxation rates calculated from simulated distributions of correlation times.

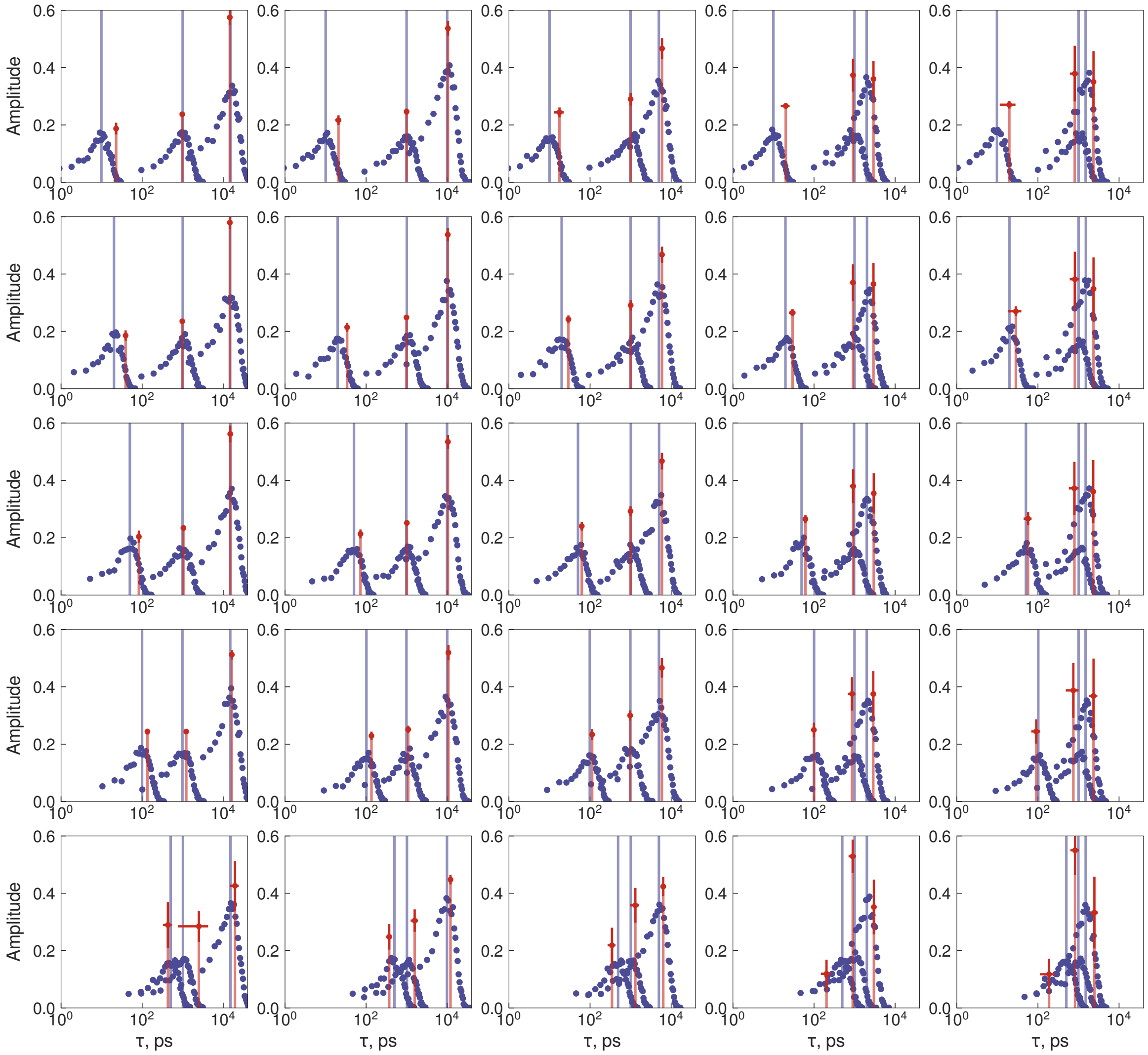

Figure App.8: Comparison of simulated distributions of correlation times used for calculation of relaxation rates with apparent dynamic modes estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated correlation times (Population ratio: 1:1:2, Ω = 0.6). Blue points represent distribution of simulated correlation times for each motion with average represented by blue bar. Orange points and bars represent average and standard deviation of estimated apparent dynamic modes on the basis of Model-Free analysis of relaxation rates calculated from simulated distributions of correlation times.

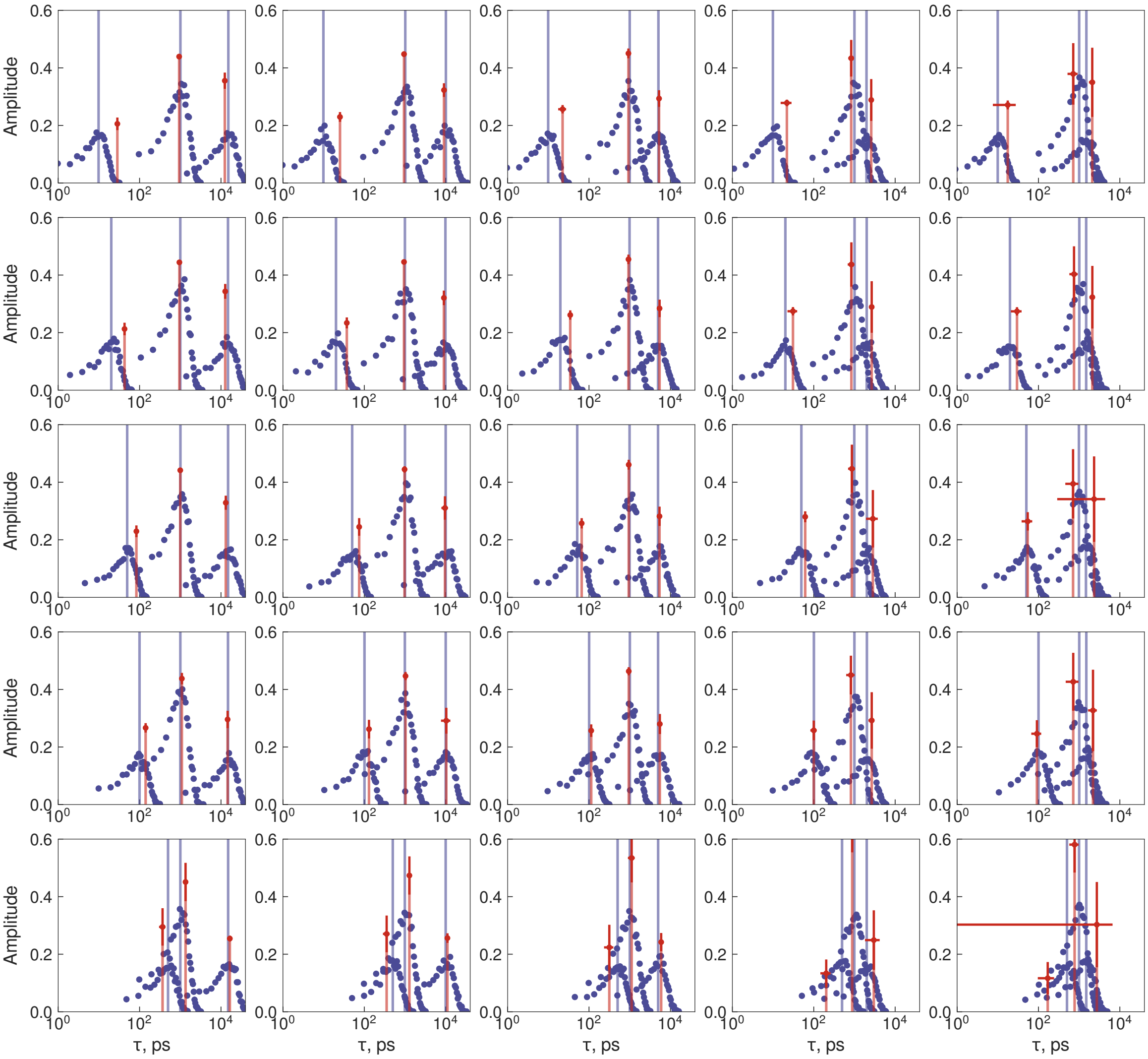

Figure App.9: Comparison of simulated distributions of correlation times used for calculation of relaxation rates with apparent dynamic modes estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated correlation times (Population ratio: 1:2:1, $\Omega = 0.6$). Blue points show simulated values classified in 33 bins for each dynamic mode with average value represented by blue bar. Orange points and bars show average and standard deviation of estimated apparent dynamic modes on the basis of Model-Free analysis of relaxation rates calculated from simulated distributions of correlation times.

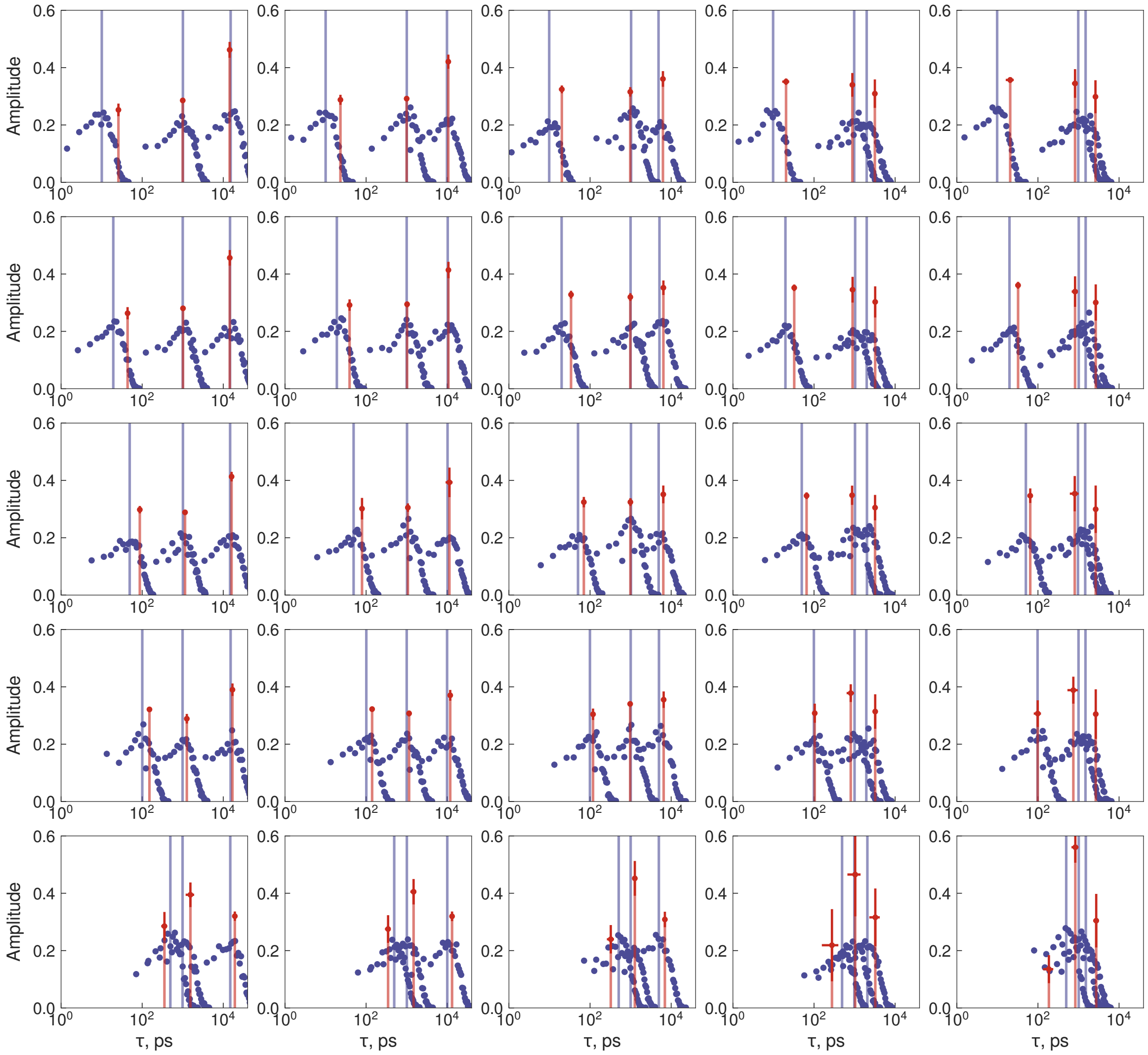

Figure App.10: Comparison of simulated distributions of correlation times used for calculation of relaxation rates with apparent dynamic modes estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated correlation times (Population ratio: 1:1:1, $\Omega = 0.9$). Blue points show distribution of simulated correlation times for each motion with average represented by blue bar. Orange points and bars show average and standard deviation of calculated apparent dynamic modes from Model-Free analysis of relaxation rates calculated on the basis of simulated distribution of timescales and amplitudes of motion.

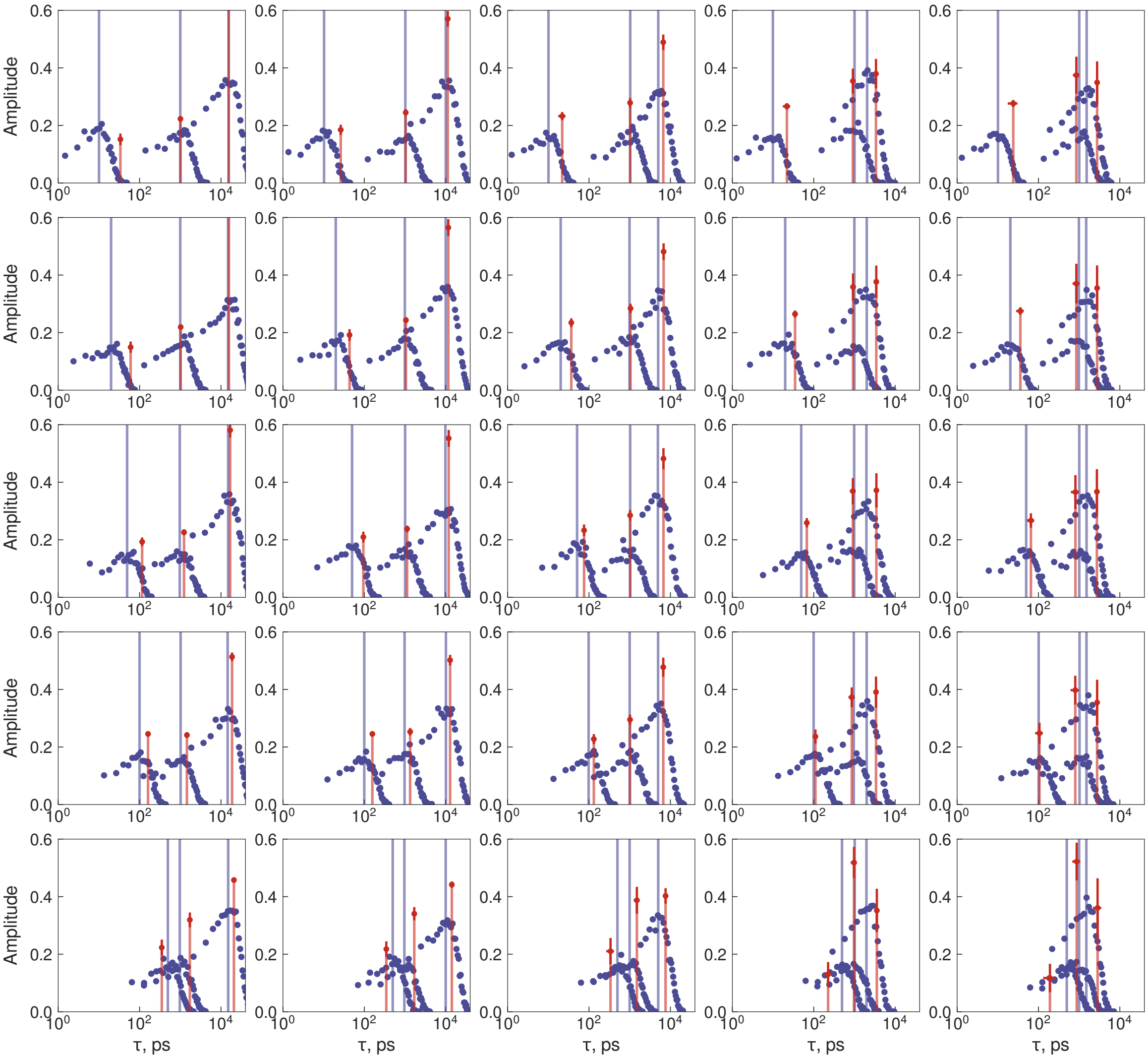

Figure App.11: Comparison of simulated values of dynamic modes used for calculation of relaxation rates with values estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated ones (Population ratio: 1:1:2, Ω = 0.9). Blue points represent distribution of simulated correlation times for each motion with average represented by blue bar. Orange points and bars represent average and standard deviation of estimated apparent dynamic modes on the basis of Model-Free analysis of relaxation rates calculated from simulated distributions of correlation times.

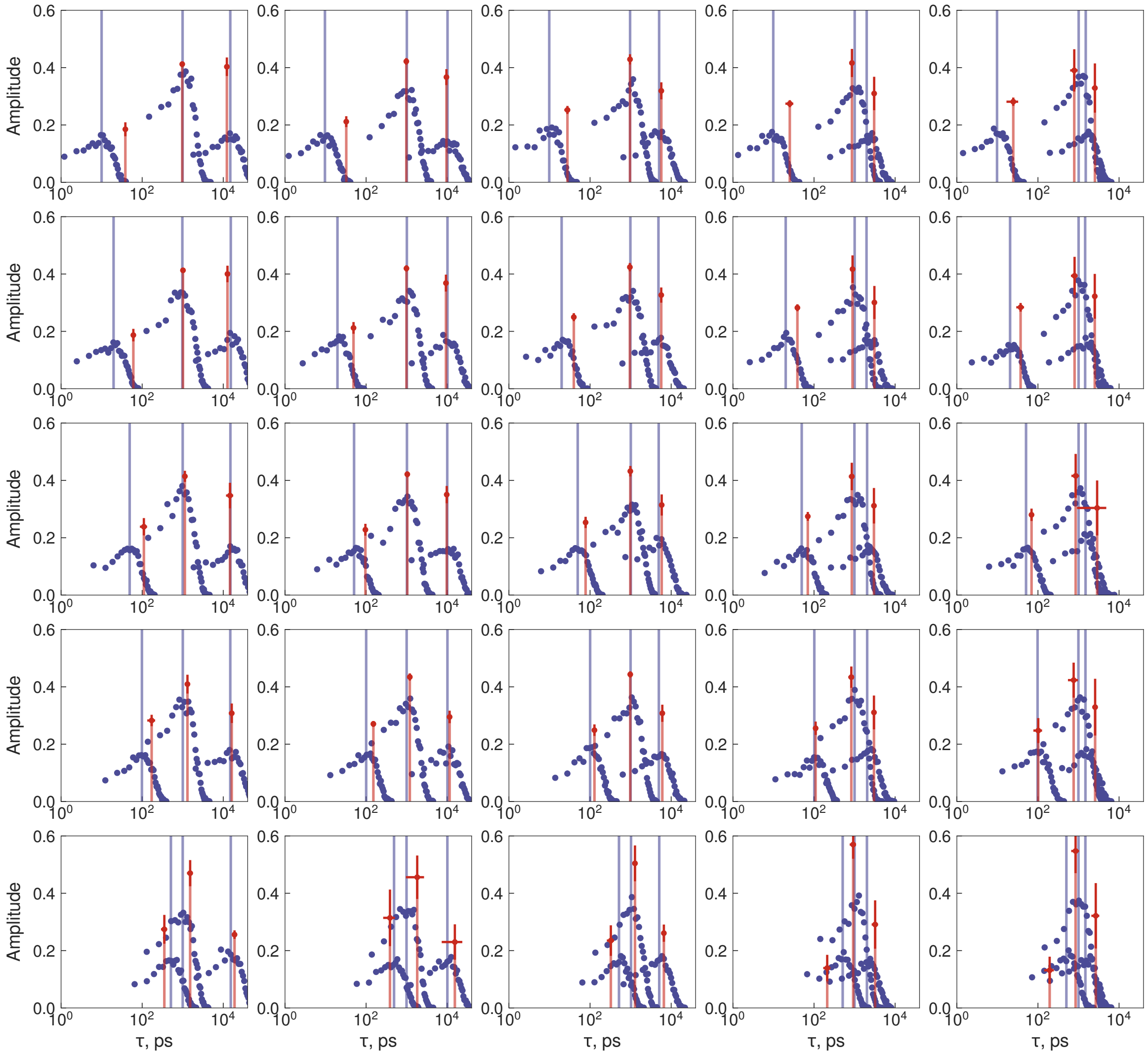

Figure App.12: Comparison of simulated distributions of correlation times used for calculation of relaxation rates with apparent dynamic modes estimated from Model-Free analysis of relaxation rates calculated on the basis of simulated correlation times (Population ratio: 1:2:1, Ω = 0.9). Blue points represent distribution of simulated correlation times for each motion with average represented by blue bar. Orange points and bars represent average and standard deviation of estimated apparent dynamic modes on the basis of Model-Free analysis of relaxation rates calculated from simulated distributions of correlation times.

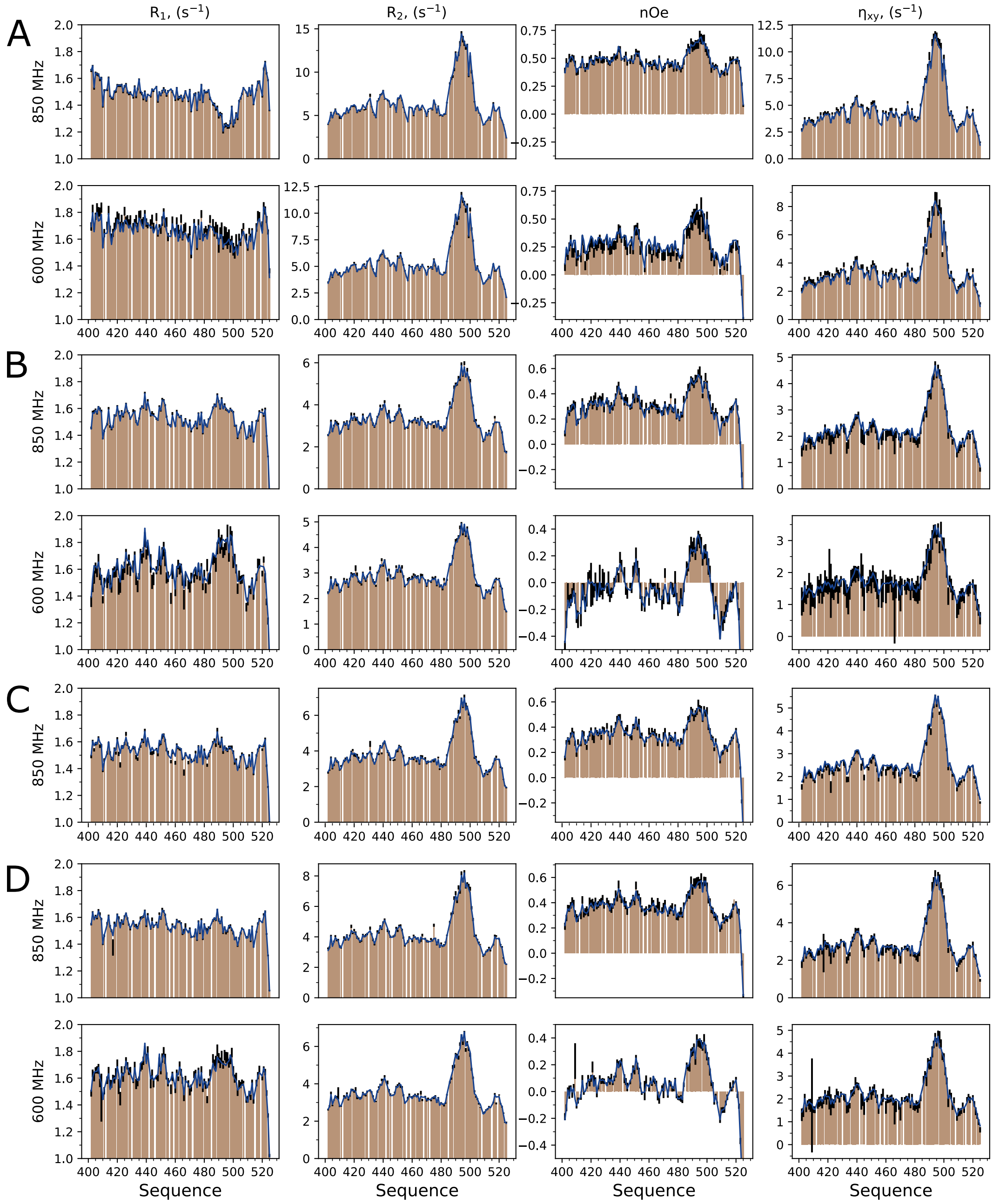

Figure App.13: Best-fit values of the combined Einstein-Arrhenius Model-Free analysis of ^{15}N-backbone relaxation of C-terminal part of Measles Nucleoprotein using Model 3 with equation 5.4. Panels A-D show experimental (bars) and calculated (line) values of relaxation rates at 278K, 0 g/L, 298K 0, 37.5 and 75 g/L.

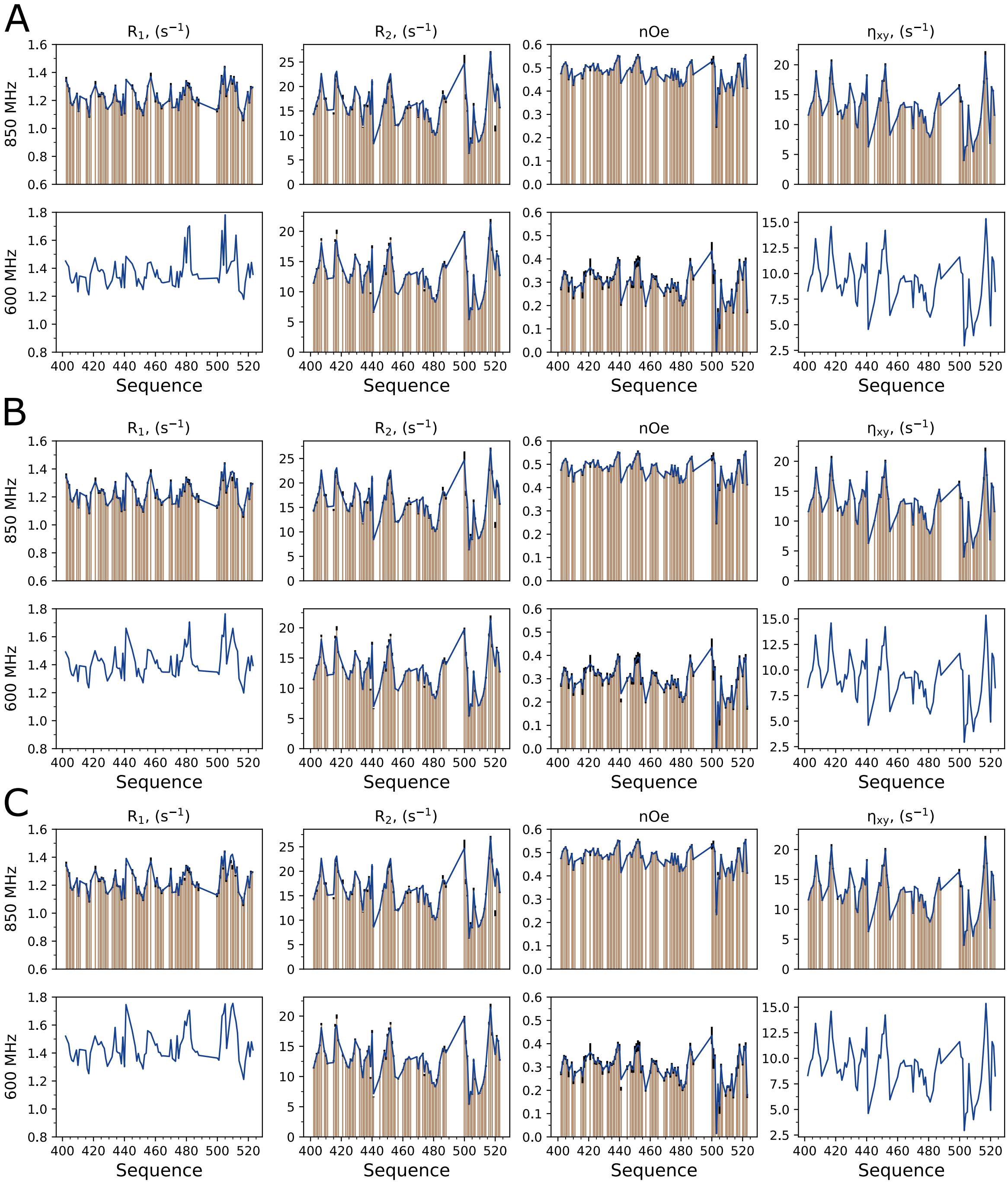

Figure App.14: Best-fit values of independent Model-Free analysis of ^{15}N-backbone relaxation of C-terminal part of Measles Nucleoprotein in the concentrated phase using Model 3, optimizing two timescales, two amplitudes and CSA-DD angle, assuming constant average value of timescales of fast motions equal to 50, 75 or 100 ps (panels A, B and C respectively).

REFERENCES

1. V. N. Uversky, J. R. Gillespie, A. L. Fink, Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* **41**, 415–427, 2000.
2. R. van der Lee *et al.*, Classification of Intrinsically Disordered Regions and Proteins. *Chemical Reviews* **114**, 6589–6631, 2014.
3. H. J. Dyson, P. E. Wright, Intrinsically unstructured proteins and their functions. *Nature Reviews. Molecular Cell Biology* **6**, 197–208, 2005.
4. M. M. Babu *et al.*, Intrinsically disordered proteins: regulation and disease. *Current Opinion in Structural Biology* **21**, 432–440, 2011.
5. A. Dunker *et al.*, Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling* **19**, 26–59, 2001.
6. P. E. Wright, H. J. Dyson, Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *Journal of Molecular Biology* **293**, 321–331, 1999.
7. J. Henriques *et al.*, On the Calculation of SAXS Profiles of Folded and Intrinsically Disordered Proteins from Computer Simulations. *Journal of Molecular Biology* **430**, 2521–2539, 2018.
8. M. R. Hermann, J. S. Hub, SAXS-Restrained Ensemble Simulations of Intrinsically Disordered Proteins with Commitment to the Principle of Maximum Entropy. *Journal of Chemical Theory and Computation* **15**, 5103–5115, 2019.
9. H. D. Mertens, D. I. Svergun, Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *Journal of Structural Biology* **172**, 128–141, 2010.
10. E. D. Holmstrom *et al.*, Accurate Transfer Efficiencies, Distance Distributions, and Ensembles of Unfolded and Intrinsically Disordered Proteins From Single-Molecule FRET. *Methods in Enzymology* **611**, 287–325, 2018.
11. E. Lerner *et al.*, FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. *eLife* **10**, e60416, 2021.

12. I. Nasir, E. P. Bentley, A. A. Deniz, Ratiometric Single-Molecule FRET Measurements to Probe Conformational Subpopulations of Intrinsically Disordered Proteins. *Current Protocols in Chemical Biology* **12**, 80, 2020.
13. L. A. Metskas, E. Rhoades, Single-Molecule FRET of Intrinsically Disordered Proteins. *Annual Review of Physical Chemistry* **71**, 391–414, 2020.
14. S. J. LeBlanc, P. Kulkarni, K. R. Weninger, Single Molecule FRET: A Powerful Tool to Study Intrinsically Disordered Proteins. *Biomolecules* **8**, E140, 2018.
15. J. Song *et al.*, Conformational Heterogeneity and FRET Data Interpretation for Dimensions of Unfolded Proteins. *Biophysical Journal* **113**, 1012–1024, 2017.
16. G. Fuertes *et al.*, Decoupling of size and shape fluctuations in heteropolymeric sequences reconciles discrepancies in SAXS vs. FRET measurements. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E6342–E6351, 2017.
17. G.-N. W. Gomes *et al.*, “Integrating multiple experimental data to determine conformational ensembles of an intrinsically disordered protein”, preprint, Biophysics, Feb. 2020.
18. M. Aznauryan *et al.*, Comprehensive structural and dynamical view of an unfolded protein from the combination of single-molecule FRET, NMR, and SAXS. *Proceedings of the National Academy of Sciences* **113**, E5389–E5398, 2016.
19. A. Soranno *et al.*, Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy. *Proceedings of the National Academy of Sciences* **109**, 17800–17806, 2012.
20. A. Soranno *et al.*, Integrated view of internal friction in unfolded proteins from single-molecule FRET, contact quenching, theory, and simulations. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E1833–E1839, 2017.
21. M. R. Jensen *et al.*, Defining Conformational Ensembles of Intrinsically Disordered and Partially Folded Proteins Directly from Chemical Shifts. *Journal of the American Chemical Society* **132**, 1270–1272, 2010.
22. T. Mittag, J. D. Forman-Kay, Atomic-level characterization of disordered protein ensembles. *Current Opinion in Structural Biology* **17**, 3–14, 2007.
23. M. Levitt, S. Lifson, Refinement of protein conformations using a macromolecular energy minimization procedure. *Journal of Molecular Biology* **46**, 269–279, 1969.

24. J. A. McCammon, B. R. Gelin, M. Karplus, Dynamics of folded proteins. *Nature* **267**, 585–590, 1977.

25. J. Schlitter, M. Engels, P. Krüger, Targeted molecular dynamics: a new approach for searching pathways of conformational transitions. *Journal of Molecular Graphics* **12**, 84–89, 1994.

26. A. Raval *et al.*, Refinement of protein structure homology models via long, all-atom molecular dynamics simulations. *Proteins* **80**, 2071–2079, 2012.

27. P. Robustelli, S. Piana, D. E. Shaw, Developing a molecular dynamics force field for both folded and disordered protein states. *Proceedings of the National Academy of Sciences of the United States of America* **115**, E4758–E4766, 2018.

28. K. Tamiola, B. Acar, F. A. A. Mulder, Sequence-Specific Random Coil Chemical Shifts of Intrinsically Disordered Proteins. *Journal of the American Chemical Society* **132**, 18000–18003, 2010.

29. W. Zhang, D. Ganguly, J. Chen, Residual structures, conformational fluctuations, and electrostatic interactions in the synergistic folding of two intrinsically disordered proteins. *PLoS computational biology* **8**, e1002353, 2012.

30. A. Bax, G. Kontaxis, N. Tjandra, Dipolar couplings in macromolecular structure determination. *Methods in Enzymology* **339**, 127–174, 2001.

31. A. Bax, Weak alignment offers new NMR opportunities to study protein structure and dynamics. *Protein Science* **12**, 1–16, 2003.

32. T. Haliloglu, A. Kolinski, J. Skolnick, Use of residual dipolar couplings as restraints in ab initio protein structure prediction. *Biopolymers* **70**, 548–562, 2003.

33. B. Vögeli, L. Yao, A. Bax, Protein backbone motions viewed by intraresidue and sequential HN-Halpha residual dipolar couplings. *Journal of biomolecular NMR* **41**, 17–28, 2008.

34. J. H. Prestegard, C. M. Bougault, A. I. Kishore, Residual dipolar couplings in structure determination of biomolecules. *Chemical Reviews* **104**, 3519–3540, 2004.

35. J. A. Marsh *et al.*, Calculation of residual dipolar couplings from disordered state ensembles using local alignment. *Journal of the American Chemical Society* **130**, 7804–7805, 2008.

36. M. R. Jensen *et al.*, Exploring Free-Energy Landscapes of Intrinsically Disordered Proteins at Atomic Resolution Using NMR Spectroscopy. *Chemical Reviews* **114**, 6632–6660, 2014.

37. L. Salmon *et al.*, NMR Characterization of Long-Range Order in Intrinsically Disordered Proteins. *Journal of the American Chemical Society* **132**, 8407–8418, 2010.

38. J. Iwahara, C. D. Schwieters, G. M. Clore, Ensemble approach for NMR structure refinement against (1)H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. *Journal of the American Chemical Society* **126**, 5879–5896, 2004.

39. G. Nodet *et al.*, Quantitative Description of Backbone Conformational Sampling of Unfolded Proteins at Amino Acid Resolution from NMR Residual Dipolar Couplings. *Journal of the American Chemical Society* **131**, 17908–17918, 2009.

40. V. Ozenne *et al.*, Flexible-meccano: a tool for the generation of explicit ensemble descriptions of intrinsically disordered proteins and their associated experimental observables. *Bioinformatics* **28**, 1463–1470, 2012.

41. R. I. Cukier, Generating Intrinsically Disordered Protein Conformational Ensembles from a Database of Ramachandran Space Pair Residue Probabilities Using a Markov Chain. *The Journal of Physical Chemistry B* **122**, 9087–9101, 2018.

42. U. R. Shrestha *et al.*, Generation of the configurational ensemble of an intrinsically disordered protein from unbiased molecular dynamics simulation. *Proceedings of the National Academy of Sciences* **116**, 20446–20452, 2019.

43. A. G. Palmer, Probing molecular motion by NMR. *Current Opinion in Structural Biology* **7**, 732–737, 1997.

44. J. Cavanagh, Ed., *Protein NMR spectroscopy: principles and practice* (Academic Press, Amsterdam; Boston, 2nd ed, 2007).

45. R. Ishima, D. A. Torchia, Protein dynamics from NMR. *Nature Structural Biology* **7**, 740–743, 2000.

46. K. J. Walters *et al.*, Characterizing protein-protein complexes and oligomers by nuclear magnetic resonance spectroscopy. *Methods in Enzymology* **339**, 238–258, 2001.

47. L. E. Kay, NMR studies of protein structure and dynamics. *Journal of Magnetic Resonance* **173**, 193–207, 2005.

48. K. Sugase, H. J. Dyson, P. E. Wright, Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025, 2007.

49. P. Vallurupalli *et al.*, Probing conformational dynamics in biomolecules via chemical exchange saturation transfer: a primer. *Journal of biomolecular NMR* **67**, 243–271, 2017.

50. F. A. Mulder *et al.*, Studying excited states of proteins by NMR spectroscopy. *Nature Structural Biology* **8**, 932–935, 2001.

51. N. J. Anthis, G. M. Clore, Visualizing transient dark states by NMR spectroscopy. *Quarterly Reviews of Biophysics* **48**, 35–116, 2015.

52. P. Vallurupalli, G. Bouvignies, L. E. Kay, Studying "Invisible" Excited Protein States in Slow Exchange with a Major State Conformation. *Journal of the American Chemical Society* **134**, 8148–8161, 2012.

53. D. Ban *et al.*, Measuring Dynamic and Kinetic Information in the Previously Inaccessible Supra-tc Window of Nanoseconds to Microseconds by Solution NMR Spectroscopy. *Molecules* **18**, 11904–11937, 2013.

54. T. R. Alderson, L. E. Kay, Unveiling invisible protein states with NMR spectroscopy. *Current Opinion in Structural Biology* **60**, 39–49, 2020.

55. H. Y. Carr, E. M. Purcell, Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments. *Physical Review* **94**, 630–638, 1954.

56. S. Meiboom, D. Gill, Modified Spin-Echo Method for Measuring Nuclear Relaxation Times. *Review of Scientific Instruments* **29**, 688–691, 1958.

57. A. G. Palmer, C. D. Kroenke, J. P. Loria, Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods in Enzymology* **339**, 204–238, 2001.

58. O. Trott, A. G. Palmer, R1rho relaxation outside of the fast-exchange limit. *Journal of Magnetic Resonance* **154**, 157–160, 2002.

59. J. G. Kempf *et al.*, Off-Resonance TROSY (R1rho - R1) for Quantitation of Fast Exchange Processes in Large Proteins. *Journal of the American Chemical Society* **125**, 12064–12065, 2003.

60. P. Lundström, M. Akke, Off-resonance rotating-frame amide proton spin relaxation experiments measuring microsecond chemical exchange in proteins. *Journal of biomolecular NMR* **32**, 163–173, 2005.

61. F. Massi, J. W. Peng, Characterizing Protein Dynamics with NMR R 1rho Relaxation Experiments. *Methods in Molecular Biology* **1688**, 205–221, 2018.

62. A. G. Palmer, H. Koss, Chemical Exchange. *Methods in Enzymology* **615**, 177–236, 2019.

63. G. Bouvignies, L. E. Kay, Measurement of Proton Chemical Shifts in Invisible States of Slowly Exchanging Protein Systems by Chemical Exchange Saturation Transfer. *The Journal of Physical Chemistry B* **116**, 14311–14317, 2012.

64. N. L. Fawzi *et al.*, Probing exchange kinetics and atomic resolution dynamics in high-molecular-weight complexes using dark-state exchange saturation transfer NMR spectroscopy. *Nature Protocols* **7**, 1523–1533, 2012.

65. V. Tugarinov, G. M. Clore, Exchange saturation transfer and associated NMR techniques for studies of protein interactions involving high-molecular-weight systems. *Journal of biomolecular NMR* **73**, 461–469, 2019.

66. H. M. McConnell, Reaction Rates by Nuclear Magnetic Resonance. *The Journal of Chemical Physics* **28**, 430–431, 1958.

67. D. Abergel, A. G. Palmer, Approximate solutions of the Bloch-McConnell equations for two-site chemical exchange. *Chemphyschem: A European Journal of Chemical Physics and Physical Chemistry* **5**, 787–793, 2004.

68. R. Schneider *et al.*, Visualizing the molecular recognition trajectory of an intrinsically disordered protein using multinuclear relaxation dispersion NMR. *Journal of the American Chemical Society* **137**, 1220–1229, 2015.

69. J. Kragelj *et al.*, Structure and dynamics of the MKK7–JNK signaling complex. *Proceedings of the National Academy of Sciences*, 2015.

70. E. Delaforge *et al.*, Deciphering the Dynamic Interaction Profile of an Intrinsically Disordered Protein by NMR Exchange Spectroscopy. *Journal of the American Chemical Society* **140**, 1148–1158, 2018.

71. C. Charlier *et al.*, Structure and Dynamics of an Intrinsically Disordered Protein Region That Partially Folds upon Binding by Chemical-Exchange NMR. *Journal of the American Chemical Society* **139**, 12219–12227, 2017.

72. S. Milles *et al.*, Plasticity of an ultrafast interaction between nucleoporins and nuclear transport receptors. *Cell* **163**, 734–745, 2015.

73. E. Delaforge *et al.*, Large-Scale Conformational Dynamics Control H5N1 Influenza Polymerase PB2 Binding to Importin a. *Journal of the American Chemical Society* **137**, 15122–15134, 2015.

74. A. R. Camacho-Zarco *et al.*, Molecular basis of host-adaptation interactions between influenza virus polymerase PB2 subunit and ANP32A. *Nature Communications* **11**, 3656, 2020.

75. V. Venditti, T. K. Egner, G. M. Clore, Hybrid Approaches to Structural Characterization of Conformational Ensembles of Complex Macromolecular Systems Combining NMR Residual Dipolar Couplings and Solution X-ray Scattering. *Chemical Reviews* **116**, 6305–6322, 2016.

76. H. J. Dyson, Early Strides in NMR Dynamics Measurements. *Biochemistry*, 2021.

77. A. Abragam, *The principles of nuclear magnetism* (Oxford Univ. Pr, Oxford, Repr, 2011).

78. B. Halle, The physical basis of model-free analysis of NMR relaxation data from proteins and complex fluids. *The Journal of Chemical Physics* **131**, 224507, 2009.

79. G. Lipari, A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *Journal of the American Chemical Society* **104**, 4546–4559, 1982.

80. D. T. Braddock *et al.*, Rapid Identification of Medium- to Large-Scale Interdomain Motion in Modular Proteins Using Dipolar Couplings. *Journal of the American Chemical Society* **123**, 8634–8635, 2001.

81. J. L. Baber, A. Szabo, N. Tjandra, Analysis of Slow Interdomain Motion of Macromolecules Using NMR Relaxation Data. *Journal of the American Chemical Society* **123**, 3953–3959, 2001.

82. P. Bernadó, M. Blackledge, Local dynamic amplitudes on the protein backbone from dipolar couplings: toward the elucidation of slower motions in biomolecules. *Journal of the American Chemical Society* **126**, 7760–7761, 2004.

83. B. Fierz *et al.*, Loop formation in unfolded polypeptide chains on the picoseconds to microseconds time scale. *Proceedings of the National Academy of Sciences* **104**, 2163–2168, 2007.

84. A. K. Simorellis, P. F. Flynn, Fast Local Backbone Dynamics of Encapsulated Ubiquitin. *Journal of the American Chemical Society* **128**, 9580–9581, 2006.

85. O. F. Lange *et al.*, Recognition Dynamics Up to Microseconds Revealed from an RDC-Derived Ubiquitin Ensemble in Solution. *Science* **320**, 1471–1475, 2008.

86. S.-H. Bae *et al.*, Prion Proteins with Pathogenic and Protective Mutations Show Similar Structure and Dynamics. *Biochemistry* **48**, 8120–8128, 2009.

87. R. B. Berlow *et al.*, Role of Backbone Dynamics in Modulating the Interactions of Disordered Ligands with the TAZ1 Domain of the CREB-Binding Protein. *Biochemistry* **58**, 1354–1362, 2019.

88. J. L. Lorieau, J. M. Louis, A. Bax, Whole-Body Rocking Motion of a Fusion Peptide in Lipid Bilayers from Size-Dispersed 15N NMR Relaxation. *Journal of the American Chemical Society* **133**, 14184–14187, 2011.

89. N. J. Anthis, G. M. Clore, The Length of the Calmodulin Linker Determines the Extent of Transient Interdomain Association and Target Affinity. *Journal of the American Chemical Society* **135**, 9648–9651, 2013.

90. W.-Y. Choy, D. Shortle, L. E. Kay, Side Chain Dynamics in Unfolded Protein States: an NMR Based 2H Spin Relaxation Study of d131d. *Journal of the American Chemical Society* **125**, 1748–1758, 2003.

91. W.-Y. Choy, L. E. Kay, Probing Residual Interactions in Unfolded Protein States Using NMR Spin Relaxation Techniques: An Application to d131d. *Journal of the American Chemical Society* **125**, 11988–11992, 2003.

92. N. V. Prabhu *et al.*, Dynamics and Entropy of a Calmodulin-Peptide Complex Studied by NMR and Molecular Dynamics. *Biochemistry* **42**, 562–570, 2003.

93. K. Yamasaki *et al.*, Characterization of the internal motions of Escherichia coli ribonuclease HI by a combination of 15N-NMR relaxation analysis and molecular dynamics simulation: examination of dynamic models. *Biochemistry* **34**, 6587–6601, 1995.

94. K. A. Sharp *et al.*, On the relationship between NMR-derived amide order parameters and protein backbone entropy changes. *Proteins: Structure, Function, and Bioinformatics* **83**, 922–930, 2015.

95. E. Johnson, S. A. Showalter, R. Brüschweiler, A Multifaceted Approach to the Interpretation of NMR Order Parameters: A Case Study of a Dynamic a-Helix. *The Journal of Physical Chemistry B* **112**, 6203–6210, 2008.

96. D.-W. Li, S. A. Showalter, R. Brüschweiler, Entropy Localization in Proteins. *The Journal of Physical Chemistry B* **114**, 16036–16044, 2010.

97. M. Akke, *Order Parameters and Conformational Entropies Derived from Combined NMR and Molecular Dynamics Approach* (American Cancer Society, 2014).
98. L. Vugmeyster *et al.*, Temperature-dependent Dynamics of the Villin Headpiece Helical Subdomain, An Unusually Small Thermostable Protein. *Journal of Molecular Biology* **320**, 841–854, 2002.
99. F. Massi, A. G. Palmer, Temperature Dependence of NMR Order Parameters and Protein Dynamics. *Journal of the American Chemical Society* **125**, 11158–11159, 2003.
100. M. L. Verteramo *et al.*, Interplay between Conformational Entropy and Solvation Entropy in Protein–Ligand Binding. *Journal of the American Chemical Society* **141**, 2012–2026, 2019.
101. D. F. Hansen, P. Vallurupalli, L. E. Kay, Measurement of Methyl Group Motional Parameters of Invisible, Excited Protein States by NMR Spectroscopy. *Journal of the American Chemical Society* **131**, 12745–12754, 2009.
102. J. R. Tolman, K. Ruan, NMR Residual Dipolar Couplings as Probes of Biomolecular Dynamics. *Chemical Reviews* **106**, 1720–1736, 2006.
103. K. Chen, N. Tjandra, The Use of Residual Dipolar Coupling in Studying Proteins by NMR. *Topics in current chemistry* **326**, 47–67, 2012.
104. G. Cornilescu, F. Delaglio, A. Bax, Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *Journal of biomolecular NMR* **13**, 289–302, 1999.
105. K. B. Briggman, J. R. Tolman, De Novo Determination of Bond Orientations and Order Parameters from Residual Dipolar Couplings with High Accuracy. *Journal of the American Chemical Society* **125**, 10164–10165, 2003.
106. J. H. Prestegard *et al.*, Determination of Protein Backbone Structures from Residual Dipolar Couplings. *Methods in enzymology* **394**, 175–209, 2005.
107. G. M. Clore, C. D. Schwieters, Amplitudes of Protein Backbone Dynamics and Correlated Motions in a Small a/b Protein: Correspondence of Dipolar Coupling and Heteronuclear Relaxation Measurements. *Biochemistry* **43**, 10678–10691, 2004.
108. G. M. Clore, C. D. Schwieters, Concordance of residual dipolar couplings, backbone order parameters and crystallographic B-factors for a small alpha/beta protein: a unified picture of high probability, fast atomic motions in proteins. *Journal of Molecular Biology* **355**, 879–886, 2006.

109. M. V. Deshmukh *et al.*, Inter-domain orientation and motions in VAT-N explored by residual dipolar couplings and 15N backbone relaxation. *Magnetic resonance in chemistry: MRC* **44 Spec No**, S89–S100, 2006.

110. Y. Ryabov, D. Fushman, Analysis of interdomain dynamics in a two-domain protein using residual dipolar couplings together with 15N relaxation data. *Magnetic resonance in chemistry: MRC* **44**, S143–151, 2006.

111. N.-A. Lakomek *et al.*, Self-consistent residual dipolar coupling based model-free analysis for the robust determination of nanosecond to microsecond protein dynamics. *Journal of Biomolecular NMR* **41**, 139, 2008.

112. L. Salmon *et al.*, Nuclear Magnetic Resonance Provides a Quantitative Description of Protein Conformational Flexibility on Physiologically Important Time Scales. *Biochemistry* **50**, 2735–2747, 2011.

113. T. Bremi, R. Brüschweiler, Locally Anisotropic Internal Polypeptide Backbone Dynamics by NMR Relaxation. *Journal of the American Chemical Society* **119**, 6672–6673, 1997.

114. S. F. Lienin *et al.*, Anisotropic Intramolecular Backbone Dynamics of Ubiquitin Characterized by NMR Relaxation and MD Computer Simulation. *Journal of the American Chemical Society* **120**, 9870–9879, 1998.

115. M. Karplus, Vicinal Proton Coupling in Nuclear Magnetic Resonance. *Journal of the American Chemical Society* **85**, 2870–2871, 1963.

116. M. Karplus, Contact Electron-Spin Coupling of Nuclear Magnetic Moments. *The Journal of Chemical Physics* **30**, 11–15, 1959.

117. A. M. Bonvin *et al.*, Rapid protein fold determination using secondary chemical shifts and cross-hydrogen bond 15N-13C scalar couplings (3hbJNC). *Journal of Biomolecular NMR* **21**, 221–233, 2001.

118. L. E. Kay *et al.*, Correlation between Dynamics and High Affinity Binding in an SH2 Domain Interaction. *Biochemistry* **35**, 361–368, 1996.

119. N. A. Farrow *et al.*, Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology 2 Domain Studied by 15N NMR Relaxation. *Biochemistry* **33**, 5984–6003, 1994.

120. R. B. Fenwick *et al.*, Weak Long-Range Correlated Motions in a Surface Patch of Ubiquitin Involved in Molecular Recognition. *Journal of the American Chemical Society* **133**, 10336–10339, 2011.

121. T. Carlomagno *et al.*, Ubiquitin Backbone Motion Studied via NHN-C’Ca Dipolar-Dipolar and C’-C’Ca / NHN CSA-Dipolar Cross-Correlated Relaxation. *Journal of the American Chemical Society* **122**, 5105–5113, 2000.

122. G. Bouvignies *et al.*, Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bond scalar couplings. *Proceedings of the National Academy of Sciences* **102**, 13885–13890, 2005.

123. G. M. Clore *et al.*, Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *Journal of the American Chemical Society* **112**, 4989–4991, 1990.

124. S.-H. Bae, H. J. Dyson, P. E. Wright, Prediction of the Rotational Tumbling Time for Proteins with Disordered Segments. *Journal of the American Chemical Society* **131**, 6814–6821, 2009.

125. J. J. Prompers, R. Brüschweiler, General Framework for Studying the Dynamics of Folded and Nonfolded Proteins by NMR Relaxation Spectroscopy and MD Simulation. *Journal of the American Chemical Society* **124**, 4522–4534, 2002.

126. J. J. Prompers, R. Brüschweiler, Reorientational Eigenmode Dynamics: A Combined MD/NMR Relaxation Analysis Method for Flexible Parts in Globular Proteins. *Journal of the American Chemical Society* **123**, 7305–7313, 2001.

127. Y. Gu, D.-W. Li, R. Brüschweiler, NMR Order Parameter Determination from Long Molecular Dynamics Trajectories for Objective Comparison with Experiment. *Journal of Chemical Theory and Computation* **10**, 2599–2607, 2014.

128. W. Sinko *et al.*, Population Based Reweighting of Scaled Molecular Dynamics. *The Journal of Physical Chemistry B* **117**, 12759–12768, 2013.

129. R. Crehuet *et al.*, Bayesian-Maximum-Entropy Reweighting of IDP Ensembles Based on NMR Chemical Shifts. *Entropy* **21**, 898, 2019.

130. J. Köfinger *et al.*, Efficient Ensemble Refinement by Reweighting. *Journal of Chemical Theory and Computation* **15**, 3390–3401, 2019.

131. J. Huang, A. D. MacKerell, Force field development and simulations of intrinsically disordered proteins. *Current Opinion in Structural Biology* **48**, 40–48, 2018.

132. S. Rauscher *et al.*, Structural Ensembles of Intrinsically Disordered Proteins Depend Strongly on Force Field: A Comparison to Experiment. *Journal of Chemical Theory and Computation* **11**, 5513–5524, 2015.

133. J. Mu *et al.*, Recent Force Field Strategies for Intrinsically Disordered Proteins. *Journal of Chemical Information and Modeling* **61**, 1037–1047, 2021.

134. S. M. Gopal *et al.*, Conformational Preferences of an Intrinsically Disordered Protein Domain: A Case Study for Modern Force Fields. *The Journal of Physical Chemistry B* **125**, 24–35, 2021.

135. R. B. Best, W. Zheng, J. Mittal, Balanced Protein-Water Interactions Improve Properties of Disordered Proteins and Non-Specific Protein Association. *Journal of Chemical Theory and Computation* **10**, 5113–5124, 2014.

136. N. Salvi, A. Abyzov, M. Blackledge, Analytical Description of NMR Relaxation Highlights Correlated Dynamics in Intrinsically Disordered Proteins. *Angewandte Chemie International Edition* **56**, 14020–14024, 2017.

137. V. Zapletal *et al.*, Choice of Force Field for Proteins Containing Structured and Intrinsically Disordered Regions. *Biophysical Journal* **118**, 1621–1633, 2020.

138. N. Salvi, A. Abyzov, M. Blackledge, Multi-Timescale Dynamics in Intrinsically Disordered Proteins from NMR Relaxation and Molecular Simulation. *The Journal of Physical Chemistry Letters* **7**, 2483–2489, 2016.

139. N. Salvi, A. Abyzov, M. Blackledge, Atomic resolution conformational dynamics of intrinsically disordered proteins from NMR spin relaxation. *Progress in Nuclear Magnetic Resonance Spectroscopy* **102-103**, 43–60, 2017.

140. N. Salvi, A. Abyzov, M. Blackledge, Solvent-dependent segmental dynamics in intrinsically disordered proteins. *Science Advances* **5**, eaax2348, 2019.

141. D. Laage, T. Elsaesser, J. T. Hynes, Water Dynamics in the Hydration Shells of Biomolecules. *Chemical Reviews* **117**, 10694–10725, 2017.

142. Y. Gavrilov, J. D. Leuchter, Y. Levy, On the coupling between the dynamics of protein and water. *Physical Chemistry Chemical Physics* **19**, 8243–8257, 2017.

143. F. Persson, P. Soderhjelm, B. Halle, How proteins modify water dynamics. *The Journal of Chemical Physics* **148**, 215103, 2018.

144. S. Capponi *et al.*, Structural Relaxation Processes and Collective Dynamics of Water in Biomolecular Environments. *The Journal of Physical Chemistry B* **123**, 480–486, 2019.

145. S.-Y. Sheu *et al.*, Surface Topography Effects of Globular Biomolecules on Hydration Water. *The Journal of Physical Chemistry B* **123**, 6917–6932, 2019.

146. K. Lindorff-Larsen *et al.*, Picosecond to Millisecond Structural Dynamics in Human Ubiquitin. *The Journal of Physical Chemistry B* **120**, 8313–8320, 2016.

147. S. A. Showalter, R. Brüschweiler, Validation of Molecular Dynamics Simulations of Biomolecules Using NMR Spin Relaxation as Benchmarks:Application to the AMBER99SB Force Field. *Journal of Chemical Theory and Computation* **3**, 961–975, 2007.

148. S. F. Lienin, R. Brüschweiler, Characterization of collective and anisotropic reorientational protein dynamics. *Physical Review Letters* **84**, 5439–5442, 2000.

149. D. Long *et al.*, Toward a Predictive Understanding of Slow Methyl Group Dynamics in Proteins. *Biophysical Journal* **101**, 910–915, 2011.

150. S. A. Showalter *et al.*, Toward Quantitative Interpretation of Methyl Side-Chain Dynamics from NMR by Molecular Dynamics Simulations. *Journal of the American Chemical Society* **129**, 14146–14147, 2007.

151. S. F. Cousin *et al.*, Time-Resolved Protein Side-Chain Motions Unraveled by High-Resolution Relaxometry and Molecular Dynamics Simulations. *Journal of the American Chemical Society* **140**, 13456–13465, 2018.

152. X. Liao *et al.*, Probing Side-Chain Dynamics in Proteins by the Measurement of Nine Deuterium Relaxation Rates Per Methyl Group. *The Journal of Physical Chemistry B* **116**, 606–620, 2012.

153. S. C. L. Kamerlin *et al.*, Coarse-grained (multiscale) simulations in studies of biophysical and chemical systems. *Annual Review of Physical Chemistry* **62**, 41–64, 2011.

154. C. Clementi, Coarse-grained models of protein folding: toy models or predictive tools? *Current Opinion in Structural Biology* **18**, 10–15, 2008.

155. B. Stauch, J. Orts, T. Carlomagno, The description of protein internal motions aids selection of ligand binding poses by the INPHARMA method. *Journal of Biomolecular Nmr* **54**, 245–256, 2012.

156. P. J. Nichols *et al.*, The Exact Nuclear Overhauser Enhancement: Recent Advances. *Molecules* **22**, E1176, 2017.

157. C. A. Smith *et al.*, Enhancing NMR derived ensembles with kinetics on multiple timescales. *Journal of Biomolecular NMR* **74**, 27–43, 2020.

158. J.-r. Huang, S. Grzesiek, Ensemble Calculations of Unstructured Proteins Constrained by RDC and PRE Data: A Case Study of Urea-Denatured Ubiquitin. *Journal of the American Chemical Society* **132**, 694–705, 2010.

159. M. M. Dedmon *et al.*, Mapping Long-Range Interactions in a-Synuclein using Spin-Label NMR and Ensemble Molecular Dynamics Simulations. *Journal of the American Chemical Society* **127**, 476–477, 2005.

160. J. Iwahara, C. Tang, G. Marius Clore, Practical aspects of 1H transverse paramagnetic relaxation enhancement measurements on macromolecules. *Journal of Magnetic Resonance* **184**, 185–195, 2007.

161. A. Ceccon, G. Marius Clore, V. Tugarinov, Towards interpretation of intermolecular paramagnetic relaxation enhancement outside the fast exchange limit. *Journal of Biomolecular NMR* **66**, 1–7, 2016.

162. J. Iwahara, G. M. Clore, Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature* **440**, 1227–1230, 2006.

163. T. K. Karamanos, V. Tugarinov, G. M. Clore, Unraveling the structure and dynamics of the human DNAJB6b chaperone by NMR reveals insights into Hsp40-mediated proteostasis. *Proceedings of the National Academy of Sciences* **116**, 21529–21538, 2019.

164. C. Hartlmüller *et al.*, NMR characterization of solvent accessibility and transient structure in intrinsically disordered proteins. *Journal of Biomolecular NMR* **73**, 305–317, 2019.

165. Y. Okuno, A. Szabo, G. M. Clore, Quantitative Interpretation of Solvent Paramagnetic Relaxation for Probing Protein–Cosolute Interactions. *Journal of the American Chemical Society* **142**, 8281–8290, 2020.

166. A. Einstein, Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik* **322**, 549–560, 1905.

167. G. Peskir, On the Diffusion Coefficient: The Einstein Relation and Beyond. *Stochastic Models* **19**, 383–405, 2003.

168. G. D. J. Phillies, Dynamics of polymers in concentrated solutions: the universal scaling equation derived. *Macromolecules* **20**, 558–564, 1987.

169. P. J. Flory, Treatment of the effect of excluded volume and deduction of unperturbed dimensions of polymer chains. Configurational parameters for cellulose derivatives. *Die Makromolekulare Chemie* **98**, 128–135, 1966.

170. A. K. Soper, The excluded volume effect in confined fluids and liquid mixtures. *Journal of Physics: Condensed Matter* **9**, 2399–2410, 1997.

171. O. Annunziata *et al.*, Effect of polyethylene glycol on the liquid–liquid phase transition in aqueous protein solutions. *Proceedings of the National Academy of Sciences* **99**, 14165–14170, 2002.

172. L.-H. Cai, S. Panyukov, M. Rubinstein, Mobility of Nonsticky Nanoparticles in Polymer Liquids. *Macromolecules* **44**, 7853–7863, 2011.

173. A. Wisniewska *et al.*, Scaling Equation for Viscosity of Polymer Mixtures in Solutions with Application to Diffusion of Molecular Probes. *Macromolecules* **50**, 4555–4561, 2017.

174. T. Kalwarczyk *et al.*, Motion of nanoprobes in complex liquids within the framework of the length-scale dependent viscosity model. *Advances in Colloid and Interface Science* **223**, 55–63, 2015.

175. P. J. Flory, S. Fisk, Effect of Volume Exclusion on the Dimensions of Polymer Chains. *The Journal of Chemical Physics* **44**, 2243–2248, 1966.

176. A. P. Minton, Models for Excluded Volume Interaction between an Unfolded Protein and Rigid Macromolecular Cosolutes: Macromolecular Crowding and Protein Stability Revisited. *Biophysical Journal* **88**, 971–985, 2005.

177. H. Hofmann *et al.*, Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proceedings of the National Academy of Sciences* **109**, 16155–16160, 2012.

178. F. Zosel *et al.*, Depletion interactions modulate the binding between disordered proteins in crowded environments. *Proceedings of the National Academy of Sciences* **117**, 13480–13489, 2020.

179. G. Rivas, A. P. Minton, Macromolecular crowding in vitro, in vivo, and in between. *Trends in biochemical sciences* **41**, 970–981, 2016.

180. H.-X. Zhou, G. Rivas, A. P. Minton, Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annual review of biophysics* **37**, 375–397, 2008.

181. A. Soranno *et al.*, Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments. *Proceedings of the National Academy of Sciences* **111**, 4874–4879, 2014.

182. T. P. Lodge, M. Muthukumar, Physical Chemistry of Polymers: Entropy, Interactions, and Dynamics. *The Journal of Physical Chemistry* **100**, 13275–13292, 1996.

183. Y.-H. Lin, J. D. Forman-Kay, H. S. Chan, Sequence-Specific Polyampholyte Phase Separation in Membraneless Organelles. *Physical Review Letters* **117**, 178101, 2016.

184. E. W. Martin *et al.*, Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694–699, 2020.

185. S. Das *et al.*, Comparative roles of charge,n, and hydrophobic interactions in sequence-dependent phase separation of intrinsically disordered proteins. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 28795–28805, 2020.

186. J. Sheu-Gruttadauria, I. J. MacRae, Phase Transitions in the Assembly and Function of Human miRISC. *Cell* **173**, 946–957, 2018.

187. T. Yoshizawa *et al.*, Nuclear Import Receptor Inhibits Phase Separation of FUS through Binding to Multiple Sites. *Cell* **173**, 693–705, 2018.

188. M. H. Levitt, *Spin dynamics: basics of nuclear magnetic resonance* (John Wiley & Sons, Chichester, England; Hoboken, NJ, 2nd ed, 2008).

189. M Goldman, Interference effects in the relaxation of a pair of unlike nuclei. *Journal of Magnetic Resonance* **60**, 437–452, 1984.

190. J. R. Tolman, J. H. Prestegard, Simultaneous collection of two transverse 15N relaxation pathways in isotopically labeled proteins. *Journal of Magnetic Resonance. Series B* **106**, 97–100, 1995.

191. N. Tjandra, A. Szabo, A. Bax, Protein Backbone Dynamics and 15N Chemical Shift Anisotropy from Quantitative Measurement of Relaxation Interference Effects. *Journal of the American Chemical Society* **118**, 6986–6991, 1996.

192. C. D. Kroenke *et al.*, Longitudinal and Transverse 1H-15N Dipolar/15N Chemical Shift Anisotropy Relaxation Interference:Unambiguous Determination of Rotational Diffusion Tensors and Chemical Exchange Effects in Biological Macromolecules. *Journal of the American Chemical Society* **120**, 7905–7915, 1998.

193. K. Loth, P. Pelupessy, G. Bodenhausen, Chemical shift anisotropy tensors of carbonyl, nitrogen, and amide proton nuclei in proteins through cross-correlated relaxation in NMR spectroscopy. *Journal of the American Chemical Society* **127**, 6062–6068, 2005.

194. P. Damberg, J. Jarvet, A. Gräslund, Limited Variations in 15N CSA Magnitudes and Orientations in Ubiquitin Are Revealed by Joint Analysis of Longitudinal and Transverse NMR Relaxation. *Journal of the American Chemical Society* **127**, 1995–2005, 2005.

195. P. Pelupessy, G. M. Espallargas, G. Bodenhausen, Symmetrical reconversion: measuring cross-correlation rates with enhanced accuracy. *Journal of Magnetic Resonance* **161**, 258–264, 2003.

196. J. W. Peng, G. Wagner, Mapping of spectral density functions using heteronuclear NMR relaxation measurements. *Journal of Magnetic Resonance* **98**, 308–332, 1992.

197. R. Ishima, K. Nagayama, Quasi-Spectral-Density Function Analysis for Nitrogen-15 Nuclei in Proteins. *Journal of Magnetic Resonance, Series B* **108**, 73–76, 1995.

198. N. A. Farrow *et al.*, Spectral density function mapping using 15N relaxation data exclusively. *Journal of Biomolecular NMR* **6**, 153–162, 1995.

199. J.-F. Lefèvre *et al.*, Internal Mobility in the Partially Folded DNA Binding and Dimerization Domains of GAL4: NMR Analysis of the N-H Spectral Density Functions. *Biochemistry* **35**, 2674–2686, 1996.

200. P. Kadeřávek *et al.*, Spectral density mapping protocols for analysis of molecular motions in disordered proteins. *Journal of Biomolecular NMR* **58**, 193–207, 2014.

201. S. N. Khan *et al.*, Distribution of Pico- and Nanosecond Motions in Disordered Proteins from Nuclear Spin Relaxation. *Biophysical Journal* **109**, 988–999, 2015.

202. D. M. LeMaster, Larmor frequency selective model free analysis of protein NMR relaxation. *Journal of Biomolecular NMR* **6**, 366–374, 1995.

203. K. Modig, F. M. Poulsen, Model-independent interpretation of NMR relaxation data for unfolded proteins: the acid-denatured state of ACBP. *Journal of biomolecular NMR* **42**, 163–177, 2008.

204. F. Ochsenbein *et al.*, Dynamical characterization of residual and non-native structures in a partially folded protein by 15N NMR relaxation using a model based on a distribution of correlation times. *Protein Science* **11**, 957–964, 2002.

205. A. Hsu, F. Ferrage, A. G. Palmer, Analysis of NMR Spin-Relaxation Data Using an Inverse Gaussian Distribution Function. *Biophysical Journal* **115**, 2301–2309, 2018.

206. A. A. Smith, M. Ernst, B. H. Meier, Because the Light is Better Here: Correlation-Time Analysis by NMR Spectroscopy. *Angewandte Chemie* **129**, 13778–13783, 2017.

207. A. A. Smith, M. Ernst, B. H. Meier, Optimized "detectors" for dynamics analysis in solid-state NMR. *The Journal of Chemical Physics* **148**, 045104, 2018.

208. A. A. Smith *et al.*, Reducing bias in the analysis of solution-state NMR data with dynamics detectors. *The Journal of Chemical Physics* **151**, 034102, 2019.

209. A. A. Smith *et al.*, How wide is the window opened by high-resolution relaxometry on the internal dynamics of proteins in solution? *Journal of Biomolecular NMR* **75**, 119–131, 2021.

210. N.-A. Lakomek, J. Ying, A. Bax, Measurement of 15N relaxation rates in perdeuterated proteins by TROSY-based methods. *Journal of Biomolecular NMR* **53**, 209–221, 2012.

211. P. Pelupessy, F. Ferrage, G. Bodenhausen, Accurate measurement of longitudinal cross-relaxation rates in nuclear magnetic resonance. *The Journal of Chemical Physics* **126**, 134508, 2007.

212. P. Kadeřávek *et al.*, Cross-correlated relaxation measurements under adiabatic sweeps: determination of local order in proteins. *Journal of biomolecular NMR* **63**, 353–365, 2015.

213. E. J. d'Auvergne, P. R. Gooley, Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces. *Journal of Biomolecular NMR* **40**, 107–119, 2008.

214. C. Charlier *et al.*, Nanosecond Time Scale Motions in Proteins Revealed by High-Resolution NMR Relaxometry. *Journal of the American Chemical Society* **135**, 18665–18672, 2013.

215. F. Kateb, P. Pelupessy, G. Bodenhausen, Measuring fast hydrogen exchange rates by NMR spectroscopy. *Journal of Magnetic Resonance* **184**, 108–113, 2007.

216. S. A. Showalter, K. B. Hall, in *Methods in Enzymology* (Elsevier, 2005), vol. 394, pp. 465–480.

217. Y. Gu, D.-W. Li, R. Brüschweiler, Decoding the Mobility and Time Scales of Protein Loops. *Journal of Chemical Theory and Computation* **11**, 1308–1314, 2015.

218. J. E. Fuchs *et al.*, Independent Metrics for Protein Backbone and Side-Chain Flexibility: Time Scales and Effects of Ligand Binding. *Journal of Chemical Theory and Computation* **11**, 851–860, 2015.

219. K. R. Peddireddy *et al.*, Unexpected entanglement dynamics in semidilute blends of supercoiled and ring DNA. *Soft Matter* **16**, 152–161, 2020.

220. D. André d'Avignon *et al.*, Thermodynamics and Kinetics of a Folded-Folded Transition at Valine-9 of a GCN4-Like Leucine Zipper. *Biophysical Journal* **76**, 2752–2759, 1999.

221. D. R. Dee, M. T. Woodside, Comparing the energy landscapes for native folding and aggregation of PrP. *Prion* **10**, 207–220, 2016.

222. J. C. Schlatterer *et al.*, Mapping the Kinetic Barriers of a Large RNA Molecule's Folding Landscape. *PLOS ONE* **9**, e85041, 2014.

223. A. R. Fadel *et al.*, Crankshaft motions of the polypeptide backbone in molecular dynamics simulations of human type-a transforming growth factor. *Journal of Biomolecular NMR* **6**, 221–226, 1995.

224. P. Guerry, L. Mollica, M. Blackledge, Mapping Protein Conformational Energy Landscapes Using NMR and Molecular Simulation. *ChemPhysChem* **14**, 3046–3058, 2013.

225. R. B. Fenwick, C. D. Schwieters, B. Vogeli, Direct Investigation of Slow Correlated Dynamics in Proteins via Dipolar Interactions. *Journal of the American Chemical Society* **138**, 8412–8421, 2016.

226. J. R. Lewandowski *et al.*, Direct observation of hierarchical protein dynamics. *Science* **348**, 578–581, 2015.

227. G. Bouvignies, P. R. L. Markwick, M. Blackledge, Characterization of protein dynamics from residual dipolar couplings using the three dimensional Gaussian axial fluctuation model. *Proteins* **71**, 353–363, 2008.

228. S.-L. Chang, N. Tjandra, Temperature dependence of protein backbone motion from carbonyl 13C and amide 15N NMR relaxation. *Journal of Magnetic Resonance* **174**, 43–53, 2005.

229. L. Vugmeyster, D. Ostrovsky, Temperature dependence of fast carbonyl backbone dynamics in chicken villin headpiece subdomain. *Journal of Biomolecular NMR* **50**, 119–127, 2011.

230. A. M. Mandel, M. Akke, A. G. Palmer, Dynamics of ribonuclease H: temperature dependence of motions on multiple time scales. *Biochemistry* **35**, 16009–16023, 1996.

231. M. J. Seewald *et al.*, The role of backbone conformational heat capacity in protein stability: temperature dependent dynamics of the B1 domain of Streptococcal protein G. *Protein Science : A Publication of the Protein Society* **9**, 1177–1193, 2000.

232. L. Vugmeyster, C. J. McKnight, Slow Motions in Chicken Villin Headpiece Subdomain Probed by Cross-Correlated NMR Relaxation of Amide NH Bonds in Successive Residues. *Biophysical Journal* **95**, 5941–5950, 2008.

233. T. Zeiske, K. A. Stafford, A. G. Palmer, Thermostability of Enzymes from Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation* **12**, 2489–2492, 2016.

234. C. Bracken *et al.*, Temperature dependence of intramolecular dynamics of the basic leucine zipper of GCN4: implications for the entropy of association with DNA 1 1Edited by P. E. Wright. *Journal of Molecular Biology* **285**, 2133–2146, 1999.

235. A. Abyzov *et al.*, Identification of Dynamic Modes in an Intrinsically Disordered Protein Using Temperature-Dependent NMR Relaxation. *Journal of the American Chemical Society* **138**, 6240–6251, 2016.

236. J.-r. Huang *et al.*, Direct Prediction of NMR Residual Dipolar Couplings from the Primary Sequence of Unfolded Proteins. *Angewandte Chemie International Edition* **52**, 687–690, 2013.

237. P. Bernado *et al.*, A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. *Proceedings of the National Academy of Sciences* **102**, 17002–17007, 2005.

238. M. Schwalbe *et al.*, Predictive Atomic Resolution Descriptions of Intrinsically Disordered hTau40 and a-Synuclein in Solution from NMR and Small Angle Scattering. *Structure* **22**, 238–249, 2014.

239. E. Johnson, A. G. Palmer, M. Rance, Temperature dependence of the NMR generalized order parameter. *Proteins: Structure, Function, and Bioinformatics* **66**, 796–803, 2007.

240. S. Pistolesi, N. Tjandra, Temperature Dependence of Molecular Interactions Involved in Defining Stability of Glutamine Binding Protein and Its Complex with l-Glutamine. *Biochemistry* **51**, 643–652, 2012.

241. H. Schwalbe *et al.*, Structural and Dynamical Properties of a Denatured Protein. Heteronuclear 3D NMR Experiments and Theoretical Simulations of Lysozyme in 8 M Urea. *Biochemistry* **36**, 8977–8991, 1997.

242. J. Klein-Seetharaman *et al.*, Long-range interactions within a nonnative protein. *Science* **295**, 1719–1722, 2002.

243. F.-X. Theillet *et al.*, Structural disorder of monomeric a-synuclein persists in mammalian cells. *Nature* **530**, 45–50, 2016.

244. K. Melkova *et al.*, Functionally specific binding regions of microtubule-associated protein 2c exhibit distinct conformations and dynamics. *Journal of Biological Chemistry* **293**, 13297–13309, 2018.

245. U. Baul, M. Bley, J. Dzubiella, Thermal Compaction of Disordered and Elastin-like Polypeptides: A Temperature-Dependent, Sequence-Specific Coarse-Grained Simulation Model. *Biomacromolecules* **21**, 3523–3538, 2020.

246. M.-K. Cho *et al.*, Amino Acid Bulkiness Defines the Local Conformations and Dynamics of Natively Unfolded a-Synuclein and Tau. *Journal of the American Chemical Society* **129**, 3032–3033, 2007.

247. J. J. Rooney, Eyring transition-state theory and kinetics in catalysis. *Journal of Molecular Catalysis A: Chemical* **96**, L1–L3, 1995.

248. A. J. B. Spaull, The theory of polymer dynamics M. Doi and S. F. Edwards. *Journal of Chemical Technology & Biotechnology* **44**, 79–80, 1989.

249. C. H. Ziener, T. Kampf, F. T. Kurz, Diffusion propagators for hindered diffusion in open geometries. *Concepts in Magnetic Resonance Part A* **44**, 150–159, 2015.

250. D. Guin, M. Gruebele, Weak Chemical Interactions That Drive Protein Evolution: Crowding, Sticking, and Quinary Structure in Folding and Function. *Chemical Reviews* **119**, 10691–10717, 2019.

251. S. B. Zimmerman, S. O. Trach, Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. *Journal of Molecular Biology* **222**, 599–620, 1991.

252. C. E. Ioan, T. Aberle, W. Burchard, Solution Properties of Glycogen. 2. Semidilute Solutions. *Macromolecules* **32**, 8655–8662, 1999.

253. C. E. Ioan, T. Aberle, W. Burchard, Light Scattering and Viscosity Behavior of Dextran in Semidilute Solution. *Macromolecules* **34**, 326–336, 2001.

254. T. Coviello *et al.*, Static and Dynamic Light Scattering by a Thermoreversible Gel from Rhizobium leguminosarum 8002 Exopolysaccharide. *Macromolecules* **30**, 2008–2015, 1997.

255. A. E. Smith *et al.*, In-cell thermodynamics and a new role for protein surfaces. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 1725–1730, 2016.

256. R. D. Cohen, G. J. Pielak, Electrostatic Contributions to Protein Quinary Structure. *Journal of the American Chemical Society* **138**, 13139–13142, 2016.

257. P. Zielenkiewicz, W. Saenger, Residue solvent accessibilities in the unfolded polypeptide chain. *Biophysical Journal* **63**, 1483–1486, 1992.

258. P. Bernadó, M. Blackledge, J. Sancho, Sequence-Specific Solvent Accessibilities of Protein Residues in Unfolded Protein Ensembles. *Biophysical Journal* **91**, 4536–4543, 2006.

259. S. Kaieda, B. Halle, Internal Water and Microsecond Dynamics in Myoglobin. *The Journal of Physical Chemistry B* **117**, 14676–14687, 2013.

260. J.-F. Bodart *et al.*, NMR observation of Tau in Xenopus oocytes. *Journal of Magnetic Resonance* **192**, 252–257, 2008.

261. C. A. Waudby *et al.*, In-Cell NMR Characterization of the Secondary Structure Populations of a Disordered Conformation of a-Synuclein within E. coli Cells. *PLOS ONE* **8**, e72286, 2013.

262. Q. Wang, A. Zhuravleva, L. M. Gierasch, Exploring Weak, Transient Protein–Protein Interactions in Crowded In Vivo Environments by In-Cell Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **50**, 9225–9236, 2011.

263. M. P. Latham, L. E. Kay, Probing non-specific interactions of Ca2+-calmodulin in E. coli lysate. *Journal of Biomolecular NMR* **55**, 239–247, 2013.

264. P. B. Crowley, E. Chow, T. Papkovskaia, Protein Interactions in the Escherichia coli Cytosol: An Impediment to In-Cell NMR Spectroscopy. *ChemBioChem* **12**, 1043–1048, 2011.

265. A. Binolfi, F.-X. Theillet, P. Selenko, Bacterial in-cell NMR of human a-synuclein: a disordered monomer by nature? *Biochemical Society Transactions* **40**, 950–954, 2012.

266. P. Selenko *et al.*, Quantitative NMR analysis of the protein G B1 domain in Xenopus laevis egg extracts and intact oocytes. *Proceedings of the National Academy of Sciences* **103**, 11904–11909, 2006.

267. C. Li *et al.*, Differential Dynamical Effects of Macromolecular Crowding on an Intrinsically Disordered Protein and a Globular Protein: Implications for In-Cell NMR Spectroscopy. *Journal of the American Chemical Society* **130**, 6310–6311, 2008.

268. E. Luchinat, L. Banci, In-Cell NMR in Human Cells: Direct Protein Expression Allows Structural Studies of Protein Folding and Maturation. *Accounts of Chemical Research* **51**, 1550–1557, 2018.

269. T. Müntener *et al.*, In-Cell Protein Structures from 2D NMR Experiments. *The Journal of Physical Chemistry Letters* **7**, 2821–2825, 2016.

270. F.-X. Theillet *et al.*, Physicochemical Properties of Cells and Their Effects on Intrinsically Disordered Proteins (IDPs). *Chemical Reviews* **114**, 6661–6714, 2014.

271. R Zwanzig, Time-Correlation Functions and Transport Coefficients in Statistical Mechanics. *Annual Review of Physical Chemistry* **16**, 67–102, 1965.

272. R. I. Cukier, Diffusion of Brownian spheres in semidilute polymer solutions. *Macromolecules* **17**, 252–255, 1984.

273. R. G. Bryant, in *New Developments in NMR*, ed. by R. Kimmich (Royal Society of Chemistry, Cambridge, 2018), pp. 207–228.

274. R. Zwanzig, Hydrodynamic fluctuations and Stokes' law friction. *Journal of Research of the National Bureau of Standards Section B Mathematics and Mathematical Physics* **68B**, 143, 1964.

275. T. Kalwarczyk *et al.*, Comparative Analysis of Viscosity of Complex Liquids and Cytoplasm of Mammalian Cells at the Nanoscale. *Nano Letters* **11**, 2157–2163, 2011.

276. D. Laage, J. T. Hynes, Reorientional dynamics of water molecules in anionic hydration shells. *Proceedings of the National Academy of Sciences* **104**, 11167–11172, 2007.

277. A. Wisniewska *et al.*, Scaling of activation energy for macroscopic flow in poly(ethylene glycol) solutions: Entangled – Non-entangled crossover. *Polymer* **55**, 4651–4657, 2014.

278. O. Rahaman *et al.*, Configurational Disorder of Water Hydrogen-Bond Network at the Protein Dynamical Transition. *The Journal of Physical Chemistry B* **121**, 6792–6798, 2017.

279. G. Zaccai, The effect of water on protein dynamics. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **359**, 1269–1275; discussion 1275, 1323–1328, 2004.

280. M. Tros *et al.*, Picosecond orientational dynamics of water in living cells. *Nature Communications* **8**, 904, 2017.

281. R. Brüschweiler, P. E. Wright, Water self-diffusion model for protein-water NMR cross relaxation. *Chemical Physics Letters* **229**, 75–81, 1994.

282. N. V. Nucci, M. S. Pometun, A. J. Wand, Site-resolved measurement of water-protein interactions by solution NMR. *Nature Structural & Molecular Biology* **18**, 245–249, 2011.

283. F. Persson, B. Halle, Transient Access to the Protein Interior: Simulation versus NMR. *Journal of the American Chemical Society* **135**, 8735–8748, 2013.

284. P. W. Fenimore *et al.*, Slaving: solvent fluctuations dominate protein dynamics and functions. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 16047–16051, 2002.

285. S. Genheden, Effect of solvent model when probing protein dynamics with molecular dynamics. *Journal of Molecular Graphics and Modelling* **71**, 80–87, 2017.

286. R. Barnes *et al.*, Spatially Heterogeneous Surface Water Diffusivity around Structured Protein Surfaces at Equilibrium. *Journal of the American Chemical Society* **139**, 17890–17901, 2017.

287. F. Pizzitutti *et al.*, How Protein Surfaces Induce Anomalous Dynamics of Hydration Water. *The Journal of Physical Chemistry B* **111**, 7584–7590, 2007.

288. F. Persson, P. Soderhjelm, B. Halle, The spatial range of protein hydration. *The Journal of Chemical Physics* **148**, 215104, 2018.

289. D. R. Martin, J. E. Forsmo, D. V. Matyushov, Complex Dynamics of Water in Protein Confinement. *The Journal of Physical Chemistry B* **122**, 3418–3425, 2018.

290. D. Laage, J. T. Hynes, A Molecular Jump Mechanism of Water Reorientation. *Science* **311**, 832–835, 2006.

291. J. O. Richardson *et al.*, Concerted hydrogen-bond breaking by quantum tunneling in the water hexamer prism. *Science* **351**, 1310–1313, 2016.

292. O. Bignucolo *et al.*, Backbone Hydration Determines the Folding Signature of Amino Acid Residues. *Journal of the American Chemical Society* **137**, 4300–4303, 2015.

293. S. A. Dames *et al.*, Residual Dipolar Couplings in Short Peptides Reveal Systematic Conformational Preferences of Individual Amino Acids. *Journal of the American Chemical Society* **128**, 13508–13514, 2006.

294. A. H. Mao *et al.*, Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proceedings of the National Academy of Sciences* **107**, 8183–8188, 2010.

295. O. Byutner, G. D. Smith, Temperature and molecular weight dependence of the zero shear-rate viscosity of an entangled polymer melt from simulation and theory. *Journal of Polymer Science Part B: Polymer Physics* **39**, 3067–3071, 2001.

296. E. Luchinat, L. Banci, In-cell NMR: a topical review. *IUCrJ* **4**, 108–118, 2017.

297. C. Kang, Applications of In-Cell NMR in Structural Biology and Drug Discovery. *International Journal of Molecular Sciences* **20**, 139, 2019.

298. R. Hänsel *et al.*, In-Cell NMR and EPR Spectroscopy of Biomacromolecules. *Angewandte Chemie International Edition* **53**, 10300–10314, 2014.

299. I. Yu *et al.*, Biomolecular interactions modulate macromolecular structure and dynamics in atomistic model of a bacterial cytoplasm. *eLife* **5**, ed. by Y. Shan, e19274, 2016.

300. A. Bhattarai, I. A. Emerson, Dynamic conformational flexibility and molecular interactions of intrinsically disordered proteins. *Journal of Biosciences* **45**, 29, 2020.

301. I. Konig *et al.*, Impact of In-Cell and In-Vitro Crowding on the Conformations and Dynamics of an Intrinsically Disordered Protein. *Angewandte Chemie International Edition* **60**, 10724–10729, 2021.

302. S. Majumder *et al.*, Probing Protein Quinary Interactions by In-Cell Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **54**, 2727–2738, 2015.

303. B. C. McNulty, G. B. Young, G. J. Pielak, Macromolecular Crowding in the Escherichia coli Periplasm Maintains a-Synuclein Disorder. *Journal of Molecular Biology* **355**, 893–897, 2006.

304. S. Zhang *et al.*, In-Cell NMR Study of Tau and MARK2 Phosphorylated Tau. *International Journal of Molecular Sciences* **20**, 90, 2019.

305. X. Song *et al.*, Characterization of Residue Specific Protein Folding and Unfolding Dynamics in Cells. *Journal of the American Chemical Society* **141**, 11363–11366, 2019.

306. L. Banci *et al.*, Molecular chaperone function of Mia40 triggers consecutive induced folding steps of the substrate in mitochondrial protein import. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 20190–20195, 2010.

307. L. Banci *et al.*, In-cell NMR in E. coli to Monitor Maturation Steps of hSOD1. *PLOS ONE* **6**, e23561, 2011.

308. L. Banci *et al.*, Functional role of two interhelical disulfide bonds in human Cox17 protein from a structural perspective. *The Journal of Biological Chemistry* **286**, 34382–34390, 2011.

309. C. Scheufler *et al.*, Structure of TPR Domain–Peptide Complexes: Critical Elements in the Assembly of the Hsp70–Hsp90 Multichaperone Machine. *Cell* **101**, 199–210, 2000.

310. S. K. Wandinger, K. Richter, J. Buchner, The Hsp90 chaperone machinery. *The Journal of Biological Chemistry* **283**, 18473–18477, 2008.

311. F. U. Hartl, A. Bracher, M. Hayer-Hartl, Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324–332, 2011.

312. P. Selenko *et al.*, In situ observation of protein phosphorylation by high-resolution NMR spectroscopy. *Nature Structural & Molecular Biology* **15**, 321–329, 2008.

313. D. Dormann *et al.*, ALS-associated fused in sarcoma (FUS) mutations disrupt Transportin-mediated nuclear import. *The EMBO Journal* **29**, 2841–2857, 2010.

314. S. Alberti, A. Gladfelter, T. Mittag, Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **176**, 419–434, 2019.

315. T. M. Franzmann, S. Alberti, Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior. *Journal of Biological Chemistry* **294**, 7128–7136, 2019.

316. M. Feric *et al.*, Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* **165**, 1686–1697, 2016.

317. D. Dormann, FG-nucleoporins caught in the act of liquid–liquid phase separation. *Journal of Cell Biology* **219**, 2019.

318. J. Nikolic *et al.*, Negri bodies are viral factories with properties of liquid organelles. *Nature Communications* **8**, 58, 2017.

319. S. Guseva *et al.*, Measles virus nucleo- and phosphoproteins form liquid-like phase-separated compartments that promote nucleocapsid assembly. *Science Advances* **6**, eaaz7095, 2020.

320. L. Pytowski *et al.*, Liquid–liquid phase separation of type II diabetes-associated IAPP initiates hydrogelation and aggregation. *Proceedings of the National Academy of Sciences* **117**, 12050–12061, 2020.

321. S. Rauscher, R. Pomès, The liquid structure of elastin. *eLife* **6**, ed. by G. Hummer, e26526, 2017.

322. C. Belott, B. Janis, M. A. Menze, Liquid-liquid phase separation promotes animal desiccation tolerance. *Proceedings of the National Academy of Sciences* **117**, 27676–27684, 2020.

323. P. E. A. Ash *et al.*, TIA1 potentiates tau phase separation and promotes generation of toxic oligomeric tau. *Proceedings of the National Academy of Sciences* **118**, 2021.

324. A. E. Conicella *et al.*, TDP-43 a-helical structure tunes liquid–liquid phase separation and function. *Proceedings of the National Academy of Sciences* **117**, 5883–5894, 2020.

325. S. E. Reichheld *et al.*, Direct observation of structure and dynamics during phase separation of an elastomeric protein. *Proceedings of the National Academy of Sciences* **114**, E4408–E4415, 2017.

326. T. C. Boothby *et al.*, Tardigrades Use Intrinsically Disordered Proteins to Survive Desiccation. *Molecular Cell* **65**, 975–984, 2017.

327. A. Siegert *et al.*, Interplay between tau and a-synuclein liquid-liquid phase separation. *Protein Science* **30**, 1326–1336, 2021.

328. S. Wegmann *et al.*, Tau protein liquid-liquid phase separation can initiate tau aggregation. *The EMBO journal* **37**, e98049, 2018.

329. J. P. Brady *et al.*, Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation. *Proceedings of the National Academy of Sciences* **114**, E8194–E8203, 2017.

330. T. Yuwen, J. P. Brady, L. E. Kay, Probing Conformational Exchange in Weakly Interacting, Slowly Exchanging Protein Systems via Off-Resonance R1rh Experiments: Application to Studies of Protein Phase Separation. *Journal of the American Chemical Society* **140**, 2115–2126, 2018.

331. S. F. Banani *et al.*, Biomolecular condensates: organizers of cellular biochemistry. *Nature reviews Molecular cell biology* **18**, 285–298, 2017.

332. P. D. Majors, J. S. McLean, J. C. M. Scholten, NMR bioreactor development for live in-situ microbial functional analysis. *Journal of Magnetic Resonance* **192**, 159–166, 2008.

333. N. G. Sharaf *et al.*, A bioreactor for in-cell protein NMR. *Journal of Magnetic Resonance* **202**, 140–146, 2010.

334. D. A. Foley *et al.*, NMR Flow Tube for Online NMR Reaction Monitoring. *Analytical Chemistry* **86**, 12008–12013, 2014.

335. S. Milles *et al.*, Self-Assembly of Measles Virus Nucleocapsid-like Particles: Kinetics and RNA Sequence Dependence. *Angewandte Chemie International Edition* **55**, 9356–9360, 2016.

336. C. J. Moreau *et al.*, in *Methods in Enzymology* (Elsevier, 2015), vol. 556, pp. 425–454.

337. J. V. Sehy, J. J. H. Ackerman, J. J. Neil, Water and lipid MRI of the Xenopus oocyte. *Magnetic Resonance in Medicine* **46**, 900–906, 2001.

338. L. Yao *et al.*, Site-Specific Backbone Amide 15N Chemical Shift Anisotropy Tensors in a Small Protein from Liquid Crystal and Cross-Correlated Relaxation Measurements. *Journal of the American Chemical Society* **132**, 4295–4309, 2010.

339. J. R. Brender, D. M. Taylor, A. Ramamoorthy, Orientation of amide-nitrogen-15 chemical shift tensors in peptides: a quantum chemical study. *Journal of the American Chemical Society* **123**, 914–922, 2001.

340. C. D. Kroenke, M. Rance, A. G. Palmer, Variability of the 15N Chemical Shift Anisotropy in Escherichia coli Ribonuclease H in Solution. *Journal of the American Chemical Society* **121**, 10119–10125, 1999.

341. C. Scheurer *et al.*, Effects of Dynamics and Environment on ^{15} N Chemical Shielding Anisotropy in Proteins. A Combination of Density Functional Theory, Molecular Dynamics Simulation, and NMR Relaxation. *Journal of the American Chemical Society* **121**, 4242–4251, 1999.

342. J. B. Hall, D. Fushman, Variability of the 15N Chemical Shielding Tensors in the B3 Domain of Protein G from 15N Relaxation Measurements at Several Fields. *Journal of the American Chemical Society* **128**, 7855–7870, 2006.

343. Q. Johnson *et al.*, Water's Contribution to the Energetic Roughness from Peptide Dynamics. *Journal of Chemical Theory and Computation* **6**, 2591–2597, 2010.

344. D. Nerukh, S. Karabasov, Water–Peptide Dynamics during Conformational Transitions. *The Journal of Physical Chemistry Letters* **4**, 815–819, 2013.

345. Z. Chang, B. Halle, Nuclear magnetic relaxation by the dipolar EMOR mechanism: General theory with applications to two-spin systems. *The Journal of Chemical Physics* **144**, 084202, 2016.

346. W. Adamski *et al.*, A Unified Description of Intrinsically Disordered Protein Dynamics under Physiological Conditions Using NMR Spectroscopy. *Journal of the American Chemical Society* **141**, 17817–17829, 2019.

347. M. Xie *et al.*, Functional protein dynamics on uncharted time scales detected by nanoparticle-assisted NMR spin relaxation. *Science Advances* **5**, eaax5560, 2019.

348. S. Wardenfelt *et al.*, Broadband Dynamics of Ubiquitin by Anionic and Cationic Nanoparticle Assisted NMR Spin Relaxation. *Angewandte Chemie International Edition* **60**, 148–152, 2021.

349. M. Xie *et al.*, Quantitative Binding Behavior of Intrinsically Disordered Proteins to Nanoparticle Surfaces at Individual Residue Level. *Chemistry – A European Journal* **24**, 16997–17001, 2018.

350. M. Xie *et al.*, Residue-Specific Interactions of an Intrinsically Disordered Protein with Silica Nanoparticles and Their Quantitative Prediction. *The Journal of Physical Chemistry C* **120**, 24463–24468, 2016.

351. D.-W. Li, M. Xie, R. Brüschweiler, Quantitative Cooperative Binding Model for Intrinsically Disordered Proteins Interacting with Nanomaterials. *Journal of the American Chemical Society* **142**, 10730–10738, 2020.

352. S. F. Cousin *et al.*, High-resolution two-field nuclear magnetic resonance spectroscopy. *Physical Chemistry Chemical Physics* **18**, 33187–33194, 2016.

353. N. Bolik-Coulon *et al.*, Theoretical and computational framework for the analysis of the relaxation properties of arbitrary spin systems. Application to high-resolution relaxometry. *Journal of Magnetic Resonance* **313**, 106718, 2020.

354. N. Bolik-Coulon *et al.*, Two-field transverse relaxation-optimized spectroscopy for the study of large biomolecules – An in silico investigation. *Journal of Magnetic Resonance Open* **4-5**, 100007, 2020.

355. P. Kadeřávek *et al.*, Protein Dynamics from Accurate Low-Field Site-Specific Longitudinal and Transverse Nuclear Spin Relaxation. *Journal of Physical Chemistry Letters* **10**, 5917–5922, 2019.

356. K. M. Ruff, R. V. Pappu, A. S. Holehouse, Conformational preferences and phase behavior of intrinsically disordered low complexity sequences: insights from multiscale simulations. *Current Opinion in Structural Biology* **56**, 1–10, 2019.

357. G. L. Dignon *et al.*, Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS computational biology* **14**, e1005941, 2018.

358. Y. R. Li *et al.*, Stress granules as crucibles of ALS pathogenesis. *Journal of Cell Biology* **201**, 361–372, 2013.

359. T. H. Kim *et al.*, Phospho-dependent phase separation of FMRP and CAPRIN1 recapitulates regulation of translation and deadenylation. *Science* **365**, 825–829, 2019.

360. S. Banjade *et al.*, Conserved interdomain linker promotes phase separation of the multivalent adaptor protein Nck. *Proceedings of the National Academy of Sciences* **112**, E6426–E6435, 2015.

361. B. Tsang *et al.*, Phosphoregulated FMRP phase separation models activity-dependent translation through bidirectional control of mRNA granule formation. *Proceedings of the National Academy of Sciences* **116**, 4218–4227, 2019.

362. Z. Monahan *et al.*, Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *The EMBO journal* **36**, 2951–2967, 2017.

363. M. Hofweber *et al.*, Phase Separation of FUS Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation. *Cell* **173**, 706–719, 2018.

364. T. J. Nott *et al.*, Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Molecular Cell* **57**, 936–947, 2015.

365. A. A. M. André, E. Spruijt, Liquid–Liquid Phase Separation in Crowded Environments. *International Journal of Molecular Sciences* **21**, 5908, 2020.

366. G. C. Carter *et al.*, N-terminal Domain of TDP43 Enhances Liquid-Liquid Phase Separation of Globular Proteins. *Journal of Molecular Biology* **433**, 166948, 2021.

367. L. E. Wong *et al.*, Tripartite phase separation of two signal effectors with vesicles priming B cell responsiveness. *Nature Communications* **11**, 848, 2020.