HHS Public Access

Author manuscript

*Arch Biochem Biophys*. Author manuscript; available in PMC 2018 August 15.

Published in final edited form as:

*Arch Biochem Biophys*. 2017 August 15; 628: 71–80. doi:10.1016/j.abb.2017.05.002.

# Applications of NMR and computational methodologies to study protein dynamics

**Chitra Narayanan**[1], **Khushboo Bafna**[2], **Louise D. Roux**[1], **Pratul K. Agarwal**[3,4], and **Nicolas Doucet**[1,5,6,*]

[1]INRS-Institut Armand-Frappier, Université du Québec, 531 Boul. des Prairies, Laval, QC H7V 1B7, Canada

[2]Genome Science and Technology, University of Tennessee, Knoxville, TN 37996, USA

[3]Department of Biochemistry, Cellular and Molecular Biology, University of Tennessee, Knoxville, TN 37996, USA

[4]Computational Biology Institute and Computer Science and Mathematics Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN 37830, USA

[5]PROTEO, the Quebec Network for Research on Protein Function, Structure, and Engineering, 1045 Avenue de la Médecine, Université Laval, Québec, QC G1V 0A6, Canada

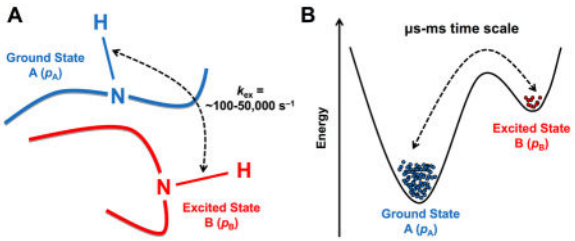
[6]GRASP, the Groupe de Recherche Axé sur la Structure des Protéines, 3649 Promenade Sir William Osler, McGill University, Montréal, QC H3G 0B1, Canada

## Abstract

Overwhelming evidence now illustrates the defining role of atomic-scale protein flexibility in biological events such as allostery, cell signaling, and enzyme catalysis. Over the years, spin relaxation nuclear magnetic resonance (NMR) has provided significant insights on the structural motions occurring on multiple time frames over the course of a protein life span. The present review article aims to illustrate to the broader community how this technique continues to shape many areas of protein science and engineering, in addition to being an indispensable tool for studying atomic-scale motions and functional characterization. Continuing developments in underlying NMR technology alongside software and hardware developments for complementary computational approaches now enable methodologies to routinely provide spatial directionality and structural representations traditionally harder to achieve solely using NMR spectroscopy. In addition to its well-established role in structural elucidation, we present recent examples that illustrate the combined power of selective isotope labeling, relaxation dispersion experiments, chemical shift analyses, and computational approaches for the characterization of conformational sub-states in proteins and enzymes.

## Graphical Abstract

[*]Contact information: nicolas.doucet@iaf.inrs.ca; Phone: (450) 687-5010, ext. 4212.




**Keywords**

Relaxation dispersion; quasi anharmonic analysis; chemical shift analysis; protein dynamics; allostery; conformational sub-states

## 1. Introduction

Proteins are not rigid entities but flexible assemblies. Despite the broad range of growth and living conditions inside eukaryotic and prokaryotic cells, a number of studies have efficiently demonstrated that changes in molecular flexibility of proteins and enzymes on various time scales results in significant loss of biological function and/or phenotypic defects [1–4]. Each protein exists as an ensemble of conformations that undergo continuous exchange within several spatial and temporal scales, relying on these motions to achieve their biological function. To fully grasp, predict, and eventually engineer efficient protein behavior at the molecular level, it is therefore crucial to understand the effects of conformational exchange and atomic-scale protein flexibility on biological function. To this day, nuclear magnetic resonance (NMR) spectroscopy represents one of the most powerful techniques available to unravel and investigate such dynamic protein behaviors on several time frames (Table 1). Technological developments and continuous improvements in processing power have recently allowed computational approaches to significantly complement NMR methodologies, providing spatial directionality and structural representations not easily or rapidly achievable by NMR. The present review article showcases several recent studies and methodologies characterizing protein and enzyme dynamics, primarily illustrating advances in NMR relaxation and computational approaches, coupled to other biophysical approaches. Recent examples demonstrating the combined power of selective isotope labeling, relaxation dispersion experiments, chemical shift analyses, and computational approaches are presented.

Excellent compendia have recently covered the theoretical aspects of NMR methodologies available for quantifying protein dynamics on a broad range of time frames, *i.e.* from picoseconds to hours [5–17]. To prevent undue repetition of the same theoretical concepts, the present report neither aims to offer a comprehensive overview of these NMR methodologies, nor to fully cover the wide array of protein systems exemplifying the importance of conformational flexibility in protein function. Rather, the goal is to illustrate to a general audience how the combination of newer developments with now established methods such as the NMR nuclear spin and relaxation dispersion experiments have become indispensable tools for providing a uniquely thorough understanding of the historically elusive protein dynamics and allosteric processes occurring in many biological systems. Non-technical

readers should nevertheless be informed that most recent advancements have seen rigorous targeting of distinct atomic probes in protein backbone and residue side chains (HN, Hα, $NH_2$, Cα, C′, $CH_3$, etc.), in addition to effectively overcoming the classical size limitation of protein complexes amenable to NMR investigation [18,19]. Protein relaxation still requires dedicated access to high-field NMR spectrometers (typically between 500 MHz and 1GHz $^1H$ frequencies) and for extended periods of time (hours/days). This is due to the relatively low signal sensitivity of the method, especially in light of long three-dimensional heteronuclear experiments. However, significant advances over the past 10–15 years have lifted many classical limitations of the NMR technique, effectively providing a more complete understanding of the time-dependent properties of exchange processes occurring during protein function and allosteric communication.

## 2. Selective isotope labeling: overcoming protein size limitation in relaxation

For a long time, protein NMR applications were limited by the size of the molecular complexes under investigation. This was primarily attributed to a number of experimental factors pertaining to physical limitations of the technique and isotope labeling methodologies [20]. One of the most important limitations in protein NMR is the fact that higher molecular weight protein monomers or multimeric complexes tumble (or rotate) more slowly in solution than their smaller counterparts. This affects relaxation parameters and results in spectral broadening (or dampening) of NMR signals. Effectively, if a high molecular weight protein is uniformly labeled with $^{13}C/^{15}N$ isotopes throughout its primary structure, the relevant spectral NMR resonances are likely to be overlapped with other resonances, and therefore masked (or completely invisible) to the experimentalist. The uniform isotope labeling methodology required to achieve NMR signal presents another limitation for NMR of large proteins. Signal resolution for any multidimensional NMR experiment on a uniformly labeled large protein or protein complex results in a significant percentage of resonance overlaps, preventing detailed atomic-scale analyses for many of the overlapping residues.

Recent years have seen significant improvements in selective isotope protein labeling methodologies, effectively eliminating some of the disadvantages of uniform labeling [20]. A convincing example of selective labeling coupled to recent protein dynamics developments was illustrated by the Kay group on high molecular weight systems, particularly the 20S core particle (CP) proteasome [21,22]. Focusing on ‘metabolic tricks’ limiting isotope labeling to the methyl groups of Ile, Leu, Met, and Val residues, the authors have efficiently characterized functionally relevant motions in this megadalton complex. The 20S core particle (CP) proteasome is a 670-kDa complex formed by a combination of 2 alpha and 2 beta ring barrel structures that build a molecular machine involved in degrading cellular proteins that have become damaged over time. Due to massive size limitations, standard uniform isotope labeling prevents clear atomic-scale investigation. Using unlabeled (therefore invisible to NMR) beta protomers and highly deuterated, selectively labeled $^{13}CH_3$-Methionine residues in the alpha protomers (5 of which are found in the N-terminal pore entrance), the authors elegantly characterized the functional gating mechanism

of this macromolecular complex by combining mutagenesis with paramagnetic resonance enhancement (PRE). The authors combined their NMR analyses with spin labels inside and outside the proteasome lumen, taking advantage of the PRE effect to study the conformational changes in the molecular gating mechanism [21,22].

This example illustrates the power of combining selective protonation with mutagenesis, NMR relaxation, and PRE effects to investigate the dynamical behavior of high molecular weight protein complexes. Originally elusive, well-defined selective labeling protocols are now readily available to any skilled protein expression experimentalist [8,19], allowing one to significantly improve the investigation of conformational exchange for large protein complexes using NMR relaxation dispersion (see below).

## 3. Investigating the microsecond to millisecond time frame to probe biologically relevant protein sub-states involved in allostery and enzyme function

Arguably one of the most influential and now routinely employed methodologies of the past 15 years in solution-state NMR is the investigation of free and ligand-bound protein complexes using relaxation dispersion. Relaxation dispersion NMR allows the characterization of low populated, 'invisible' sub-states sampled by proteins in solution, either in their free states or as they perform their catalytic function in the presence of binding partners. For instance, the Carr-Purcell-Meiboom-Gill (CPMG) and $R_{1\rho}$ rotating frame experiments have been collectively employed to investigate conformational exchange rates ($k_{ex}$) occurring in proteins over time frames that roughly span ~100 to ~50,000 events per second ($s^{-1}$) [9,23] (Figure 1). In addition to their relatively simple usage and implementation on multiple atomic probes (most of which are now available in standard and TROSY versions) [19,24], relaxation dispersion experiments are especially popular in light of the overlapping time scale of the characterized rates of motion with the slower microsecond-millisecond (μs-ms) time frame of relevant biological processes such as protein folding, enzyme catalysis, ligand binding, etc. Recent advances have emerged to improve the breadth of $^{15}$N-labeled protein investigation by complementing the fast time scale (fast picosecond-nanosecond [ps-ns]) $R_1$, $R_2$, and heteronuclear NOE (hNOE) nuclear spin relaxation with CPMG and/or $R_{1\rho}$. This was done by selectively targeting $^{13}$C, $^{1}$H and/or $^{2}$H dynamic spin probes and by improving spectral sensitivity [25–27].

More recent methodologies, such as the NMR chemical exchange saturation transfer (CEST) experiment [6,28,29], probe the same ms time frame and have provided improved sensitivity, in addition to allowing subtler conformational changes and better characterization of slower exchange processes (*e.g.* $k_{ex}$ = 50–150 $s^{-1}$). The CEST experiment can also simultaneously take advantage of the CEST-derived $R_1$ and $R_2$ data to extract $S^2$ order parameters reporting on the ps-ns time scale [30]. In spin-relaxation NMR, the order parameter $S^2$ is typically extracted from the model-free approach using experimental $R_1$, $R_2$, and hNOE NMR data [31–33]. This parameter is used to characterize the amplitude of the internal motions of a $^{1}$H-$^{15}$N bond vector on the ps-ns time scale. $S^2$ measures the degree of spatial restriction of each backbone $^{1}$H-$^{15}$N vector, describing an array of atomic-site motions ranging from fully

unrestricted ($S^2 = 0$) to completely restricted ($S^2 = 1$) on this particular time frame. The CEST approach showed remarkably good agreement with standard $R_1$, $R_2$, and hNOE nuclear spin relaxation data, in addition to considerably improving NMR acquisition time (hours instead of days). The CPMG method can further be extended at the other end of the time spectrum, allowing the investigation of systems with exchange rates as high as 6,000 $s^{-1}$ [34]. The combined analysis of both amide $^{15}N$ and $^{1}H^{N}$ CPMG profiles and major state exchange-induced $^{15}N$ chemical shifts measured in $^{1}H^{N}$-$^{15}N$ heteronuclear multiple-quantum coherence (HMQC) and heteronuclear single-quantum coherence (HSQC) data sets allows the accurate extraction of exchange parameters and chemical shift difference between the interconversion states of systems exchanging with rates of ~6,000 $s^{-1}$.

In addition to the biological importance of the μs-ms time frame, and relative ease of technical implementation, recent examples of relaxation dispersion NMR have been employed with selective labeling and other biophysical methodologies to yield a complete portrait of protein allostery in various systems. A comprehensive enzymatic characterization also typically takes advantage of other biophysical techniques such as circular dichroism spectroscopy (CD) and/or isothermal titration calorimetry (ITC) [35,36]. A study on the allosteric mechanism of the heterodimeric 51-kDa glutamine amidotransferase imidazole glycerol phosphate synthase (IGPS) recently exemplified a subtle yet functionally relevant behavior observed through long-range allosteric motional transmission upon small-molecule effector binding to this enzyme involved in purine and histidine biosynthetic pathways in bacteria [37,38]. In this work, the authors combined ITC with $^{1}H$,$^{13}C$-TROSY-HMQC and methyl-TROSY multiple-quantum CPMG dispersion — all solely permitted because of selective isotope labeling of Ile, Leu, and Val methyl groups — to illustrate that the IGPS complexes with allosteric effectors distinctively respond to long-range (>15 Å) millisecond dynamics. Binding of allosteric effectors, monitored using ITC, revealed favorable enthalpic and entropic contributions to the free energy change, suggesting a shallow binding pocket for the effectors, consistent with docking and MD simulations. Interestingly, the authors demonstrate a direct linear correlation between the number of flexible residues on the μs-ms time scale and the catalytic efficiency ($k_{cat}/K_m$) of the glutaminase activity, further confirming the dynamic behavior of amino acid residues that were previously shown to be critical for relaying allosteric information from independent biochemical studies [37,38].

These results highlight how protein-ligand complexes between IGPS and allosteric effectors with varying modulating efficiencies result in significantly distinct dynamic residue networks on the catalytic time frame of the enzyme. Further strengthening these observations, the authors observe that only the catalytically productive ternary complex between IGPS (the proper allosteric effector) and the PRFAR substrate yields a dynamic pathway with a shared, global exchange-rate constant ($k_{ex}$) for nearly all residues undergoing millisecond exchange in the enzyme. Although other IGPS allosteric activators also trigger millisecond motions, the exchange rates of individual dynamic residues are not concerted nor localized at selective residue positions essential for optimal allosteric communication. In addition to perturbing the unique profile of IGPS conformation, these suboptimal effectors also show concomitant gaps in their abilities to enhance glutamine hydrolysis [37,38].

This work further illustrates the power of combining biophysical, computational, and new NMR methodologies to improve our understanding of allostery, offering the means to observe 'invisible' conformational sub-states that would have remained experimentally elusive just a few years ago. Performing these experiments using classical non-TROSY $^{1}$H,$^{15}$N-HSQC titrations, coupled with standard backbone $^{1}$H,$^{15}$N-CPMG relaxation would have yielded overly crowded spectra without significant resolution to observe the subtle molecular and motional effects described in these studies. Also, reports have illustrated the lack of significant structural changes caused by effector binding to IGPS (root-mean-square deviation of 0.41 Å between the apo and holo forms) [37]. As a result, previous crystal structures [39] and fluorescence quenching experiments examining solvent accessibility [40] were unable to provide definitive information on the intricate atomic-scale details underlying allosteric transmission in this enzyme system, further illustrating the power of the NMR methodology combination described here.

## 4. Uncovering potential allosteric sites using the NMR chemical shift covariance (CHESCA) and projection (CHESPA) analyses

Recent years have also seen the development of a variety of approaches that investigate NMR chemical shift perturbations to identify potential allosteric networks and structural dynamics in proteins [41–46]. The Melacini group introduced the chemical shift covariance analysis (CHESCA) approach [44] to identify dynamically driven intramolecular amino acid networks. CHESCA uses a combination of agglomerative clustering and singular value decomposition to identify amino acid networks that show correlated changes in chemical shifts due to perturbations arising from ligand binding or mutations.

The CHESCA approach involves the following steps: first, a matrix **M** of combined $^{1}$H and $^{15}$N chemical shifts of a protein in the apo and multiple perturbed states is used to calculate inter-residue Pearson's correlation coefficients. The resulting symmetric matrix **R** of pairwise correlation coefficients is analyzed using the agglomerative clustering (AC) method [47] to identify cluster(s) of coupled amino acid residues. AC assigns the first intracluster link for a residue pair with the highest correlation coefficient ($r_{ij}$). Subsequent intracluster links are assigned between the first pair and nearest neighbor based on the highest value of correlation coefficient to either of the residues in the original pair. These intracluster links, which can be visualized using dendrograms, identify cluster(s) of residues that show concerted response to perturbations. Clusters with more than three residues are grouped into 'networks'. At this point, however, it is unclear whether these networks play a role in allostery.

To assign a function to the identified network(s), a correlation coefficient matrix $\mathbf{R_I}$ is computed using combined $^{1}$H and $^{15}$N chemical shifts for a subset of amino acid residues ($\mathbf{M_I}$) of the network(s) identified using AC, as described above. $\mathbf{R_I}$ is analyzed again using AC to cluster functional states (such as active *vs.* inactive states) based on residues rather than grouping residues based on functional states. Selvaratnam *et al.* applied this approach to the cAMP binding domain of EPAC using chemical shifts of the apo and activator (analog) bound states.

The authors identified two networks of amino acids, one separating active and inactive states of the protein, and another separating the bound and apo states. These observations indicate the role of these two networks in allostery. They further used singular value decomposition (SVD) [44,48] to validate the functional networks identified through hierarchical clustering of AC. By using combined chemical shift differences relative to an antagonist-bound form of EPAC as reference, the authors identified residue networks involved in ligand binding and allostery, respectively.

The Melacini group also introduced the chemical shift projection analysis (CHESPA) to quantify the size and direction of perturbations induced by ligand binding or point mutations (Figure 2) [43]. In CHESPA, two residue-specific descriptors of this perturbation, *i.e.* the direction ($cos\theta$) and magnitude (fractional shift $X$), are calculated based on the chemical shift changes observed in the perturbed system. The fractional shift $X$ is calculated as the ratio of vector A along vector B and the magnitude of vector B:

$$X = \frac{\vec{A} \cdot \vec{B}}{\left|\vec{B}\right|^2}$$

The fractional shift $X$ is a scalar quantity and is complemented by $cos\theta$, which reports on the relative orientation of vectors A and B, $\theta$ being the angle between vectors A and B:

$$\cos\theta = \frac{\vec{A} \cdot \vec{B}}{\left|\vec{A}\right|\left|\vec{B}\right|}$$

Of note is the fact that this approach is valid when the exchange in the free and ligand-bound conformations is fast on the NMR time scale. The presence of two or more states in the fast exchange regime could compromise the linear chemical shift pattern exploited by the CHESPA method, resulting in $\cos\theta$ values below unity. Thus, $\cos\theta$ values approaching unity suitably report on two-state equilibrium through the fractional shift $X$.

Much like CHESCA, the CHESPA approach was first applied on the EPAC model system by comparing the chemical shift changes arising from mutations and ligand binding to identify amino acid residues that (de)stabilize protein conformations associated with ligand binding [43]. In this study, the fractional shift $X$ is positive if the mutation shifts the equilibrium towards the active state, and negative if the displacement is in the opposite direction. The value of $X$ approaches 0 if variations caused by the mutation are negligible relative to those caused by ligand binding and/or if vectors A and B are orthogonal. Finally, $X \approx 1$ indicates that the magnitude of the chemical shift is identical for mutation and ligand binding events. The direction of movement is quantified using $\cos\theta$ (Figure 2). Residues with $\cos\theta \approx 1$ show displacements in the same direction for both apo-mutant and ligand-bound states, whereas residues significantly affected by mutation display $\cos\theta < 1$. In this example, the L237W mutant of EPAC was shown to stabilize the inactive state, while cAMP binding activated wild type (WT) EPAC. The chemical shift difference between apo-WT and apo-mutant induced by amino acid substitution was calculated as the magnitude of vector A

connecting apo-WT and apo-mutant peaks in the plane of $^{1}H$ and scaled $^{15}N$ spectra (Figure 2). Similarly, the chemical shift difference induced upon ligand binding was computed as the magnitude of vector B connecting apo-WT and cAMP-bound WT peaks. The projection of vector A onto vector B is a measure of the shift of the conformational equilibrium between protein conformations found in free and bound states. In a recent study, Selvaratnam *et al.* combined the CHESPA and CHESCA approaches to characterize the effects of mutations on the inactive *vs.* active conformational equilibria of apo EPAC [49]. The CHESPA analysis was used to remove the false positives in the allosteric clusters obtained through the hierarchical clustering of residue pairs in CHESCA.

The CHESPA approach can also be used to investigate the effect of long-range residue network behavior occurring upon distinct ligands binding to an enzyme, as recently illustrated by the binding of 3′-UMP and 5′-AMP to human angiogenin (Figure 3) [50]. In this study, the authors highlighted the existence of two distinct, long-range clusters of residue networks either displaying correlated or uncorrelated chemical shift trajectories upon binding of each ligand to the purine- or pyrimidine-specific subsites of the angiogenin catalytic cleft, suggesting a structural rearrangement or allosteric response linked to ligand binding in this enzyme. Finally, using alanine to glycine substitutions, Axe *et al.* probed structural and dynamical changes in the α-subunit of tryptophan synthase [51]. By combining the projection (CHESPA) and covariance (CHESCA) analyses for the free and ligand-bound forms of the wild-type and mutant enzymes, they identified point mutations that stabilized three conformational states, corresponding to the free form, glyceraldehyde-3-phosphate (product) bound and active (indole-bound) states. The authors further illustrated that the dynamical properties of the allosteric networks of amino acid residues identified from the covariance analysis differed in each of these states.

## 5. Enhancing the structural and time-evolution throughput of NMR spectroscopy using computational and theoretical approaches

### 5.1. Combining NMR observables and MD simulations to characterize protein motion

Molecular dynamics (MD) simulations provide a time-evolution ‘movie’ of atomic-level motions based on Newton’s laws of motions. Classical molecular mechanics describe the potential energy of a molecule, such as a protein, as a function of its atomic coordinates and dictate the interactions between atoms over the course of an MD simulation [52]. The various force-constants associated with inter-atomic interactions are collectively referred to as a force-field. MD trajectories, which correspond to a collection of conformational snapshots sampled during the simulation, are analyzed using a variety of approaches to identify conformational states that are potentially important for function [53]. While conformational motions in proteins are observed over a wide range of time scales — from the fast femtosecond-picosecond (fs-ps) to the slower millisecond-second (ms-s) — the range of time scales accessible by MD is currently limited due to the speed of central/graphical processing units and related computer hardware. Until recently, sampling from MD simulations was typically on the order of hundreds of nanoseconds. Enhanced sampling techniques, alongside advances in computational resources, now enable routine conformational sampling on the μs time scale, with specialized supercomputing facilities

reporting longer ms time scale simulations [54,55]. MD simulations have been used in combination with NMR to identify allosteric networks of conformational motions that play an important role in enzyme function [10,56,57].

A diverse array of other sequence and structure based computational approaches have also been used to identify and characterize protein flexibility on the fast time scale [58,59], in addition to characterizing allosteric networks that display motions on the slower μs-ms time frame *via* dynamical coupling of distal residues in proteins [60–62]. For example, DynaMine is a sequence-based approach that provides a statistical and quantitative analysis of protein backbone flexibility on the fast time scale (ps-ns) for each amino acid of the protein [59]. Statistical coupling analysis is another approach that uses sequence datasets of protein families to identify co-evolving amino acid networks important for distinct biochemical functions [63,64]. Using this approach, Reynolds *et al.* identified a co-evolving amino acid network in the dihydrofolate reductase family undergoing millisecond conformational dynamics that was strongly correlated with enzyme catalysis [61].

Accurate prediction of NMR observables such as chemical shifts, residual dipolar couplings (RDCs), and order parameters ($S^2$) from protein coordinates has facilitated the characterization of protein dynamics through the comparison between NMR and molecular simulation ensembles [65]. Algorithms such as SHIFTX2, SHIFTS, CamShift, SPARTA+, and others have been used to predict chemical shifts from structural coordinates obtained from MD simulations [66–69]. Chemical shifts have also been used to compute order parameters ($S^2$) for MD ensemble coordinates. Comparison of the predicted $S^2$ with experimental observations facilitates the quantitative characterization of fast time scale dynamics of the protein backbone, providing a range of conformations associated with the observed order parameters [65,70,71]. Further, MD can complement structural observations that are missing from NMR relaxation [72,73].

In recent years, a number of studies have combined MD and NMR chemical shift analyses to probe conformational dynamics in proteins. For example, Robustelli *et al.* compared dynamically averaged chemical shifts predicted from MD simulation ensembles with experimental NMR shifts to characterize the effect of conformational dynamics on the observed chemical shift pattern in proteins. Using 100 ns and 1 μs simulations of two homologs of ribonuclease H (RNase H), they showed that ensemble averaged chemical shifts calculated from MD are in better agreement with NMR observables than those predicted from static X-ray structures [42]. By probing the dynamical properties of regions displaying improved chemical shift agreement with experiment, the authors showed that these improvements result from: 1) a population weighted sampling of multiple conformational states, and 2) sampling within individual conformational basins. Further, analysis of averaged chemical shifts from MD simulations also identified erroneous conformations resulting from inadequacies of the force-field. These observations highlight the use of dynamically averaged chemical shifts to probe conformational dynamics in proteins.

NMR chemical shifts have also been used as replica-averaged structural restraints for MD simulations to characterize the dynamics of ribonuclease A (RNase A) [31–33]. Camilloni *et al.*

showed that agreement between calculated and experimental chemical shifts was improved in the chemical shift restrained ensembles relative to unrestrained MD ensembles. Comparison of the free energy landscape of RNase A for the unrestrained and restrained MD ensembles showed that the restrained MD simulations sample the major and minor conformational states reported for RNase A, with the calculated population of the minor state consistent with previous experimental reports. These observations highlight the importance of using chemical shifts to bias the conformational landscape to reproduce dynamical properties of a protein in molecular dynamics simulations.

A number of NMR observables other than chemical shifts have also been used as restraints for MD simulations, either alone or by combining several observables [11,75]. Several studies have compared NMR residual dipolar couplings (RDCs) with predicted RDCs from MD ensembles to characterize backbone and side-chain dynamics over a range of time scales [12,13,76–79]. Markwick *et al.* combined plain and accelerated MD (AMD) simulations with experimental RDCs to characterize conformational dynamics of the GB3 protein over the faster ns and slower ms time scales [78]. In a subsequent study, the authors combined a novel AMD/SVD approach with extensive NMR observables to determine the time-resolved structural dynamics of ubiquitin over a broad time scale (ps-ms) [80]. SVD allowed the determination of optimal scalar *J*-coupling and RDC alignment tensor parameters. The population weighted average scalar couplings were shown to be in good agreement with experimental observations.

Computational MD simulations and experimentally extracted NMR relaxation parameters are now being routinely used as complementary approaches to characterize protein dynamics on multiple time scales. Among recent examples, this combined method was used to uncover amino acid residues associated with allosteric communication in Pin1 and IGPS (as reviewed above and in ref. [14]). Cross-correlation analyses of protein motions from MD trajectories have highlighted changes in correlated motions upon binding of effector molecules to IGPS [38,81]. NMR and MD were also recently integrated to characterize the dynamical properties of the *Streptomyces lividans* xylanase B2 (XlnB2) in the free and ligand-bound states [82]. Ligand binding was shown to induce enhanced conformational dynamics of residues that interact with the ligand in the thumb loop and finger regions of the enzyme. Finally, by comparing order parameters from MD simulations with NMR observables, Fisette *et al.* showed that better agreement is observed with structured elements (α-helices and β-sheets) while sampling of loop regions by MD were insufficient to reproduce the experimental observations for these regions [83]. In a subsequent study, the authors characterized the dynamical properties of two β-lactamases upon substrate binding [84] (also reviewed in ref. [65]). By comparing order parameters obtained from MD and NMR for the free form, they showed excellent agreement for ordered regions while differences were observed for the less ordered loop regions. For the substrate-bound state, back-calculated order parameters from MD showed localized differences upon binding near catalytic residues.

### 5.2. Conformational fluctuations and sub-states characterization using the Quasi Anharmonic Analysis (QAA)

NMR techniques such as CPMG can provide the rate of exchange between distinct conformations constituting discrete conformational *sub-states*, enabled by concerted motions of residue networks that occur on the μs-ms time scale. For instance, many residues in a loop can experience similar individual exchange rates ($k_{ex}$), suggesting a concerted group (or global) motion on this particular time frame. Rates of conformational exchange of surface loop regions have been shown to coincide with the catalytic turnover in several enzyme systems [85–87]. This has led to a significant interest in linking events associated with these dynamic exchange rates with particular molecular mechanisms involved in protein function. Unfortunately, while possible, NMR is somewhat limited in its ability to provide rapid atomistic details on directionality and length-scales for these movements. Therefore, integration of computational and NMR techniques is becoming increasingly popular for the characterization of conformational substates and their populations. In combination with experimental NMR relaxation data, MD simulation techniques can be used to build a hierarchy of conformational states. With statistical sampling, MD can provide information on populations associated with each state and can also be extended to calculate the kinetic information about the interconversion between the states (Figure 1B). The rates and magnitudes of interconversion are correlated with the CPMG-based relaxation parameters. Once validated, the computationally obtained conformational sub-states can provide detailed information about the orientation and other structural (and dynamical) properties of individual atoms, residues, and whole proteins.

A novel computational approach, the Quasi Anharmonic Analysis (QAA), was developed to identify sub-states in the conformational landscape of proteins [88,89]. QAA uses higher order statistics of positional variations associated with sampling of conformational states from MD simulations and provides insights into the role of conformational fluctuations on longer time scales. It permits the identification of conformational sub-states along a reaction pathway that exhibit structural and dynamical properties important for function. Although a number of alternate approaches exist, the advantage of QAA lies in its ability to correctly identify energetically homogeneous conformational sub-states that correlate well with experimental information. Using T4 lysozyme and ubiquitin as model systems, Ramanathan *et al.* applied QAA to identify and characterize the hierarchical organization of conformational sub-states in these proteins [88]. The authors also characterized conformational sub-states associated with the cis/trans isomerization of proline peptide bonds catalyzed by the enzyme cyclophilin A (CypA)[88]. They identified a separate conformational sub-state corresponding to the transition state of the reaction pathway in CypA, where motions of clusters of residues were shown to adopt conformations that promote the transition state.

In combination with NMR relaxation dispersion experiments, QAA was recently used to identify the conformational sub-states associated with the substrate and product bound states of wild-type (WT) and a single point mutant form of RNase A [90]. MD simulation trajectories of the apo, reactant, and product-bound states were used for QAA. Analysis of the top QAA modes revealed reduced global conformational motions in the mutant, corresponding to the interconversion between the reactant to product states, relative to the

WT enzyme. Further, enhanced conformational dynamics of the surface loop 4 was observed in the apo and ligand-bound states of the mutant form. These enhanced loop motions were shown to be consistent with significant changes observed in the ligand binding and catalytic properties of the enzyme. Further analysis revealed a lack of correlated motions of two surface loop regions (loops 1 and 4) in the mutant, suggesting a role of this correlated motion for optimal catalysis in the WT enzyme [90].

Further illustrating the power of this combined experimental-theoretical method, Gagné *et al.* used $^{15}$N-CPMG, microsecond time-scale MD simulations, and QAA to investigate the conformational fluctuations associated with the binding of substrates in xylanase B2 from *Streptomyces lividans* (XlnB2) [82]. Using simulations of the apo enzyme and 6- and 9-unit xylan substrates bound to the enzyme as input for QAA, they characterized the dynamic modes and structural interactions of the enzyme that facilitate substrate binding (Figure 4). Details of the role of structural interactions of protein residues with ligand during the binding process and the range of motions (amplitude) that each residue undergoes were identified. It was discovered that the apo protein also samples the subset of conformations required for ligand binding. Quantitative characterization of these states would help in identifying the relationship between the rate of conformational exchange and the rate of ligand binding.

## 6. Conclusion

Technological advancements remain critical for improving NMR resolution and spectroscopic data acquisition when studying atomic-scale protein flexibility. The present review nevertheless illustrates that recent developments stem from the combination of clever isotope labeling schemes and carefully selected relaxation methods that either selectively target specific spin probes and/or improve spectral sensitivity, in addition to taking advantage of other biophysical techniques such as circular dichroism spectroscopy and/or isothermal titration calorimetry to characterize protein motions on multiple time frames. Finally, the combination of experimental NMR with computational methods aimed at predicting NMR chemical shifts, order parameters, and/or conformational sub-state populations (*e.g.* the quasi-anharmonic analysis, or QAA approach) now pave the way to an ever-increasing understanding of functionally relevant time-dependent motional events essential to biological function at the atomic level.

## Acknowledgments

The authors thank David Bernard (INRS) for providing experimental NMR relaxation dispersion data. This work was supported in part by a multi-PI grant from the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health (NIH) under award number R01GM105978 to N.D. and P.K.A., and a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant under award number RGPIN-2016-05557 (to N.D.). C.N. holds a Postdoctoral Fellowship from the Fondation Universitaire Armand-Frappier de l'INRS. N.D. holds a Fonds de Recherche Québec – Santé (FRQS) Research Scholar Junior 2 Career Award.

## References


1. Bhabha G, Lee J, Ekiert DC, Gam J, Wilson IA, Dyson HJ, Benkovic SJ, Wright PE. A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis. Science. 2011; 332:234–238. [PubMed: 21474759]
2. Hammes-Schiffer S, Benkovic SJ. Relating protein motion to catalysis. Annu Rev Biochem. 2006; 75:519–541. [PubMed: 16756501]
3. Luk LYP, Javier Ruiz-Pernía J, Dawson WM, Roca M, Loveridge EJ, Glowacki DR, Harvey JN, Mulholland AJ, Tuñón I, Moliner V, Allemann RK. Unraveling the role of protein dynamics in dihydrofolate reductase catalysis. Proc Natl Acad Sci U S A. 2013; 110:16344–16349. [PubMed: 24065822]
4. Warner LR, Gatzeva-Topalova PZ, Doerner PA, Pardi A, Sousa MC. Flexibility in the periplasmic domain of BamA is important for function. Structure. 2016 In Press.
5. Palmer AG, Massi F. Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. Chem Rev. 2006; 106:1700–1719. [PubMed: 16683750]
6. Anthis NJ, Clore GM. Visualizing transient dark states by NMR spectroscopy. Q Rev Biophys. 2015; 48:35–116. [PubMed: 25710841]
7. Kovermann M, Rogne P, Wolf-Watz M. Protein dynamics and function from solution state NMR spectroscopy. Q Rev Biophys. 2016; 49:e6. [PubMed: 27088887]
8. Lundström P, Vallurupalli P, Hansen DF, Kay LE. Isotope labeling methods for studies of excited protein states by relaxation dispersion NMR spectroscopy. Nat Protoc. 2009; 4:1641–1648. [PubMed: 19876024]
9. Sekhar A, Kay LE. NMR paves the way for atomic level descriptions of sparsely populated, transiently formed biomolecular conformers. Proc Natl Acad Sci U S A. 2013; 110:12867–12874. [PubMed: 23868852]
10. Boulton S, Melacini G. Advances in NMR methods to map allosteric sites: from models to translation. Chem Rev. 2016; 116:6267–6304. [PubMed: 27111288]
11. Grutsch S, Brüschweiler S, Tollinger M. NMR methods to study dynamic allostery. PLoS Comput Biol. 2016; 12:e1004620. [PubMed: 26964042]
12. Salmon L, Bouvignies G, Markwick P, Blackledge M. Nuclear magnetic resonance provides a quantitative description of protein conformational flexibility on physiologically important time scales. Biochemistry. 2011; 50:2735–2747. [PubMed: 21388216]
13. Guerry P, Salmon L, Mollica L, Ortega Roldan JL, Markwick P, van Nuland NAJ, McCammon JA, Blackledge M. Mapping the population of protein conformational energy sub-states from NMR dipolar couplings. Angew Chem Int Ed Engl. 2013; 52:3181–3185. [PubMed: 23371543]
14. Lisi GP, Loria JP. Solution NMR spectroscopy for the study of enzyme allostery. Chem Rev. 2016; 116:6323–6369. [PubMed: 26734986]
15. Clore GM, Iwahara J. Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. Chem Rev. 2009; 109:4108–4139. [PubMed: 19522502]
16. Jarymowycz VA, Stone MJ. Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences. Chem Rev. 2006; 106:1624–1671. [PubMed: 16683748]
17. Chen K, Tjandra N. The use of residual dipolar coupling in studying proteins by NMR. Top Curr Chem. 2012; 326:47–67. [PubMed: 21952837]
18. Sprangers R, Velyvis A, Kay LE. Solution NMR of supramolecular complexes: providing new insights into function. Nat Methods. 2007; 4:697–703. [PubMed: 17762877]
19. Xu Y, Matthews S. TROSY NMR spectroscopy of large soluble proteins. Top Curr Chem. 2013; 335:97–119. [PubMed: 21928013]
20. Zhang H, van Ingen H. Isotope-labeling strategies for solution NMR studies of macromolecular assemblies. Curr Opin Struct Biol. 2016; 38:75–82. [PubMed: 27295425]
21. Kay LE. New views of functionally dynamic proteins by solution NMR spectroscopy. J Mol Biol. 2016; 428:323–331. [PubMed: 26707200]

22. Sprangers R, Kay LE. Quantitative dynamics and binding studies of the 20S proteasome by NMR. Nature. 2007; 445:618–622. [PubMed: 17237764]
23. Palmer AG. Chemical exchange in biomacromolecules: past, present, and future. J Magn Reson. 2014; 241:3–17. [PubMed: 24656076]
24. Pervushin K, Riek R, Wider G, Wüthrich K. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci U S A. 1997; 94:12366–12371. [PubMed: 9356455]
25. Yuwen T, Vallurupalli P, Kay LE. Enhancing the sensitivity of CPMG relaxation dispersion to conformational exchange processes by multiple-quantum spectroscopy. Angew Chem Int Ed Engl. 2016; 55:11490–11494. [PubMed: 27527986]
26. Rennella E, Schuetz AK, Kay LE. Quantitative measurement of exchange dynamics in proteins via (13)C relaxation dispersion of (13)CHD2-labeled samples. J Biomol NMR. 2016; 65:59–64. [PubMed: 27251650]
27. Fenwick RB, Oyen D, Wright PE. Multi-probe relaxation dispersion measurements increase sensitivity to protein dynamics. Phys Chem Chem Phys PCCP. 2016; 18:5789–5798. [PubMed: 26426424]
28. Vallurupalli P, Bouvignies G, Kay LE. Studying “invisible” excited protein states in slow exchange with a major state conformation. J Am Chem Soc. 2012; 134:8148–8161. [PubMed: 22554188]
29. Bouvignies G, Kay LE. A 2D $^{13}$C-CEST experiment for studying slowly exchanging protein systems using methyl probes: an application to protein folding. J Biomol NMR. 2012; 53:303–310. [PubMed: 22689067]
30. Gu Y, Hansen AL, Peng Y, Brüschweiler R. Rapid determination of fast protein dynamics from NMR chemical exchange saturation transfer data. Angew Chem Int Ed Engl. 2016; 55:3117–3119. [PubMed: 26821600]
31. Mandel AM, Akke M, Palmer AG. Backbone dynamics of Escherichia coli ribonuclease HI: correlations with structure and function in an active enzyme. J Mol Biol. 1995; 246:144–163. [PubMed: 7531772]
32. Lipari G, Szabo A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1 Theory and range of validity. J Am Chem Soc. 1982; 104:4546–4559.
33. Lipari G, Szabo A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2 Analysis of experimental results. J Am Chem Soc. 1982; 104:4559–4570.
34. Vallurupalli P, Bouvignies G, Kay LE. Increasing the exchange time-scale that can be probed by CPMG relaxation dispersion NMR. J Phys Chem B. 2011; 115:14891–14900. [PubMed: 22077866]
35. Freiburger L, Auclair K, Mittermaier A. Global ITC fitting methods in studies of protein allostery. Methods. 2015; 76:149–161. [PubMed: 25573261]
36. Hansen LD, Transtrum MK, Quinn C, Demarse N. Enzyme-catalyzed and binding reaction kinetics determined by titration calorimetry. Biochim Biophys Acta. 2016; 1860:957–966. [PubMed: 26721335]
37. Lisi GP, Manley GA, Hendrickson H, Rivalta I, Batista VS, Loria JP. Dissecting dynamic allosteric pathways using chemically related small-molecule activators. Structure. 2016; 24:1155–1166. [PubMed: 27238967]
38. Rivalta I, Lisi GP, Snoeberger NS, Manley G, Loria JP, Batista VS. Allosteric communication disrupted by a small molecule binding to the imidazole glycerol phosphate synthase protein-protein interface. Biochemistry. 2016; 55:6484–6494. [PubMed: 27797506]
39. Chaudhuri BN, Lange SC, Myers RS, Davisson VJ, Smith JL. Toward understanding the mechanism of the complex cyclization reaction catalyzed by imidazole glycerolphosphate synthase: crystal structures of a ternary complex and the free enzyme. Biochemistry. 2003; 42:7003–7012. [PubMed: 12795595]
40. Lipchock JM, Loria JP. Nanometer propagation of millisecond motions in V-type allostery. Structure. 2010; 18:1596–1607. [PubMed: 21134639]

41. Cembran A, Kim J, Gao J, Veglia G. NMR mapping of protein conformational landscapes using coordinated behavior of chemical shifts upon ligand binding. Phys Chem Chem Phys PCCP. 2014; 16:6508–6518. [PubMed: 24604024]
42. Robustelli P, Stafford KA, Palmer AG. Interpreting protein structural dynamics from NMR chemical shifts. J Am Chem Soc. 2012; 134:6365–6374. [PubMed: 22381384]
43. Selvaratnam R, VanSchouwen B, Fogolari F, Mazhab-Jafari MT, Das R, Melacini G. The projection analysis of NMR chemical shifts reveals extended EPAC autoinhibition determinants. Biophys J. 2012; 102:630–639. [PubMed: 22325287]
44. Selvaratnam R, Chowdhury S, VanSchouwen B, Melacini G. Mapping allostery through the covariance analysis of NMR chemical shifts. Proc Natl Acad Sci U S A. 2011; 108:6133–6138. [PubMed: 21444788]
45. Boulton S, Akimoto M, Selvaratnam R, Bashiri A, Melacini G. A tool set to map allosteric networks through the NMR chemical shift covariance analysis. Sci Rep. 2014; 4:7306. [PubMed: 25482377]
46. Konuma T, Lee YH, Goto Y, Sakurai K. Principal component analysis of chemical shift perturbation data of a multiple-ligand-binding system for elucidation of respective binding mechanism. Proteins. 2013; 81:107–118. [PubMed: 22927212]
47. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A. 1998; 95:14863–14868. [PubMed: 9843981]
48. Sakurai K, Goto Y. Principal component analysis of the pH-dependent conformational transitions of bovine beta-lactoglobulin monitored by heteronuclear NMR. Proc Natl Acad Sci U S A. 2007; 104:15346–15351. [PubMed: 17878316]
49. Selvaratnam R, Mazhab-Jafari MT, Das R, Melacini G. The auto-inhibitory role of the EPAC hinge helix as mapped by NMR. PloS One. 2012; 7:e48707. [PubMed: 23185272]
50. Gagné D, Narayanan C, Doucet N. Network of long-range concerted chemical shift displacements upon ligand binding to human angiogenin. Protein Sci. 2015; 24:525–533. [PubMed: 25450558]
51. Axe JM, Boehr DD. Long-range interactions in the α subunit of tryptophan synthase help to coordinate ligand binding, catalysis, and substrate channeling. J Mol Biol. 2013; 425:1527–1545. [PubMed: 23376097]
52. Ponder JW, Case DA. Force fields for protein simulations. Adv Protein Chem. 2003; 66:27–85. [PubMed: 14631816]
53. Agarwal PK, Doucet N, Chennubhotla C, Ramanathan A, Narayanan C. Conformational sub-states and populations in enzyme catalysis. Methods Enzymol. 2016; 578:273–297. [PubMed: 27497171]
54. Narayanan C, Bernard D, Doucet N. Role of conformational motions in enzyme function: selected methodologies and case studies. Catalysts. 2016; 6:81. [PubMed: 28367322]
55. Hampton, S., Agarwal, PK., Alam, SR., Crozier, PS. Towards microsecond biological molecular dynamics simulations on hybrid processors. Proceedings of International Conference on High Performance Computing and Simulation; IEEE; 2010. p. 98-107.
56. Guo J, Zhou HX. Protein allostery and conformational dynamics. Chem Rev. 2016; 116:6503–6515. [PubMed: 26876046]
57. Wei G, Xi W, Nussinov R, Ma B. Protein ensembles: how does nature harness thermodynamic fluctuations for life? The diverse functional roles of conformational ensembles in the cell. Chem Rev. 2016; 116:6516–6551. [PubMed: 26807783]
58. Zhang H, Kurgan L. Sequence-based Gaussian network model for protein dynamics. Bioinforma Oxf Engl. 2014; 30:497–505.
59. Cilia E, Pancsa R, Tompa P, Lenaerts T, Vranken WF. From protein sequence to dynamics and disorder with DynaMine. Nat Commun. 2013; 4:2741. [PubMed: 24225580]
60. Feher VA, Durrant JD, Van Wart AT, Amaro RE. Computational approaches to mapping allosteric pathways. Curr Opin Struct Biol. 2014; 25:98–103. [PubMed: 24667124]
61. Reynolds KA, McLaughlin RN, Ranganathan R. Hot spots for allosteric regulation on protein surfaces. Cell. 2011; 147:1564–1575. [PubMed: 22196731]

62. Ramanathan A, Savol AJ, Langmead CJ, Agarwal PK, Chennubhotla CS. Discovering conformational sub-states relevant to protein function. PloS One. 2011; 6:e15827. [PubMed: 21297978]

63. Halabi N, Rivoire O, Leibler S, Ranganathan R. Protein sectors: evolutionary units of three-dimensional structure. Cell. 2009; 138:774–786. [PubMed: 19703402]

64. Lockless SW, Ranganathan R. Evolutionarily conserved pathways of energetic connectivity in protein families. Science. 1999; 286:295–299. [PubMed: 10514373]

65. Fisette O, Lagüe P, Gagné S, Morin S. Synergistic applications of MD and NMR for the study of biological systems. J Biomed Biotechnol. 2012; 2012:254208. [PubMed: 22319241]

66. Kohlhoff KJ, Robustelli P, Cavalli A, Salvatella X, Vendruscolo M. Fast and accurate predictions of protein NMR chemical shifts from interatomic distances. J Am Chem Soc. 2009; 131:13894–13895. [PubMed: 19739624]

67. Shen Y, Bax A. SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. J Biomol NMR. 2010; 48:13–22. [PubMed: 20628786]

68. Moon S, Case DA. A new model for chemical shifts of amide hydrogens in proteins. J Biomol NMR. 2007; 38:139–150. [PubMed: 17457516]

69. Han B, Liu Y, Ginzinger SW, Wishart DS. SHIFTX2: significantly improved protein chemical shift prediction. J Biomol NMR. 2011; 50:43–57. [PubMed: 21448735]

70. Liu Q, Shi C, Yu L, Zhang L, Xiong Y, Tian C. General order parameter based correlation analysis of protein backbone motions between experimental NMR relaxation measurements and molecular dynamics simulations. Biochem Biophys Res Commun. 2015; 457:467–472. [PubMed: 25600810]

71. Berjanskii MV, Wishart DS. The RCI server: rapid and accurate calculation of protein flexibility using chemical shifts. Nucleic Acids Res. 2007; 35:W531–537. [PubMed: 17485469]

72. Pastor N, Amero C. Information flow and protein dynamics: the interplay between nuclear magnetic resonance spectroscopy and molecular dynamics simulations. Front Plant Sci. 2015; 6:306. [PubMed: 25999971]

73. Krepl M, Cléry A, Blatter M, Allain FHT, Sponer J. Synergy between NMR measurements and MD simulations of protein/RNA complexes: application to the RRMs, the most common RNA recognition motifs. Nucleic Acids Res. 2016; 44:6452–6470. [PubMed: 27193998]

74. Camilloni C, Cavalli A, Vendruscolo M. Assessment of the use of NMR chemical shifts as replica-averaged structural restraints in molecular dynamics simulations to characterize the dynamics of proteins. J Phys Chem B. 2013; 117:1838–1843. [PubMed: 23327201]

75. Esteban-Martín S, Bryn Fenwick R, Salvatella X. Synergistic use of NMR and MD simulations to study the structural heterogeneity of proteins. Wiley Interdiscip Rev Comput Mol Sci. 2012; 2:466–478.

76. Showalter SA, Johnson E, Rance M, Brüschweiler R. Toward quantitative interpretation of methyl side-chain dynamics from NMR by molecular dynamics simulations. J Am Chem Soc. 2007; 129:14146–14147. [PubMed: 17973392]

77. Bouvignies G, Markwick PRL, Blackledge M. Simultaneous definition of high resolution protein structure and backbone conformational dynamics using NMR residual dipolar couplings. Chemphyschem. 2007; 8:1901–1909. [PubMed: 17654630]

78. Markwick PRL, Bouvignies G, Blackledge M. Exploring multiple timescale motions in protein GB3 using accelerated molecular dynamics and NMR spectroscopy. J Am Chem Soc. 2007; 129:4724–4730. [PubMed: 17375925]

79. Farès C, Lakomek NA, Walter KFA, Frank BTC, Meiler J, Becker S, Griesinger C. Accessing ns-micros side chain dynamics in ubiquitin with methyl RDCs. J Biomol NMR. 2009; 45:23–44. [PubMed: 19652920]

80. Markwick PRL, Bouvignies G, Salmon L, McCammon JA, Nilges M, Blackledge M. Toward a unified representation of protein structural dynamics in solution. J Am Chem Soc. 2009; 131:16968–16975. [PubMed: 19919148]

81. Manley G, Rivalta I, Loria JP. Solution NMR and computational methods for understanding protein allostery. J Phys Chem B. 2013; 117:3063–3073. [PubMed: 23445323]

82. Gagné D, Narayanan C, Nguyen-Thi N, Roux LD, Bernard DN, Brunzelle JS, Couture JF, Agarwal PK, Doucet N. Ligand binding enhances millisecond conformational exchange in xylanase B2 from Streptomyces lividans. Biochemistry. 2016; 55:4184–4196. [PubMed: 27387012]
83. Fisette O, Morin S, Savard PY, Lagüe P, Gagné SM. TEM-1 backbone dynamics-insights from combined molecular dynamics and nuclear magnetic resonance. Biophys J. 2010; 98:637–645. [PubMed: 20159160]
84. Fisette O, Gagné S, Lagüe P. Molecular dynamics of class A β-lactamases-effects of substrate binding. Biophys J. 2012; 103:1790–1801. [PubMed: 23083723]
85. Cole R, Loria JP. Evidence for flexibility in the function of ribonuclease A. Biochemistry. 2002; 41:6072–6081. [PubMed: 11994002]
86. Henzler-Wildman K, Kern D. Dynamic personalities of proteins. Nature. 2007; 450:964–972. [PubMed: 18075575]
87. Boehr DD, McElheny D, Dyson HJ, Wright PE. The dynamic energy landscape of dihydrofolate reductase catalysis. Science. 2006; 313:1638–1642. [PubMed: 16973882]
88. Ramanathan A, Agarwal PK. Evolutionarily conserved linkage between enzyme fold, flexibility, and catalysis. PLoS Biol. 2011; 9:e1001193. [PubMed: 22087074]
89. Ramanathan A, Savol A, Burger V, Chennubhotla CS, Agarwal PK. Protein conformational populations and functionally relevant substates. Acc Chem Res. 2014; 47:149–156. [PubMed: 23988159]
90. Gagné D, French RL, Narayanan C, Simonović M, Agarwal PK, Doucet N. Perturbation of the conformational dynamics of an active-site loop alters enzyme activity. Structure. 2015; 23:2256–2266. [PubMed: 26655472]
91. Li GC, Srivastava AK, Kim J, Taylor SS, Veglia G. Mapping the hydrogen bond networks in the catalytic subunit of protein kinase A using H/D fractionation factors. Biochemistry. 2015; 54:4042–4049. [PubMed: 26030372]
92. Skinner JJ, Lim WK, Bédard S, Black BE, Englander SW. Protein dynamics viewed by hydrogen exchange. Protein Sci. 2012; 21:996–1005. [PubMed: 22544544]
93. Kumar A, Balbach J. Real-time protein NMR spectroscopy and investigation of assisted protein folding. Biochim Biophys Acta. 2015; 1850:1965–1972. [PubMed: 25497212]
94. Rogne P, Sparrman T, Anugwom I, Mikkola JP, Wolf-Watz M. Realtime (31)P NMR investigation on the catalytic behavior of the enzyme adenylate kinase in the matrix of a switchable ionic liquid. ChemSusChem. 2015; 8:3764–3768. [PubMed: 26494201]
95. Farrow NA, Zhang O, Forman-Kay JD, Kay LE. A heteronuclear correlation experiment for simultaneous determination of 15N longitudinal decay and chemical exchange rates of systems in slow equilibrium. J Biomol NMR. 1994; 4:727–734. [PubMed: 7919956]
96. Li Y, Palmer AG. TROSY-selected ZZ-exchange experiment for characterizing slow chemical exchange in large proteins. J Biomol NMR. 2009; 45:357–360. [PubMed: 19890725]
97. Khirich G, Loria JP. Complexity of protein energy landscapes studied by solution NMR relaxation dispersion experiments. J Phys Chem B. 2015; 119:3743–3754. [PubMed: 25680027]
98. Sapienza PJ, Lee AL. Using NMR to study fast dynamics in proteins: methods and applications. Curr Opin Pharmacol. 2010; 10:723–730. [PubMed: 20933469]
99. Venditti V, Schwieters CD, Grishaev A, Clore GM. Dynamic equilibrium between closed and partially closed states of the bacterial Enzyme I unveiled by solution NMR and X-ray scattering. Proc Natl Acad Sci U S A. 2015; 112:11565–11570. [PubMed: 26305976]
100. Kempf JG, Loria JP. Measurement of intermediate exchange phenomena. Methods Mol Biol Clifton NJ. 2004; 278:185–231.

## Highlights

- Atomic flexibility on multiple time scales plays a defining role in protein function
- NMR relaxation allows the characterization of protein conformational sub-states
- Computational and biophysical applications further expand the range of NMR methods
- This review presents recent application progress to a general biochemical audience

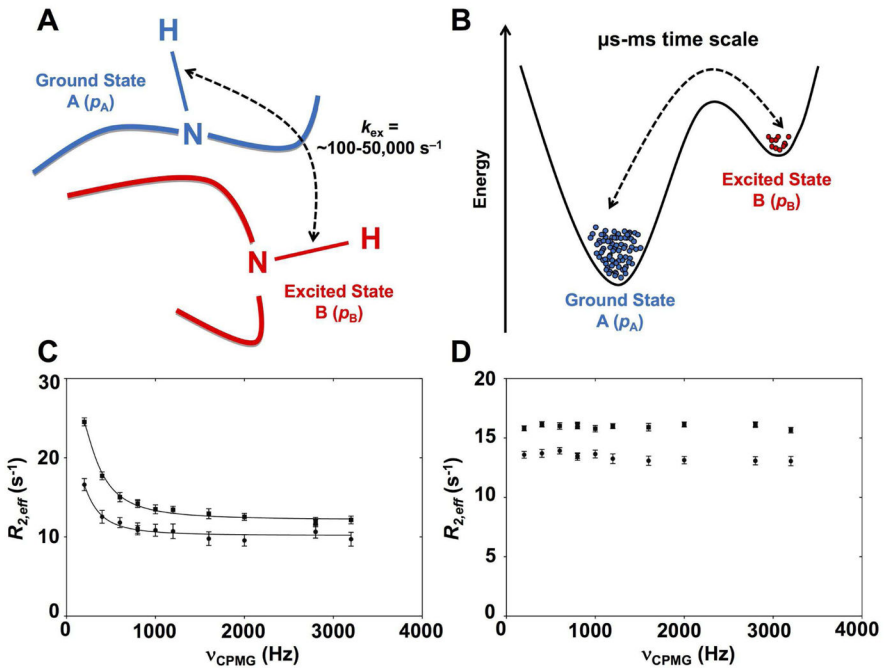



**Figure 1. Schematic depiction of exchange between two protein conformational substates on the µs-ms time scale**

A) Schematic representation of a backbone residue HN vector experiencing conformational exchange between ground state A and excited state B on the µs-ms time scale. The popular Carr-Purcell-Meiboom-Gill (CPMG) and $R_{1\rho}$ rotating frame relaxation dispersion experiments have been collectively employed to investigate conformational exchange rates ($k_{ex}$) for residues experiencing conformational exchange in proteins over time frames that roughly span ~100 to ~50,000 events per second ($s^{-1}$), overlapping the time scale of relevant biological events. B) Energetic representation of the two-site exchange between ground state A (higher population, $p_A$), and excited state B (lower population $p_B$, often invisible on fast and intermediate NMR time scales) [23,53,100]. NMR relaxation dispersion experiments can provide rates of exchange ($k_{ex}$), populations ($p_A p_B$), and chemical shifts between interconverting species ($\Delta\omega$). C) Representative experimental $^{15}$N-CPMG NMR curves at 500 MHz (circles) and 800 MHz (squares) for a backbone HN vector experiencing conformational exchange on the µs-ms time scale in a protein. D) Flat experimental $^{15}$N-CPMG NMR profiles at 500 MHz (circles) and 800 MHz (squares) for a backbone HN vector that does not experience conformational exchange on the µs-ms time scale. The somewhat limited atomistic details on directionality and length-scales provided by NMR for these movements can readily be complemented by MD simulations, which offer full atomistic details of the conformational populations.

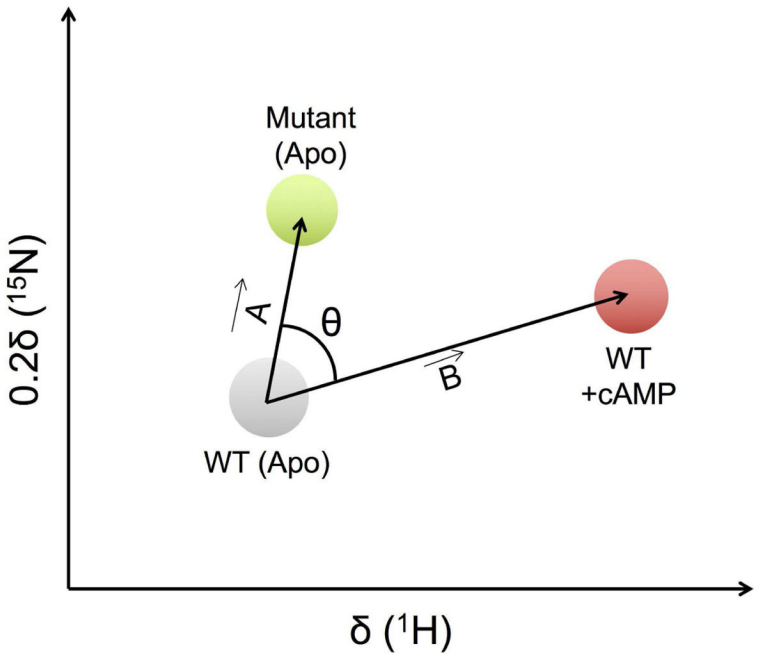



**Figure 2. Illustration of the CHESPA approach**

$^{1}H$-$^{15}N$ resonance assignments for apo, mutant and cAMP-bound forms of EPAC are represented as grey, green, and red circles, respectively. The compounded chemical shift upon mutation and ligand binding is calculated as the magnitude of vectors A and B, respectively (see text for details). $\theta$ represents the angle between vector A and B. Figure adapted from ref. [43].

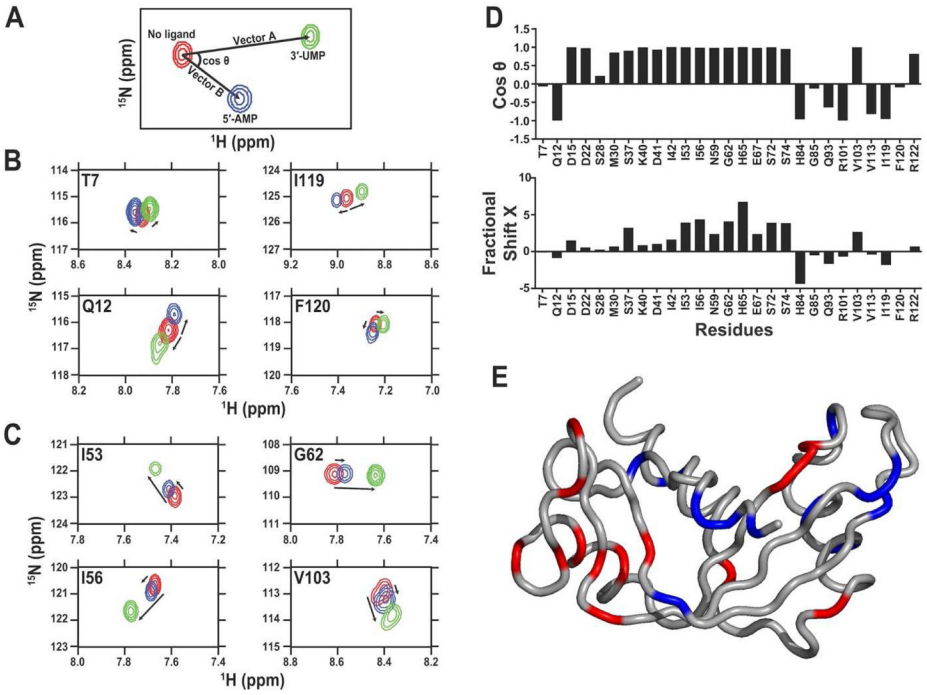



**Figure 3. Application of CHESPA to human angiogenin**

Projection analysis describing independent and coordinated residue variations upon 5′-AMP and 3′-UMP binding to human angiogenin subsites. (A) Graphical representation of the CHESPA approach described in ref. [43]. The $^{1}$H-$^{15}$N position of the peak is represented for the free (red) and bound forms (5′-AMP in blue and 3′-UMP in green) of the enzyme. Arrows indicate the movement of the $^{1}$H-$^{15}$N chemical shift (length and direction) for each peak from its origin (in red) to its saturated position (blue and green). Projection analysis of eight selected residues responding in (B) independent or (C) coordinated manner in human angiogenin. (D) Direction $\cos\theta$ and magnitude (fractional shift $X$) of the chemical shift perturbation of a subset of residues of angiogenin upon binding to saturation of the 5′-AMP and 3′-UMP. The $\cos\theta$ quantifies the angle between vectors A (3′-UMP) and B (5′-AMP) from its initial position in the free form of the enzyme. The fractional shift $X$ represents the fractional composite of vectors A and B induced by the ligand-induced chemical shift change. Collectiveness was observed for residues with a $\cos\theta \sim 1$, as previously described [43]. (E) Residues showing coordinated displacements ($\cos\theta \sim 1$) are represented in red on the three-dimensional structure of human angiogenin, while residues showing uncoordinated displacements ($\cos\theta \neq 1$) are shown in blue. Reprinted with permission from ref. [50]. Copyright 2015 Wiley.

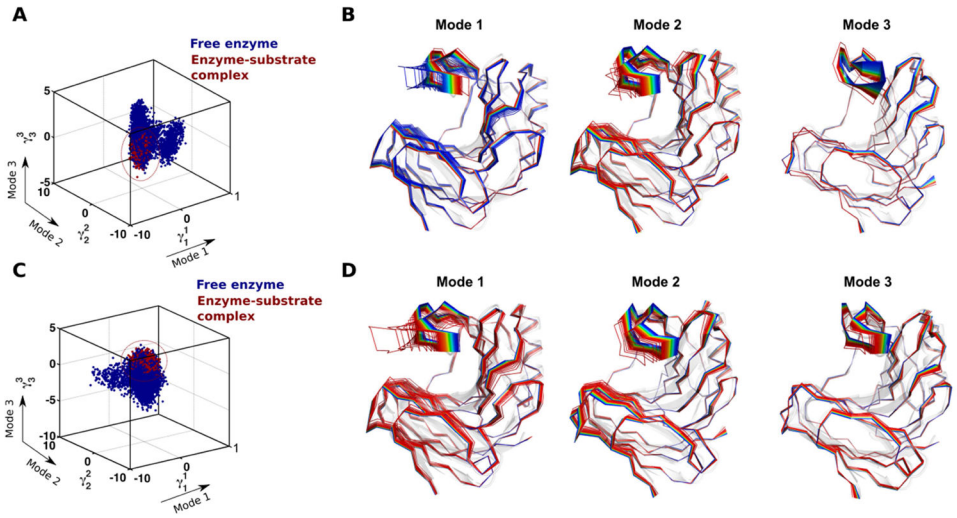



**Figure 4. Conformational sub-states of xylanase B2 (XlnB2) from *Streptomyces lividans* determined using computational simulations and QAA**

Representative conformations along the top three modes for the interconversion between free and (A–B) X6-bound and (C–D) X9-bound XlnB2 binary complexes. A total of 200,000 conformational snapshots obtained from the MD simulations were used as input for QAA to identify the top QAA-independent component vectors for characterizing the primary dynamics associated with the substrate binding process in XlnB2. Reprinted with permission from ref. [82]. Copyright 2016 American Chemical Society.

**Table 1**

General outline of different NMR methodologies employed to quantify protein dynamics on a broad range of time frames. Partly adapted from ref. [10].

| Method | Description | Information provided | Time Scale of Protein Motions Investigated | Exchange Rate[a] ($k_{ex}$) Characterized | Selected References |
|---|---|---|---|---|---|
| **H/D Exchange** | Detection of hydrogen exchange with a fully deuterated solvent: disappearance of amide proton signals over time | Solvent accessibility, hydrogen bond strength, slow conformational exchange | Seconds-minutes | | 91,92 |
| **Real-Time NMR** (RT NMR) | Direct detection of dynamic processes by quantifying the time-dependence of NMR signal intensity after a perturbation of the system | Protein folding, solvent hydrogen-exchange, slow conformational exchange | Seconds | $k_{ex} < 1\ s^{-1}$<br>$k_{ex} \ll \Delta\omega$ | 93,94 |
| **ZZ-exchange** (Exchange Spectroscopy; EXSY) | Simultaneous determination of $^{15}N$ longitudinal relaxation and chemical exchange during a mixing delay | Slow conformational exchange such as domain movement, ligand binding and release, topological interconversion of secondary structure, *cis-trans* isomerization | Milliseconds-seconds | $k_{ex} \approx 0.2–100\ s^{-1}$<br>$k_{ex} \ll \Delta\omega$ | 95,96 |
| **Chemical-and Dark-state Exchange Saturation Transfer** (CEST/DEST) | Saturation transfer by chemical exchange between invisible (minor, excited) state and visible (major, ground) state | Information on invisible states, populations of the states, exchange rates and slow-intermediate conformational exchange such as small domain movements | Milliseconds | $k_{ex} \approx 10–100\ s^{-1}$<br>$k_{ex} \leq \Delta\omega$ | 6,28 |
| **Carr-Purcell Meiboom-Gill Relaxation Dispersion** (CPMG RD) | Attenuation of chemical exchange broadening by applying refocusing pulses: reduce transverse relaxation $R_2$ sensitive to motions | Kinetic, thermodynami c, and structural information for intermediate-fast exchange processes such as side chain reorientation, loop motions, secondary structure changes and hinged domain movements | Microseconds-milliseconds | $k_{ex} \approx 200–6{,}000\ s^{-1}$<br>$k_{ex} \approx \Delta\omega$ | 37,82,97 |
| **$R_{1\rho}$ rotating frame relaxation dispersion** (RF RD) | Attenuation of chemical exchange broadening by applying refocusing pulses: reduce relaxation $R_{1\rho}$ sensitive to motions | Fast-intermediate conformational exchange such as motions of loops, side chains and secondary structure elements | Microseconds -milliseconds | $k_{ex} \approx 10{,}000–50{,}000\ s^{-1}$<br>$k_{ex} \geq \Delta\omega$ | 5 |
| **Paramagnetic Relaxation Enhancement** (PRE) | Increased rate of relaxation due to magnetic dipolar interaction between a nucleus and an unpaired electron of a spin label | Lowly populated states of macromolecules and their complexes, non-specific interactions, ligand binding sites, fast dynamic processes | Microseconds | $k_{ex} \approx 100{,}000\ s^{-1}$<br>$k_{ex} \gg \Delta\ \omega$ | 15 |
| **$R_1$, $R_2$ and heteronuclear NOE relaxation** (nuclear spin relaxation) | Analysis of longitudinal and transverse relaxation rates and the associated heteronuclear NOEs, sensitive to protein motions | Fast conformational exchange such as bond vibration and libration, side chain rotamer interconversion, random coil, loop motions, and backbone torsion angle rotation | Picoseconds - nanoseconds | | 16,98 |
| **Residual Dipolar Coupling** (RDC) | Combined analysis of dipolar coupling constants measured in an array of experimental conditions | Structural states of a protein in solution, orientation of bond vectors, fast conformational exchange | Picoseconds-milliseconds | | 17,99 |
| **Chemical Shift Covariance** (CHESCA) **or Projection** (CHESPA) **Analysis** | Analysis of residue networks that show correlated changes in chemical shifts due to perturbations in the | Allosteric networks, ligand binding sites | | | 43,44,49–51 |

| Method | Description | Information provided | Time Scale of Protein Motions Investigated | Exchange Rate[a] ($k_{ex}$) Characterized | Selected References |
|---|---|---|---|---|---|
| | system (mutations, ligand binding, etc.) | | | | |

[a] The exchange rate ($k_{ex}$) is proportional to the chemical shift difference (in Hertz) between two exchanging states ($\Delta\omega$).