



# Methyl TROSY Spectroscopy to Study Large Biomolecular Complexes 21

Milos A. Cvetkovic and Remco Sprangers

## Contents

Introduction .....	454
Methyl Group Labeling Schemes .....	455
Methyl Group Resonance Assignments .....	456
Divide and Conquer .....	457
Mutations to Assign Methyl Groups .....	457
Structure-Based Assignment Methods .....	457
The Methyl TROSY Experiment .....	458
Applications .....	459
Intermolecular Interactions .....	459
Methionine Scanning .....	460
Protein Dynamics .....	460
Elucidating Complex Structures .....	462
Conclusion .....	463
References .....	463

## Abstract

Solution state NMR spectroscopy is a powerful technique in structural biology that can provide unique information regarding the structure, dynamics, and interactions of biomolecular complexes. For a long time, its experimental range was limited to proteins of modest size. However, in the recent decades, the applicability of the method has been extended such that assemblies with molecular weights far over 100 kDa became amenable to detailed analyses. The breakthroughs that enabled these advances include the development of TROSY-based NMR techniques and procedures to produce samples that are labeled in specific methyl groups.

---

M. A. Cvetkovic · R. Sprangers (✉)  
Max Planck Institute for Developmental Biology, Tübingen, Germany  
e-mail: [milos.cvetkovic@tuebingen.mpg.de](mailto:milos.cvetkovic@tuebingen.mpg.de); [remco.sprangers@tuebingen.mpg.de](mailto:remco.sprangers@tuebingen.mpg.de)

Here, we discuss these novel approaches to the study of high molecular weight systems, explaining briefly the theoretical background behind the advancements and giving several recent practical examples. The major applications of methyl TROSY NMR spectroscopy are mentioned: studies of intermolecular interactions, protein dynamics, and complex biomolecular structures. With all this, we substantiate our notion that NMR spectroscopy will continue to be a highly valuable and relevant method for investigating large biomolecular complexes that is complementary to other structural techniques.

---

**Keywords**

Methyl TROSY · Large molecular machine · Methyl-group labeling · Biomolecular structure · Biomolecular interactions · Protein dynamics · Methyl-group resonance assignment · Relaxation dispersion · Longitudinal exchange · Methionine scanning · LEGO NMR · Segmental isotope labeling · RNA protein interactions

---

**Introduction**

Over the years, NMR spectroscopy has demonstrated its strength in studying structure, dynamics, and interactions of various biomolecules at atomic resolution and under near-physiological solution conditions. The size and complexity of biomolecules that can be studied has, however, been posing a limitation to its applicability. Two major difficulties associated with NMR studies of larger proteins and biomolecular complexes are signal overlap and line broadening. Signal overlap is the consequence of the vast number of NMR active nuclei present in large systems and results in overcrowded spectra that are impossible to analyze. Line broadening arises from the relaxation of transverse magnetization, which is faster in large systems due to longer rotational correlation times and more abundant spin-spin interactions. As a result, both sensitivity and resolution of the measurement decrease significantly for proteins larger than 20 kDa [1]

Large biomolecular complexes perform some of the most important processes in the cell and it is, therefore, important to extend the size limitations of the NMR spectroscopy. Initially, NMR studies of unlabeled proteins were conducted by  $^1\text{H}$  homonuclear experiments that were limited to systems of less than 10 kDa [2, 3]. Heteronuclear experiments on proteins uniformly labeled with  $^{15}\text{N}$  and  $^{13}\text{C}$  isotopes pushed the size limitation to around 25 kDa [4–6]. Protein deuteration had long been known to enhance resolution and sensitivity of the NMR experiments by reducing the number of possible spin relaxation pathways [7, 8]. Combined with  $^{15}\text{N}$  and  $^{13}\text{C}$  isotope labeling, deuteration extended the applicability of NMR spectroscopy to proteins of up to 50 kDa [9, 10]. Today,  $^{15}\text{N}$ ,  $^{13}\text{C}$ , and  $^2\text{H}$  labeling remains a standard procedure in studies of large protein systems. A revolutionary development in protein NMR spectroscopy was the introduction of transverse relaxation optimized spectroscopy (TROSY) [11], which enabled structure determination of proteins as large as 80 kDa [12] and analysis of proteins up to 900 kDa [13].

However, the adverse effects associated with size and complexity are often too pronounced to allow recording of useful  $^1\text{H}$ - $^{15}\text{N}$ -based spectra for very large macromolecular assemblies. Utilization of isotopically labeled methyl groups [ $^{13}\text{CH}_3$ ] in a fully deuterated background, in concert with the corresponding methyl TROSY spectra, has proven to be a successful approach for studying high molecular weight complexes. This strategy benefits from excellent relaxation properties of methyl groups where its three symmetrically arranged protons and fast rotation around its threefold symmetry axis produce highly sensitive and well-resolved NMR signal [14]. Methyl groups are usually located in the hydrophobic interior of proteins and along binding surfaces [15], which makes them valuable reporters of structural integrity, conformational changes, dynamics, and interactions. In summary, methyl TROSY NMR spectroscopy now enables recording high-quality spectra of assemblies with a molecular weight over 1 MDa [16].

---

## Methyl Group Labeling Schemes

Development of sophisticated isotope labeling schemes and improvements in the pulse sequence design have played pivotal roles for extending the applicability of NMR spectroscopy to large biomolecular systems. Optimal isotope labeling involves, essentially, finding the right balance between having an adequate amount, distribution, and/or localization of the NMR-active nuclei ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ,  $^1\text{H}$ ) to provide enough information about the system and, at the same time, not compromising quality and manageability of the recorded spectra with too many NMR-active nuclei, which would otherwise lead to signal overlap and broadening.

Isotope labeling can be uniform, affecting the whole protein, and selective, affecting particular residues or groups of residues. In methyl TROSY NMR spectroscopy, both approaches are applied: protein is uniformly deuterated, while labeled methyl ( $^{13}\text{CH}_3$ ) groups are selectively introduced in particular amino acids. Hydrogens ( $^1\text{H}$ ), naturally abundant in proteins, cause extensive relaxation of transverse magnetization through interactions with each other and with other NMR-active nuclei ( $^{15}\text{N}$ ,  $^{13}\text{C}$ ), thereby diminishing sensitivity and resolution. Hence, full deuteration (perdeuteration) is essential for studying large proteins as it minimizes this effect, due to 6.7-fold lower gyromagnetic ratio of  $^2\text{H}$  compared to  $^1\text{H}$ . Routinely, deuteration is achieved by growing the bacterial culture and overexpressing the protein in  $\text{D}_2\text{O}$ -based minimal medium in the presence of deuterated glucose [17].

Methyl ( $^{13}\text{CH}_3$ ) isotopes can be introduced into individual amino acids or in a combination of several amino acids. Strategies for production of such proteins depend on the metabolic pathways of the respective amino acids [18]. The simplest way to achieve methyl group labeling is to add a methyl-labeled biosynthetic precursor to the bacterial culture before inducing protein overexpression. This approach is possible if there is no subsequent crossing between the precursor and the metabolic pathways of other amino acids. The precursor 2-ketobutyrate can, hence, be used for production of Ile- $\delta_1$  [ $^{13}\text{CH}_3$ ] [19, 20] and 2-hydroxy-2-ethyl-3-ketobutyrate for production of Ile- $\gamma_2$  [ $^{13}\text{CH}_3$ ] [21]. The biosynthesis of valine and leucine, on the other hand, is connected

and their mutual precursor, 2-keto-3-isovalerate, is used for combined labeling of these amino acids [19]. Coproduction of stereospecifically labeled valine and leucine can be achieved using 2-acetolactate [22], while for separate labeling of valine and leucine residues one needs to prevent the scrambling of the precursor into both amino acids. As an example, stereospecifically labeled valine can be obtained either with labeled precursor 2-acetolactate and deuterated leucine [23] or with presynthesized stereospecifically labeled valine and deuterated leucine [24]. In both cases, the deuterated leucine prevents that the precursor is used in the leucine metabolic pathway. Metabolic pathways of the methyl groups of methionine, alanine, and threonine are intermingled with other metabolites and they need to be added before induction of protein expression, in their final, methyl-labeled forms: Met- $\epsilon$  [ $^{13}\text{CH}_3$ ] [25], Ala- $\beta$  [ $^{13}\text{CH}_3$ ] [26], Thr- $\gamma_2$  [ $^{13}\text{CH}_3$ ] [27].

Some of the strategies for the production of methyl-labeled proteins can be combined, which enables incorporation of different methyl reporters within a single protein simultaneously. The metabolic pathways used to produce labeled amino acids need to be compatible and the respective methyl TROSY resonances should not overlap in the spectrum. Combined labeling schemes that have been efficiently used are, for example, ILV [16, 28], MILV [29], AILV [30], and AMILVT [31].

Studies of large multidomain proteins could benefit from combining methyl TROSY NMR spectroscopy with the segmental isotope labeling. With the use of segmental isotope labeling, a specific isotope labeling strategy can be applied to a particular protein domain, while the remaining part of the protein remains NMR invisible. This simplifies the spectrum and allows for investigation of full-length proteins that would otherwise be out of reach for NMR spectroscopy [32]. Labeled and unlabeled parts of the protein are first produced separately by expressions under different labeling conditions and are subsequently ligated so that a peptide bond is formed between them. Ligation of the two parts of the protein is a critical step and can be performed either by using inteins, internally placed protein domains that are able to self-excise from a protein [33, 34] or by using transpeptidase Sortase A [35].

Work with large asymmetric protein complexes further stresses the need to reduce the number of signals in the spectrum to prevent signal overlap. This can be achieved by applying isotope labeling to only one or to a subset of subunits within the complex. In most protein systems, individual subunits may be insufficiently stable to sustain separate expression and purification, while in others *in vitro* reconstitution of the complex may be problematic. These issues can be overcome with the LEGO-NMR approach where subunits are sequentially coexpressed using different promoters, such that complex reconstitution takes place *in vivo* while only a subset of proteins is isotopically labeled [36].

---

## Methyl Group Resonance Assignments

A detailed analysis of NMR spectra requires that the resonances are assigned to the corresponding residues in the protein. The assignment of the methyl resonances in high molecular weight proteins and complexes poses a serious challenge.

For proteins up to ~25 kDa, methyl resonances can be assigned conventionally, using the general strategy for assigning aliphatic side chains, which relies on the direct correlation between methyl spin systems and the already assigned protein backbone [37]. For larger proteins, an advanced assignment method has been developed where unassigned methyl group resonances and assigned amide group resonances are independently correlated with C $\alpha$  and C $\beta$ , thus providing a link between the corresponding methyl and amide group [28, 38, 39]. However, both methods depend on the feasibility of backbone assignments, which are usually difficult or impossible to obtain for proteins larger than 50 kDa. Several strategies can be used to overcome this obstacle.

## **Divide and Conquer**

The “divide and conquer” strategy is based on dissecting the high molecular weight system into smaller building blocks [16, 29]. Since smaller building blocks usually show better spectral quality than the whole system, their methyl group resonances can be assigned using previously described standard approaches and afterwards linked to the corresponding methyl group signals of the whole system. This strategy has been successfully applied for protein complexes, where smaller building blocks represent individual subunits [16], as well as for large proteins, where smaller building blocks represent specific protein domains [29, 40].

## **Mutations to Assign Methyl Groups**

If backbone-based assignment methods and the divide and conquer approach fail, methyl group resonances can be assigned by mutagenesis. Individual amino acid residues containing a labeled methyl group are substituted with alternative residues. By comparing the respective spectra before and after mutagenesis, the signal(s) belonging to a particular methyl group can be readily identified [41, 42]. It should be noted, however, that this strategy can be complicated by secondary chemical shift perturbations, where the mutation of a single amino acid results in significant changes in the methyl TROSY spectrum.

## **Structure-Based Assignment Methods**

Structure-based procedures for methyl group assignments require detailed knowledge regarding the structure of the complex. Methyl-methyl distance information obtained from NOE data can, in those cases, be mapped with distances that are known from the structure, thereby assigning the methyl group resonances [43]. Another structure-based approach involves introduction of a spin-label at specific sites of the protein. Residues in the vicinity of such a label will consequently experience pseudocontact shifts (PCSs) [44] or paramagnetic relaxation enhancement (PRE) [45], enabling assignment of affected methyl groups.

## The Methyl TROSY Experiment

The intensity of an NMR signal directly depends on the amount of the initial magnetization and for that reason methyl groups that contain three protons are sensitive NMR probes. Furthermore, the NMR signal intensity depends on the decay rate of the magnetization during the experiment and significant improvements in spectral quality can be achieved with TROSY-based experiments.

During an NMR experiment, a large number of magnetization terms are created that all lose their magnetization with distinct relaxation rates. Due to destructive interference of relaxation mechanisms, a number of coherences can relax significantly slower than others. In TROSY-based experiments, these slow relaxing components are selected and special care is taken to prevent that they are mixed with fast relaxing ones. Such a mixing would result in an overall faster relaxation of the NMR signal and thus in NMR spectra of reduced quality.

The TROSY-based experiment was initially introduced for proton-nitrogen spectra that only detect the slow relaxing quarter of the NMR signal [11]. For large proteins, this results in NMR spectra of significantly improved quality compared to more traditional HSQC based experiments, where fast and slow relaxing coherence are mixed and added. The slow relaxation of specific coherences of H-N groups is caused by the destructive interference between the proton-nitrogen dipole-dipole and the nitrogen CSA (chemical shift anisotropy) relaxation interactions. The CSA depends on the magnetic field strength and an optimal cancellation of both relaxation mechanisms takes place at a magnetic field strength of around 1.1 GHz [11].

The carbon CSA in methyl groups is too small to induce destructive interference with the proton-carbon dipole. However, in the macromolecular limit, destructive interference occurs between the proton-carbon dipole-dipole and the multiple proton-proton dipolar relaxation interactions [14, 46]. To analyze the relaxation properties of a methyl group one needs to consider 16 different energy levels that are connected by 4 fast and 6 slowly relaxing proton transitions, 2 fast and 6 slowly relaxing carbon transitions, and 4 fast and 6 slowly relaxing  $^1\text{H}$ - $^{13}\text{C}$  double-/zero-quantum transitions. As was shown by Kay and coworkers, the rapidly and the slowly relaxing transitions are never mixed in the HMQC pulse sequence [14]. This renders the HMQC a highly efficient methyl TROSY experiment, where the final NMR signal of an isolated methyl group only results from slowly relaxing coherences. This is in strong contrast to the more popular HSQC-based experiment in which  $90^\circ$  proton pulses result in the mixing of fast and slow relaxing coherences and thus in NMR spectra with broader signals, especially for large proteins. To fully exploit the methyl TROSY effect, the methyl group of interest should be embedded in an otherwise fully deuterated protein (see above) as dipolar interactions with external protons cause interconversion of fast and slow relaxing coherences. Finally, it is worth noting that the methyl TROSY effect does not depend on the magnetic field strength as it results from the destructive interference between dipolar relaxation interactions [14].

## Applications

As previously explained, methyl TROSY NMR spectroscopy benefits significantly from excellent spectral features of the methyl groups and offers a wide range of possibilities to study large biomolecular systems. It is able to yield the same quantitative information about a complex system as was only available for small proteins until recently [47]. The applications include studies of intermolecular interactions and protein dynamics, revealing novel mechanisms of biomolecular processes, as well as providing new insights into the structure of very large biomolecular systems.

## Intermolecular Interactions

NMR spectroscopy is applicable to binding events with affinities that range from the nM to mM regime. Methyl bearing side-chains are generally well-suited probes for investigating binding surfaces as they often play an important role in biomolecular interactions. Changes in the local chemical environment, due to ligand binding, will affect methyl groups positioned within the binding interface. This causes chemical shift perturbations (CSPs) [48] that provide both qualitative and quantitative information about the interaction. Here, we describe a number of recent examples, where methyl TROSY NMR-based binding experiments have provided pivotal insights into biological function.

The molecular chaperone Hsp90, which forms a homodimer of 170 kDa, has an important role in protein folding. ATP and various cochaperones, including its binding partner p23, control its function. Hsp90 consists of three domains: an N-terminal domain, a middle domain, and a C-terminal domain. Stability of the isolated domains enabled successful assignments of Ile- $\delta_1$  methyl groups through divide and conquer strategy [49]. By observing CSPs in the methyl TROSY spectrum upon the addition of ligands, it was shown that ATP binds only to the N-terminal domain, while p23 cochaperone binds both the N-terminal and the middle domains of Hsp90 [49]. In addition, a 106 Å long interface was identified on Hsp90 that mediates the interaction with the intrinsically disordered Tau protein through many low-affinity contacts [50].

Binding studies in even larger complexes, like the 20S proteasome (670 kDa), show that methyl TROSY-based experiments are also feasible for systems of that size. The proteasome is responsible for degradation of damaged and dispensable proteins and its barrel-shaped core particle consists of four homo-heptameric stacked rings. The two outer rings, which form the entrance for substrates, can each bind the 150 kDa 11S activator, resulting in a complex with a molecular weight of 1.1 MDa. Methyl labeling of the outer subunits at Ile- $\delta_1$ , Leu- $\delta$ , and Val- $\gamma$  positions, assignment of their resonances through the divide and conquer strategy, and successful reconstitution of the 20S complex enabled mapping of the 11S binding surface on the 20S and determination of the associated affinity [16].

Ile- $\delta_1$ , Leu- $\delta$ , and Val- $\gamma$  methyl labeling was also used for the study of the interaction between the molecular chaperons ClpB (580 kDa) and DnaK (70 kDa)

that is crucial for protein disaggregation. Methyl labeling of only one component of the complex at a time, while keeping the other parts NMR invisible, provided high quality methyl TROSY spectra that were exploited to map interaction surfaces in a quantitative manner [51].

## Methionine Scanning

Methyl-containing amino acids that can be used in methyl TROSY experiments might not be present on the protein surface with the quantity and distribution to allow for a detailed mapping of interaction sites. To overcome this drawback, the methionine scanning approach has recently been introduced, which provides an increased coverage of the protein surface with methyl probes [52, 53]. It involves the strategic substitution, one-at-a-time, of solvent exposed residues with methyl-labeled methionine (Met- $\epsilon$ ). The signal of the introduced reporter methionine appears as a novel resonance in the methyl TROSY spectrum and can, thus, be instantly assigned. After the addition of the ligand, a new methyl TROSY spectrum is recorded. If the reporter Met is located inside the binding interface, it will experience new local chemical environment, which will be manifested as CSP. If the reporter is located outside of the binding interface, there will be no CSP of the corresponding signal. If a residue that is pivotal for the interaction (a hot-spot) is substituted with a reporter methionine, the interaction will be abolished, which is noticeable in the absence of CSPs of the naturally occurring methyl groups.

Methionine scanning was recently successfully employed for the study of RNA: protein interactions within the archaeal exosome. The exosome complex is an important molecular machine responsible for RNA 3' to 5' processing and degradation. The archaeal complex has a molecular weight of 270 kDa and is composed of a hexameric core and a trimeric cap. The hexameric core is a trimer of Rrp41-Rrp42 heterodimers, while the cap is made of three copies of the Rrp4 protein. The Rrp4 cap forms the opening through which RNA substrate enters the catalytic interior of the assembly. Based on the methionine scanning approach, a 50 Å long RNA binding path on each Rrp4 protomer was identified. The interaction between the Rrp4 cap and the RNA substrate, which was unattainable by the crystallographic data, proved to be crucial for the efficient recruitment and channeling of the RNA substrate towards the active sites [54].

## Protein Dynamics

### Experiments for Millisecond Methyl Dynamics

Proteins are highly dynamic biomolecules that can adopt multiple conformations. Enzymes, in particular, sample structurally different states to perform biological functions. As enzyme turnover rates in biology are often in the range between 0.1 and 5000 per second, there is a special interest in detecting motions on these timescales. Here, we briefly discuss longitudinal exchange and CPMG (Carr–Purcell–Meiboom–Gill)



relaxation dispersion experiments that are both able to detect and quantify such motions. We focus on the applicability of these methods to ( $^{13}\text{CH}_3$ ) methyl groups in large protein complexes, as these samples are ideally suited for methyl TROSY spectroscopy [55].

Longitudinal exchange experiments are applicable to systems where the spin of interest exchanges between two states (A and B) with a rate that is slow compared to the chemical shift difference between the two states ( $k_{\text{ex}} \ll \Delta\omega$ ; where  $k_{\text{ex}}$  is the exchange rate  $k_{\text{AB}} + k_{\text{BA}}$  and  $\Delta\omega$  is the chemical shift difference between states A and B). As a result, one single methyl group gives rise to two different signals in the NMR spectrum, one where the protein adopts conformation A and one where it adopts conformation B. Central to the HMQC-based longitudinal exchange experiment is a delay that is sandwiched between the carbon and proton chemical shift evolution times. In this delay the protein can change its conformation such that state A becomes state B and vice versa. In the NMR spectrum, this results in the appearance of resonances at the carbon chemical shift of state A (or B) and at the proton chemical shift of state B (or A). The dependence of the intensities of these “cross peaks” on the length of the delay directly reports on the kinetics behind the exchange process. Despite the fact that the methyl TROSY principle cannot be exploited during the complete NMR pulse sequence, longitudinal exchange experiments have been successfully applied to very large protein complexes [41, 56–58].

CPMG relaxation dispersion experiments [59] are applicable to systems where exchange of a protein from state A to state B results in the broadening of the NMR resonances. Such broadening is induced when the exchange takes place on a timescale that is comparable to the chemical shift difference between the two states ( $k_{\text{ex}} \sim \Delta\omega$ ), which, in turn, results in a dephasing of the magnetization. The extent of this dephasing depends on the difference in the chemical shift between states A and B ( $\Delta\omega$ ), on the exchange rate ( $k_{\text{ex}}$ ), and on the populations of the two states. Interestingly, significant exchange broadening of the resonance of state A can also occur when state B is only sparsely populated (e.g., less than 5%) and therefore not directly observable in the NMR spectrum. In such a situation, line broadening of the resonance of state A can report on the presence of an “invisible” state B. The dephasing of the magnetization (and thus the line-broadening that is induced by the exchange process) can be suppressed by a train of refocusing pulses. The dependence of the line broadening on the frequency with which the refocusing pulses are applied is then used to extract the kinetic parameters that underlie the exchange process. CPMG relaxation dispersion experiments can be recorded in a variety of different manners, depending on the magnetization state during the time when the refocusing pulses are applied. First,  $^1\text{H}$ - $^{13}\text{C}$  multiple quantum (MQ)-based experiments, which can be recorded in a manner that is fully compatible with the methyl TROSY principle, report on the proton and carbon chemical shift differences of states A and B [60]. The information content of MQ dispersion experiments is, thus, very high, which can make the analysis of the data complicated. Second, single quantum (SQ)-based experiments can be recorded on  $^{13}\text{CH}_3$  labeled methyl groups. These experiments can be designed such that only the  $^{13}\text{C}$  [61, 62] or only the  $^1\text{H}$  [63] chemical shift difference is sensed in the relaxation dispersion profiles.

Compared to the MQ experiments, the analysis of the SQ relaxation dispersion data is less complicated. These SQ experiments are, however, not as sensitive as the MQ experiment since only a part of the magnetization is selected and the methyl TROSY effect cannot be fully exploited. Nevertheless,  $^{13}\text{C}$  SQ relaxation dispersion experiments have been successfully applied to protein complexes over 100 kDa to quantify exchange processes [64]. Finally, a  $^1\text{H}$  triple quantum (TQ) relaxation dispersion experiment has been introduced recently, where the dispersion profiles depend on three times the proton chemical shift difference between states A and B [65]. Importantly, the dispersions in these experiments are by a factor of 10 larger than in the  $^1\text{H}$  (SQ) relaxation dispersion experiments and thus applicable to larger protein complexes and to a wider range of exchange processes.

In general, extracting accurate exchange parameters (exchange rates, populations, and chemical shift differences) from a single measurement at a single magnetic field strength is very challenging. To improve the accuracy of the extracted parameters, the relaxation dispersion experiment can be repeated on multiple field strengths, as the chemical shift difference depends on the spectrometer field. Alternatively, different relaxation dispersion experiments (SQ/MQ/TQ) can be analyzed simultaneously to improve the robustness of the fitting of the data.

### Example of Dynamics

As mentioned above, dynamic processes are often related with enzymatic function and NMR relaxation experiments are ideally suited to reveal potential correlations between dynamic processes and catalytic activity. DspS is an 80 kDa homodimeric enzyme that catalyzes the hydrolysis the 5' cap structure from short eukaryotic mRNAs, as the final step of their degradation pathway. The enzyme consists of a smaller N-terminal domain, which is flexibly connected to a larger C-terminal domain. Substrate can be positioned between the N and C-terminal domains, on both sides of the homodimer. The two binding sites strongly influence each other and the two substrates interact with the enzyme with significantly different affinities [58]. To quantify the motions of the N-terminal domain, longitudinal exchange experiments were performed on the enzyme that was produced with methyl-labeled Ile- $\delta_1$  and Met- $\epsilon$ . The study of intramolecular dynamics of this flexible system and its association with the substrate binding and activity revealed that the excess of flipping motion, induced by the binding of a second ligand, hampers the catalytic activity [58]. This finding highlights the relationship between intramolecular dynamics and activity and, hence, also the importance of understanding and quantifying intramolecular motions in enzymes.

### Elucidating Complex Structures

Complex assemblies that contain several different subunits and/or co-factors perform many important biological processes. In order to understand the structural and functional relationships between the components involved in these processes, it is necessary to combine different structural, biophysical, and biochemical methods.

The archaeal box C/D ribonucleoprotein enzyme is a highly complex system that contains the proteins L7Ae, Nop5 and fibrillarin, plus a guide sRNA. This 390 kDa complex methylates ribosomal RNA at the 2'-O-ribose, which is an important part of the pre-rRNA maturation and a necessary step for subsequent ribosome assembly. The structure of certain isolated segments of this ribonucleoprotein system was known, and these structures were assembled into the complete C/D RNP complex with the help of methyl CSPs, PRE experiments, SAXS (small angle X-ray scattering), and SANS (small angle neutron scattering) data. Furthermore, methyl CSPs observed upon addition of the RNA substrate were instrumental for understanding the elaborate mechanism of sequential site-specific methylation [66].

SecB is a molecular chaperone with a strong antifolding activity. This 70 kDa tetramer displays methyl TROSY spectra of high quality when expressed with methyl-labeled Ala- $\beta$  Val- $\gamma$ , Leu- $\delta$ , Met- $\epsilon$ , Thr- $\gamma$ 2, and Ile- $\delta$ 1. These methyl groups were used for determining the structure of the complex that SecB forms with its client proteins MBP and PhoA in their unfolded state. This clearly revealed the mechanism by which the molecular chaperons are able to keep client proteins in an unfolded state. In brief, it was shown that SecB forms long, continuous hydrophobic grooves that bind multiple hydrophobic segments exposed across the unfolded client protein [31]. Multivalent binding mode of this interaction leads to a structure where the client protein is wrapped around SecB, which affects the folding kinetics of the substrate and keeps it in an unfolding state.

---

## Conclusion

Here, we discussed the methodology that can extend the applicability of solution state NMR spectroscopy to systems that are orders of magnitude larger than those that are traditionally studied by this technique. Indeed, the combination of methyl-labeled samples and methyl TROSY experiments can provide quantitative insights in assemblies up to 1 MDa in molecular weight. Importantly, NMR spectroscopy is able to localize and quantify interactions and dynamic processes on a per residue basis. This information is highly relevant to biomolecular processes and often hidden in static structures. In that light, NMR spectroscopy is able to provide unique insights that are fully complementary to X-ray crystallography and cryo-EM methods. Indeed, recently the complementarity of methyl TROSY NMR spectroscopy and cryo-EM was impressively illustrated for a large AAA+ unfoldase [67]. We anticipate that approaches like that will be increasingly important in the future and we look forward to studies that will unravel a wide range of yet unexplored molecular mechanisms.

---

## References

1. Goto NK, Kay LE. New developments in isotope labeling strategies for protein solution NMR spectroscopy. *Curr Opin Struct Biol.* 2000;10(5):585–92.

2. Wagner G, Wuthrich K. Sequential resonance assignments in protein <sup>1</sup>H nuclear magnetic resonance spectra. Basic pancreatic trypsin inhibitor. *J Mol Biol.* 1982;155(3):347–66.
3. Williamson MP, Havel TF, Wuthrich K. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by <sup>1</sup>H nuclear magnetic resonance and distance geometry. *J Mol Biol.* 1985;182(2):295–315.
4. Fesik SW, Gampe Jr RT, Zuiderweg ER, Kohlbrenner WE, Weigl D. Heteronuclear three-dimensional NMR spectroscopy applied to CMP-KDO synthetase (27.5 kD). *Biochem Biophys Res Commun.* 1989;159(2):842–7.
5. Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM, et al. Overcoming the overlap problem in the assignment of <sup>1</sup>H NMR spectra of larger proteins by use of three-dimensional heteronuclear <sup>1</sup>H-<sup>15</sup>N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 beta. *Biochemistry.* 1989;28(15):6150–6.
6. Ikura M, Kay LE, Bax A. A novel approach for sequential assignment of <sup>1</sup>H, <sup>13</sup>C, and <sup>15</sup>N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry.* 1990;29(19):4659–67.
7. Crespi HL, Rosenberg RM, Katz JJ. Proton magnetic resonance of proteins fully deuterated except for <sup>1</sup>H-leucine side chains. *Science.* 1968;161(3843):795–6.
8. Markley JL, Putter I, Jardetzky O. High-resolution nuclear magnetic resonance spectra of selectively deuterated staphylococcal nuclease. *Science.* 1968;161(3847):1249–51.
9. Garrett DS, Seok YJ, Liao DI, Peterkofsky A, Gronenborn AM, Clore GM. Solution structure of the 30 kDa N-terminal domain of enzyme I of the *Escherichia coli* phosphoenolpyruvate: sugar phosphotransferase system by multidimensional NMR. *Biochemistry.* 1997;36(9):2517–30.
10. Mueller GA, Choy WY, Yang D, Forman-Kay JD, Venters RA, Kay LE. Global folds of proteins with low densities of NOEs using residual dipolar couplings: application to the 370-residue maltodextrin-binding protein. *J Mol Biol.* 2000;300(1):197–212.
11. Pervushin K, Riek R, Wider G, Wuthrich 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(23):12366–71.
12. Tugarinov V, Choy WY, Orekhov VY, Kay LE. Solution NMR-derived global fold of a monomeric 82-kDa enzyme. *Proc Natl Acad Sci U S A.* 2005;102(3):622–7.
13. Fiaux J, Bertelsen EB, Horwich AL, Wuthrich K. NMR analysis of a 900K GroEL GroES complex. *Nature.* 2002;418(6894):207–11.
14. Tugarinov V, Hwang PM, Ollershaw JE, Kay LE. Cross-correlated relaxation enhanced <sup>1</sup>H [bond]<sup>13</sup>C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J Am Chem Soc.* 2003;125(34):10420–8.
15. Janin J, Miller S, Chothia C. Surface, subunit interfaces and interior of oligomeric proteins. *J Mol Biol.* 1988;204(1):155–64.
16. Sprangers R, Kay LE. Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature.* 2007;445(7128):618–22.
17. Gardner KH, Kay LE. The use of <sup>2</sup>H, <sup>13</sup>C, <sup>15</sup>N multidimensional NMR to study the structure and dynamics of proteins. *Annu Rev Biophys Biomol Struct.* 1998;27:357–406.
18. Kerfah R, Plevin MJ, Sounier R, Gans P, Boisbouvier J. Methyl-specific isotopic labeling: a molecular tool box for solution NMR studies of large proteins. *Curr Opin Struct Biol.* 2015;32:113–22.
19. Goto NK, Gardner KH, Mueller GA, Willis RC, Kay LE. A robust and cost-effective method for the production of Val, Leu, Ile (delta 1) methyl-protonated <sup>15</sup>N-, <sup>13</sup>C-, <sup>2</sup>H-labeled proteins. *J Biomol NMR.* 1999;13(4):369–74.
20. Gardner KH, Kay LE. Production and incorporation of N-<sup>15</sup>, C-<sup>13</sup>, H-<sup>2</sup> (H-1-delta 1 methyl) isoleucine into proteins for multidimensional NMR studies. *J Am Chem Soc.* 1997;119(32):7599–600.
21. Ruschak AM, Velyvis A, Kay LE. A simple strategy for (1)(3)C, (1)H labeling at the Ile-gamma2 methyl position in highly deuterated proteins. *J Biomol NMR.* 2010;48(3):129–35.

22. Gans P, Hamelin O, Sounier R, Ayala I, Dura MA, Amero CD, et al. Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. *Angew Chem.* 2010;49(11):1958–62.
23. Mas G, Crublet E, Hamelin O, Gans P, Boisbouvier J. Specific labeling and assignment strategies of valine methyl groups for NMR studies of high molecular weight proteins. *J Biomol NMR.* 2013;57(3):251–62.
24. Miyanoiri Y, Takeda M, Okuma K, Ono AM, Terauchi T, Kainosho M. Differential isotope-labeling for Leu and Val residues in a protein by *E. coli* cellular expression using stereo-specifically methyl labeled amino acids. *J Biomol NMR.* 2013;57(3):237–49.
25. Fischer M, Kloiber K, Hausler J, Ledolter K, Konrat R, Schmid W. Synthesis of a  $^{13}\text{C}$ -methyl-group-labeled methionine precursor as a useful tool for simplifying protein structural analysis by NMR spectroscopy. *Chembiochem.* 2007;8(6):610–2.
26. Ayala I, Sounier R, Use N, Gans P, Boisbouvier J. An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. *J Biomol NMR.* 2009;43(2):111–9.
27. Velyvis A, Ruschak AM, Kay LE. An economical method for production of  $(2)\text{H}$ ,  $(13)\text{CH}_3$ -threonine for solution NMR studies of large protein complexes: application to the 670 kDa proteasome. *PLoS One.* 2012;7(9):e43725.
28. Tugarinov V, Kay LE. Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *J Am Chem Soc.* 2003;125(45):13868–78.
29. Gelis I, Bonvin AM, Keramisanou D, Koukaki M, Gouridis G, Karamanou S, et al. Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell.* 2007;131(4):756–69.
30. Godoy-Ruiz R, Guo C, Tugarinov V. Alanine methyl groups as NMR probes of molecular structure and dynamics in high-molecular-weight proteins. *J Am Chem Soc.* 2010;132(51):18340–50.
31. Huang C, Rossi P, Saio T, Kalodimos CG. Structural basis for the antifolding activity of a molecular chaperone. *Nature.* 2016;537(7619):202–6.
32. Rosenzweig R, Farber P, Velyvis A, Rennella E, Latham MP, Kay LE. ClpB N-terminal domain plays a regulatory role in protein disaggregation. *Proc Natl Acad Sci U S A.* 2015;112(50):E6872–81.
33. Mootz HD, Blum ES, Tyszkiewicz AB, Muir TW. Conditional protein splicing: a new tool to control protein structure and function in vitro and in vivo. *J Am Chem Soc.* 2003;125(35):10561–9.
34. Berrade L, Camarero JA. Expressed protein ligation: a resourceful tool to study protein structure and function. *Cell Mol Life Sci.* 2009;66(24):3909–22.
35. Freiburger L, Sonntag M, Hennig J, Li J, Zou P, Sattler M. Efficient segmental isotope labeling of multi-domain proteins using Sortase A. *J Biomol NMR.* 2015;63(1):1–8.
36. Mund M, Overbeck JH, Ullmann J, Sprangers R. LEGO-NMR spectroscopy: a method to visualize individual subunits in large heteromeric complexes. *Angew Chem.* 2013;52(43):11401–5.
37. Grzesiek S, Bax A. Amino acid type determination in the sequential assignment procedure of uniformly  $^{13}\text{C}/^{15}\text{N}$ -enriched proteins. *J Biomol NMR.* 1993;3(2):185–204.
38. Tugarinov V, Kay LE. Side chain assignments of Ile delta 1 methyl groups in high molecular weight proteins: an application to a 46 ns tumbling molecule. *J Am Chem Soc.* 2003;125(19):5701–6.
39. Mishra SH, Frueh DP. Assignment of methyl NMR resonances of a 52 kDa protein with residue-specific 4D correlation maps. *J Biomol NMR.* 2015;62(3):281–90.
40. Ogunjimi AA, Wiesner S, Briant DJ, Varelas X, Sicheri F, Forman-Kay J, et al. The ubiquitin binding region of the Smurf HECT domain facilitates polyubiquitylation and binding of ubiquitylated substrates. *J Biol Chem.* 2010;285(9):6308–15.
41. Sprangers R, Gribun A, Hwang PM, Houry WA, Kay LE. Quantitative NMR spectroscopy of supramolecular complexes: dynamic side pores in ClpP are important for product release. *Proc Natl Acad Sci U S A.* 2005;102(46):16678–83.

42. Amero C, Asuncion Dura M, Noirclerc-Savoie M, Perollier A, Gallet B, Plevin MJ, et al. A systematic mutagenesis-driven strategy for site-resolved NMR studies of supramolecular assemblies. *J Biomol NMR*. 2011;50(3):229–36.
43. Xiao Y, Warner LR, Latham MP, Ahn NG, Pardi A. Structure-based assignment of Ile, Leu, and Val methyl groups in the active and inactive forms of the mitogen-activated protein kinase extracellular signal-regulated kinase 2. *Biochemistry*. 2015;54(28):4307–19.
44. John M, Schmitz C, Park AY, Dixon NE, Huber T, Otting G. Sequence-specific and stereospecific assignment of methyl groups using paramagnetic lanthanides. *J Am Chem Soc*. 2007;129(44):13749–57.
45. Venditti V, Fawzi NL, Clore GM. Automated sequence- and stereo-specific assignment of methyl-labeled proteins by paramagnetic relaxation and methyl-methyl nuclear Overhauser enhancement spectroscopy. *J Biomol NMR*. 2011;51(3):319–28.
46. Ollerenshaw JE, Tugarinov V, Kay LE. Methyl TROSY: explanation and experimental verification. *Magn Reson Chem*. 2003;41(10):843–52.
47. Rosenzweig R, Kay LE. Bringing dynamic molecular machines into focus by methyl-TROSY NMR. *Annu Rev Biochem*. 2014;83:291–315.
48. Williamson MP. Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Magn Reson Spectrosc*. 2013;73:1–16.
49. Karagoz GE, Duarte AM, Ippel H, Utrecht C, Sinnige T, van Rosmalen M, et al. N-terminal domain of human Hsp90 triggers binding to the cochaperone p23. *Proc Natl Acad Sci U S A*. 2011;108(2):580–5.
50. Karagoz GE, Duarte AM, Akoury E, Ippel H, Biernat J, Moran Luengo T, et al. Hsp90-Tau complex reveals molecular basis for specificity in chaperone action. *Cell*. 2014;156(5):963–74.
51. Rosenzweig R, Moradi S, Zarrine-Afsar A, Glover JR, Kay LE. Unraveling the mechanism of protein disaggregation through a ClpB-DnaK interaction. *Science*. 2013;339(6123):1080–3.
52. Stoffregen MC, Schwer MM, Renschler FA, Wiesner S. Methionine scanning as an NMR tool for detecting and analyzing biomolecular interaction surfaces. *Structure*. 2012;20(4):573–81.
53. Mari S, Ruetalo N, Maspero E, Stoffregen MC, Pasqualato S, Polo S, et al. Structural and functional framework for the autoinhibition of Nedd4-family ubiquitin ligases. *Structure*. 2014;22(11):1639–49.
54. Cvetkovic MA, Wurm JP, Audin MJ, Schutz S, Sprangers R. The Rrp4-exosome complex recruits and channels substrate RNA by a unique mechanism. *Nat Chem Biol*. 2017. <https://doi.org/10.1038/nchembio.2328>.
55. Religa TL, Kay LE. Optimal methyl labeling for studies of supra-molecular systems. *J Biomol NMR*. 2010;47(3):163–9.
56. Audin MJ, Dorn G, Fromm SA, Reiss K, Schutz S, Vorlander MK, et al. The archaeal exosome: identification and quantification of site-specific motions that correlate with cap and RNA binding. *Angew Chem*. 2013;52(32):8312–6.
57. Religa TL, Sprangers R, Kay LE. Dynamic regulation of archaeal proteasome gate opening as studied by TROSY NMR. *Science*. 2010;328(5974):98–102.
58. Neu A, Neu U, Fuchs AL, Schlager B, Sprangers R. An excess of catalytically required motions inhibits the scavenger decapping enzyme. *Nat Chem Biol*. 2015;11(9):697–704.
59. Palmer 3rd AG, Kroenke CD, Loria JP. Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol*. 2001;339:204–38.
60. Korzhnev DM, Kloiber K, Kanelis V, Tugarinov V, Kay LE. Probing slow dynamics in high molecular weight proteins by methyl-TROSY NMR spectroscopy: application to a 723-residue enzyme. *J Am Chem Soc*. 2004;126(12):3964–73.
61. Skrynnikov NR, Mulder FA, Hon B, Dahlquist FW, Kay LE. Probing slow time scale dynamics at methyl-containing side chains in proteins by relaxation dispersion NMR measurements: application to methionine residues in a cavity mutant of T4 lysozyme. *J Am Chem Soc*. 2001;123(19):4556–66.

62. Lundstrom P, Vallurupalli P, Religa TL, Dahlquist FW, Kay LE. A single-quantum methyl  $^{13}\text{C}$ -relaxation dispersion experiment with improved sensitivity. *J Biomol NMR*. 2007;38(1):79–88.
63. Tugarinov V, Kay LE. Separating degenerate  $^1\text{H}$  transitions in methyl group probes for single-quantum  $^1\text{H}$ -CPMG relaxation dispersion NMR spectroscopy. *J Am Chem Soc*. 2007;129(30):9514–21.
64. Audin MJ, Wurm JP, Cvetkovic MA, Sprangers R. The oligomeric architecture of the archaeal exosome is important for processive and efficient RNA degradation. *Nucleic Acids Res*. 2016;44(6):2962–73.
65. Yuwen T, Vallurupalli P, Kay LE. Enhancing the sensitivity of CPMG relaxation dispersion to conformational exchange processes by multiple-quantum spectroscopy. *Angew Chem*. 2016;55(38):11490–4.
66. Lapinaite A, Simon B, Skjaerven L, Rakwalska-Bange M, Gabel F, Carlomagno T. The structure of the box C/D enzyme reveals regulation of RNA methylation. *Nature*. 2013;502(7472):519–23.
67. Huang R, Ripstein ZA, Augustyniak R, Lazniewski M, Ginalski K, Kay LE, et al. Unfolding the mechanism of the AAA+ unfoldase VAT by a combined cryo-EM, solution NMR study. *Proc Natl Acad Sci U S A*. 2016;113(29):E4190–9.