

Chapter 2

Isotope-Aided Methods for Biological NMR Spectroscopy: *Past, Present, and Future*

Masatsune Kainosho, Yohei Miyanoiri and Mitsuhiro Takeda

Abstract This chapter starts by providing a historical background of our research endeavors over the past half-century to develop various isotope-aided methods in biological NMR spectroscopy, since innovations bloom only on the rich ground cultivated by previous investigators. We then focused on the stereo-array isotope-labeling (SAIL) method, one of our recent accomplishments, which culminates the isotope-aided NMR technologies for structural studies of proteins from various aspects: accurate structural determinations of large proteins, elaboration for automated structural determination, highly efficient and versatile residue-selective methyl labeling with newly developed auxotrophic *E. coli* strains, large-amplitude slow-breathing motion (LASBM) as revealed by the aromatic ring flipping of the residues in ligand-binding interfaces, and applications of the deuterium-induced ^{13}C -NMR isotope shift to investigate the hydrogen exchange phenomena of side-chain polar groups. Meanwhile, the expected role of NMR spectroscopy has been rapidly shifting from structure determinations to dynamics studies of biologically interesting targets, such as membrane proteins and larger protein complexes. The dynamic aspects of protein–protein and protein–ligand interactions are closely related to their biological functions and can be efficiently studied by using proteins residue selectively labeled with amino acids bearing optimized labeling

M. Kainosho (✉)

Graduate School of Science and Engineering, Tokyo Metropolitan University,
1-1 Minami-ohsawa, Hachioji 192-0397, Tokyo, Japan
e-mail: kainosho@tmu.ac.jp

M. Kainosho · Y. Miyanoiri

Structural Biology Research Center, Graduate School of Science, Nagoya University,
Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan
e-mail: y-miyanoiri@protein.osaka-u.ac.jp

Y. Miyanoiri

Research Center for State-of-the-Art Functional Protein Analysis, Institute for Protein
Research, Osaka University, 3-2 Yamadaoka, Suita 565-0871, Osaka, Japan

M. Takeda

Department of Structural BioImaging, Faculty of Life Sciences, Kumamoto University,
5-1, Oe-honmachi, Chuo-ku, Kumamoto 862-0973, Japan
e-mail: takeda@structbiol.com

patterns, prepared by cellular expression. We are absolutely confident that biological NMR spectroscopy will continually develop with further innovations of isotope-labeling technologies in the coming era, featuring ultrahigh field spectrometers beyond 1 GHz.

Keywords Isotope-aided NMR method · ^{13}C , ^{15}N double labeling
SAIL method · Aromatic ring flipping · Stereo-specific methyl labeling
Hydrogen exchange rates of polar groups · Deuterium-induced ^{13}C isotope shift

2.1 Historical Background of Our Research to Develop Isotope-Aided NMR Methods

Over the past half-century, we have exploited isotope-aided NMR methods to investigate biological molecules, such as amino acids, peptides, proteins, and nucleic acids. Many of our results were initially presented at the annual meetings of the NMR Society of Japan, and some of our novel techniques are still being used worldwide. One of the prominent features of our approach is utilizing site- and stereo-specifically isotope-labeled amino acids and nucleosides, which can be efficiently prepared by combining microbial fermentations, enzymatic reactions, and chiral organic syntheses—all areas in which Japan has world-leading technological expertise. In this chapter, after a brief recollection of the early days in the development of isotope-aided biological NMR methods, we describe some of the past and current advances achieved mostly in our laboratory. However, because of the space limitation, the methods for studying nucleic acids are not included.

At the first annual symposium of the NMR Society of Japan, which was held in Tokyo, 1961, two papers on NMR studies of amino acids were presented. Namely, Fujiwara et al. reported the 56.4 MHz ^1H -NMR spectra of aqueous solutions of various amino acids and Takeuchi et al. reported the 40 MHz ^1H -NMR spectra of threonine and *allo*-threonine. To the best of our knowledge, these were the first biological NMR applications ever reported in Japan. At that time, there was no systematic NMR research on proteins, except for a short communication on ribonuclease A [1]. Unfortunately, the reported 40 MHz spectrum showed only four broad overlapped signals and did not provide detailed structural information. It was obvious that some groundbreaking methodologies were absolutely necessary to investigate proteins by NMR. Jardetzky proposed an excellent idea at the “International Symposium on Nuclear Magnetic Resonance,” which was held in Tokyo, 1965, aiming to settle the problem [2]. In this very exciting international meeting, gathering many of the eminent NMR pioneers, he explained a perspective of biosynthetic selective deuteration to simplify the ^1H -NMR spectra of proteins for obtaining structural information related to their biological functions. Surprisingly, it was only three years later when his colleagues published the first 100 MHz ^1H -NMR spectra of selectively deuterated nucleases, clearly showing that the

proposed strategy actually works well [3]. This enlightening work strongly motivated us to further develop isotope-labeling technologies for NMR studies of biological systems ever since [4].

2.1.1 Stereo-Specific Deuteration of Prochiral Methylene Protons—Conformational Analysis of Amino Acids and Peptides

Until the early 1970s, most biological NMR studies were still focused on small molecules, such as amino acids and oligopeptides. One of the crucial issues to be addressed in those days was the unambiguous discrimination of the side-chain β -methylene proton signals. Without the explicit stereo-specific assignments for the prochiral methylene protons, the two gauche conformations around the C_α - C_β bonds of amino acids, estimated by the Karplus relationship for the vicinal ^1H - ^1H coupling constants, could not be distinguished unambiguously. We solved this long-standing problem for the first time by using amino acids with one of the prochiral methylene protons stereo-specifically deuterated, prepared by either enzymatic reactions or organic syntheses. By using stereo-specifically deuterated

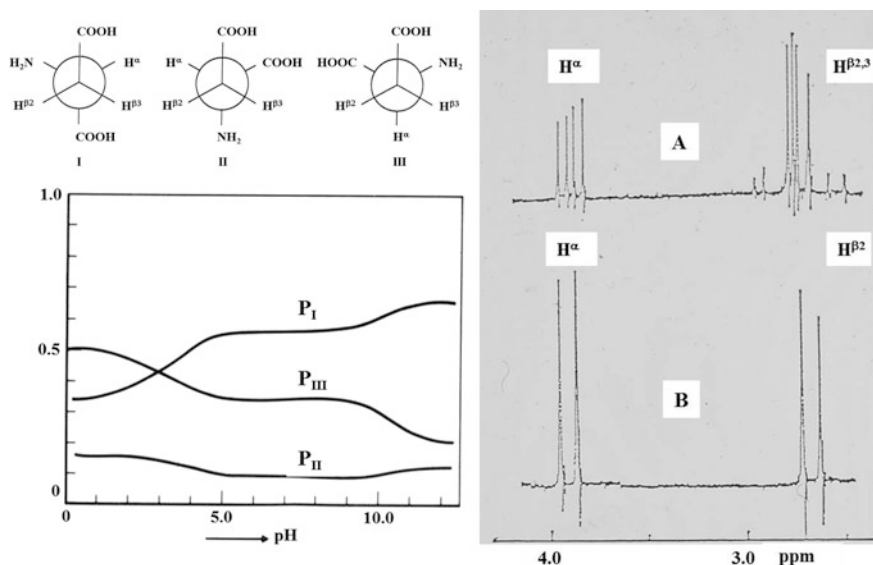


Fig. 2.1 Conformational analysis of aspartic acid using L-Asp and L-[β_3 -D]-Asp at various pHs. The relative populations of the three conformers were estimated at various pHs, by using the vicinal ^1H - ^1H spin coupling constants with the stereo-specifically assigned β -protons and α -proton. 100 MHz continuous-wave ^1H -NMR spectra: A L-Asp; B L-[β_3 -D]Asp under deuterium decoupling using the deuterium lock channel [6]

amino acids, we established the stereo-specific assignments for the prochiral β -methylene proton signals and thus discriminated the two gauche conformations for various amino acids and small peptides [5, 6] (Fig. 2.1). Despite our efforts, the approach along this line could not be extended for studying larger peptides or proteins, since no practical experimental methods to incorporate stereo-specifically deuterated amino acids into peptides or proteins were established at that time. As a matter of fact, more than 30 years after our early work, the idea blossomed into the stereo-array isotope-labeling (SAIL) method with the help of two emerging key technologies: chiral organic synthesis and recombinant DNA methods for protein expression [4].

2.1.2 Selective ^{13}C , ^{15}N Double-Labeling Method for the Sequential Assignment of Backbone Amide NMR Signals in Large Proteins

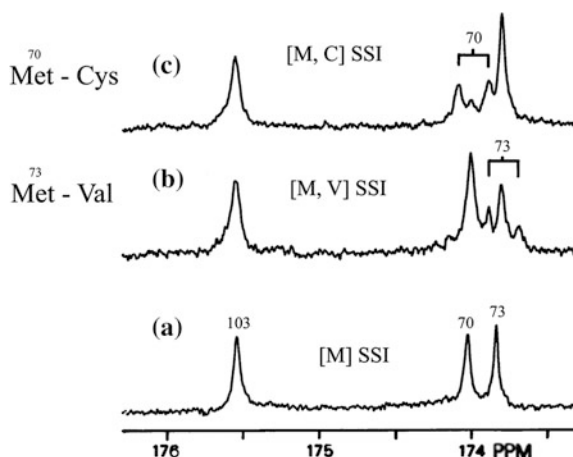
During the 1970s, there was a multidisciplinary collaborative research group, known as the “Research Consortium on *Streptomyces* Subtilisin Inhibitor (SSI),” which focused on SSI as a shared target, aiming to promote biophysical and biochemical protein research activities in Japan. SSI, which was isolated from the culture broth of *Streptomyces albobriseolus*, is a 23-kDa dimeric protein composed of two identical subunits, and it strongly inhibits serine proteinases, especially subtilisin family proteinases. One of the controversial issues for proteinase–inhibitor interactions in the late 1970s was the state of the active site peptide bond, i.e., the “scissile bond,” in the inhibitors complexed with proteinases. X-ray crystallographic analyses of various proteinase–inhibitor complexes initially concluded the existence of the “tetrahedral intermediate,” which was thought to be formed by a covalent bond between the active site Ser O_γ of a proteinase and the carbonyl of a scissile bond. It was assumed that the tetrahedral intermediate was trapped due to the overwhelming stabilization by the “oxy-anion hole” of serine proteinases, which serves as a molecular device to enormously accelerate the enzymatic hydrolysis of substrate peptides. However, this remarkable model, which was cited in most of the biochemistry textbooks at that time to highlight the beauty of the enzymatic functions of proteinases, was becoming dubious. Namely, as higher-resolution X-ray structures became available, they revealed that the atomic distance between the Ser O_γ and the carbonyl carbon atom of the scissile bond was apparently a little too long to form the covalent bond. In addition to this serious concern, others still remained about the state of the proteinase–inhibitor complex, which could not be solved by crystallographic studies, such as the following: Does it exist in solution as a single intermediate or as an equilibrium mixture of multiple intermediates? Is the scissile peptide bond in the complexes planar, as usually found for peptide bonds, or distorted by the effect of the nearby Ser O_γ ? In principle, all of these questions could be investigated by solution NMR spectroscopy. However, it was difficult to

think that such work was feasible since the molecular weight of the SSI–subtilisin complex, 78 kDa, was too large. We overcame this problem by a unique isotope-labeling strategy, as described below [7].

Since the scissile peptide bond of SSI is formed between Met⁷³ and Val⁷⁴, its state in the proteinase complex would be precisely manifested by the ¹³C-NMR signal for the carbonyl carbon of Met⁷³, if we could observe a single ¹³C-NMR signal for the 78-kDa protein. Although each of the two identical SSI subunits contains three Met residues, i.e., Met⁷⁰, Met⁷³, and Met¹⁰³, their C-terminal neighbors are all different, i.e., Cys⁷¹, Val⁷⁴, and Asn¹⁰⁴, respectively. Taking advantage of this sequential diversity, we contrived the ¹³C, ¹⁵N double-labeling method to sequentially assign the three carbonyl carbon signals. By using SSI samples doubly labeled with [1-¹³C]Met and [¹⁵N]Cys, and with [1-¹³C]Met and [¹⁵N]Val, we should be able to discriminate the carbonyl ¹³C signals of Met⁷⁰ and Met⁷³ through the spin couplings between the directly bonded ¹³C and ¹⁵N, which are known to be about 15 Hz. As illustrated in Fig. 2.2, this strategy based solely on the amino acid sequence information worked perfectly [7]. In addition to the unambiguous assignment, this method provided even more crucial information about the state of the scissile bond in the SSI–subtilisin complex, through the ¹³C chemical shift of Met⁷³ and the ¹³C–¹⁵N spin coupling values between Met⁷³ and Val⁷⁴. We finally proved that the Michaelis complex with the intact, undistorted scissile bond is the only stable form of the SSI–subtilisin complex in solution [8, 9].

The idea of the sequential assignment for the backbone carbonyl carbons by the selective ¹³C, ¹⁵N double-labeling method was extended to assign the peptide ¹⁵N and side-chain ¹³C signals, through the ¹³C–¹⁵N and ¹³C–¹³C connectivities [10, 11]. The method can be regarded as the historic prototype of triple-resonance sequential assignment methods, using uniformly ¹³C, ¹⁵N double-labeled proteins. We were actually one of the first groups to suggest the idea of establishing sequential assignment methods by extending the ¹³C, ¹⁵N double-labeling method, which was presented at the XI ICMRBS in Goa, India, in 1984 [12].

Fig. 2.2 Sequence-specific assignment of the carbonyl ¹³C signals for the three Met residues in *Streptomyces* subtilisin inhibitor (SSI) by the selective ¹³C, ¹⁵N double-labeling method. 75 MHz ¹³C-NMR spectra of SSI labeled with **a** [1-¹³C]Met; **b** [1-¹³C]Met/[¹⁵N]Val; **c** [1-¹³C]Met/[¹⁵N]Cys [7]



2.1.3 Revisiting the Stereo-Specific Isotope-Labeling Approach for Studying Proteins: A Long March to the SAIL Method

In 1997, with financial support from the newly launched grant “Core Research and Evolutional Science and Technology,” also known as CREST, aiming to promote basic sciences in Japan, we started a 5-year project to develop a breakthrough isotope-aided NMR technology for studying large proteins. At that time, NMR technologies using uniformly ^{13}C , ^{15}N double-labeled proteins were firmly established, but they could only be applied to determine the three-dimensional structures of small proteins. However, a variety of key technologies, which were not available back in the 1970s, facilitated further innovations of isotope-aided NMR methods. For example, multinuclear multidimensional NMR spectroscopy, chiral organic syntheses, and protein expression using recombinant DNA techniques were all quite mature techniques by then. Therefore, we had a unique opportunity to revisit the old idea to explore cutting-edge methods for studying larger proteins. In order to encompass the advents of various multidimensional NMR methods, we exploited novel synthetic routes for the regio- and stereo-specifically D, ^{13}C , ^{15}N triple-labeled amino acids. With the help of state-of-the-art chiral organic synthetic methods, together with enzymatic reactions and microbial fermentations, we actually successfully synthesized all of the protein component amino acids with a variety of labeling patterns [13–15]. Fortunately, with further support from a second CREST grant for another five years, we completed the development of the “stereo-array isotope-labeling (SAIL) method” by using those labeled amino acids. Although it took more than 30 years after our early work on the stereo-specific deuteration of amino acids and peptides, the SAIL method has been proven to be extremely useful for studying the structures as well as the dynamics of larger proteins, for which previous NMR methods were difficult to apply [16–19].

2.2 The SAIL Method: An Optimized Isotope-Labeling Strategy for the Structural Study of Proteins by NMR Spectroscopy

NMR spectra of larger proteins are typically characterized by numerous overlapped signals, which are severely broadened by dipolar interactions between nearby protons. Therefore, it was difficult to obtain sufficient amounts of NMR information for proteins larger than 20–25 kDa, even with sophisticated multidimensional methods. For example, for a long time it was thought to be virtually impossible to analyze the prochiral methylene proton signals, especially for large proteins, even

though the information is absolutely required for accurately determining the side-chain conformations, as described above for amino acids [5, 6]. Conceptually, however, we may not necessarily need all of the NMR data for the determination of protein structures, since many of the amino acid side chains contain somewhat redundant information. For example, if we could stereo-specifically observe either one of the prochiral groups, i.e., methyls and methylenes, we could compensate for the missing information about the geminal counterparts. The SAIL method creates this type of situation for all of the amino acid residues in proteins, by trimming away the redundant information by the optimized isotope-labeling patterns, as described below [20, 21].

Our original labeling design concepts for the SAIL amino acids were the following: (1) stereo-specific labeling for one of the methylene protons by deuterium, i.e., $^{13}\text{C}^*\text{HD}$; (2) stereo-specific labeling for the geminal methyl groups of Leu and Val residues with $^{13}\text{CHD}_2$ and $^{12}\text{CD}_3$; (3) alternating the labeling of aromatic rings with ^{13}CH and ^{12}CD ; and (4) ^{13}C labeling for the methine groups of Ile, Leu, and Val. Therefore, all of the methyl, methylene, and methine groups labeled with ^{13}C have a single proton, i.e., $^{13}\text{CHD}_2$, $^{13}\text{C}^*\text{HD}$, and ^{13}CH . In addition to these labeling patterns, all of the nitrogens and the carbonyl carbons are labeled with ^{15}N and ^{13}C , respectively. We successfully synthesized the 20 protein component amino acids based on these design concepts, as shown in Fig. 2.3 [20]. The SAIL labeling patterns preserve the through-bond ^{13}C - ^{13}C and ^{13}C - ^{15}N connectivity paths for the backbone and side-chain sequential assignments and completely eliminate the ambiguity of stereo-specific assignments for prochiral groups. More importantly, the density of the remaining protons attached to the ^{13}C atoms in a SAIL protein, which is exclusively composed of the SAIL amino acids, is reduced down to 50–60% as compared to that of a fully protonated protein. The significantly decreased proton density for a SAIL protein mitigates the spin diffusion and thus facilitates the acquisition of accurate inter-proton NOEs, even for larger proteins. One might recollect at this point that a traditional strategy using “random fractional deuteration” was used a while ago, to reduce the NMR line widths of the remaining proton signals [22]. However, the random fractional deuteration results in an enormous number of isotopomer proteins with chemical shift heterogeneities and the concomitant loss of intensities for the remaining signals. All of these problems are completely eliminated by the SAIL method. It is quite important to emphasize that, even though the level of deuteration is as high as that of random fractionally deuterated proteins, there is always a single isotopomer for a SAIL protein. Therefore, the proton concentrations for the protonated sites are not decreased at all. Actually, the terminology “stereo-array isotope labeling” seeks to highlight this striking feature of the SAIL method.

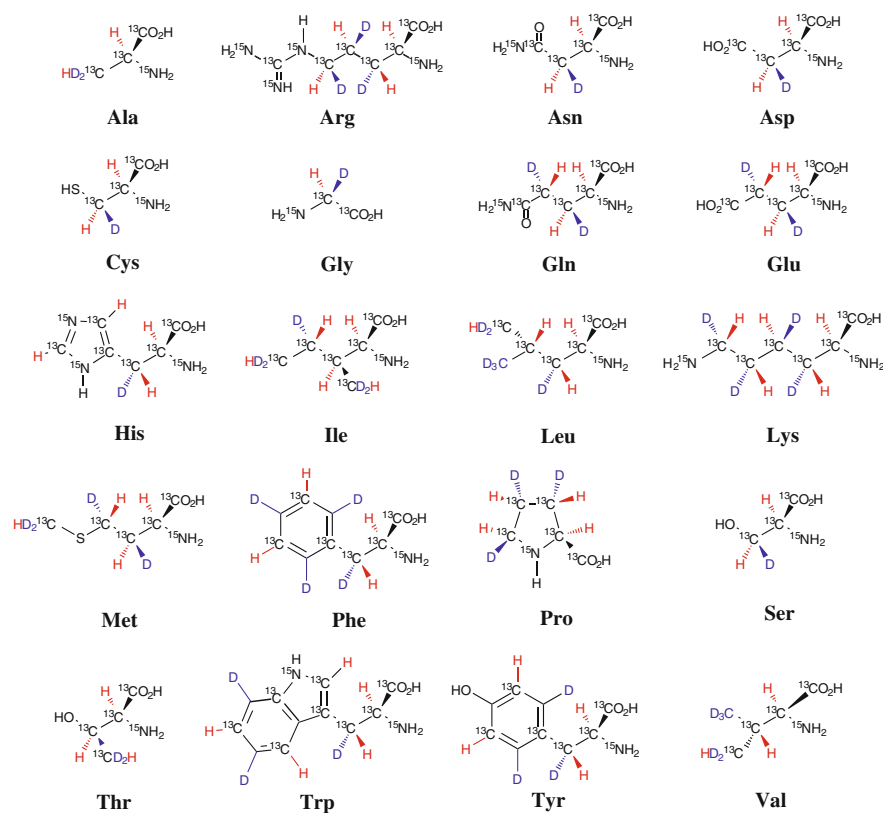


Fig. 2.3 Structures of the SAIL amino acids with the original isotope-labeling patterns [20]. There are many other SAIL and SAIL-related amino acids with various labeling patterns optimized for obtaining specific information, as described in this chapter

2.2.1 Cell-Free Expression and NMR Spectra of SAIL Proteins

As described above, the SAIL method facilitates structural analyses of larger proteins, through the rational isotope-labeling design for the component amino acid residues. Therefore, it is necessary to incorporate SAIL amino acids into a target protein while preserving their original labeling patterns. In this respect, conventional cellular protein expression using recombinant DNA may not be a good choice, since metabolic scrambling reactions and isotope dilution with unlabeled amino acids are not completely avoidable for some amino acid residues. All these problems are largely circumvented by using *in vitro* expression systems. Fortunately for us, the *in vitro* protein expression using the *E. coli* cell-free extract

became available at around the time we nearly completed the synthesis of the SAIL amino acids. By using the cell-free extract prepared from the *E. coli* cells, we successfully prepared a sufficient amount of a SAIL protein for an NMR study. SAIL proteins, which are composed exclusively of SAIL amino acids, are typically obtained with ~ 10 wt% yields calculated from the amount of the amino acid mixture and show virtually no isotopic dilution or metabolic scrambling [23–25].

The NMR spectra of SAIL proteins actually exhibit profoundly better sensitivity and resolution, as compared to those of conventional ^{13}C , ^{15}N uniformly labeled (UL) proteins. Figure 2.4 shows such superior NMR features for the case of 17-kDa SAIL calmodulin (SAIL-CaM) [20]. By comparing the ^1H - ^{13}C ct-HSQC spectra for SAIL- and UL-CaM, it becomes obvious that the ^1H - ^{13}C cross-peaks are considerably sharper and well dispersed for SAIL-CaM (Fig. 2.4a) than for UL-CaM (Fig. 2.4d). A closer comparison of the methyl (Fig. 2.4b, e) and methylene regions (Fig. 2.4c, f) of the two proteins reveals that only one of the prochiral pairs shows up as an NMR signal in SAIL-CaM. The benefit of the universal stereo-specific deuteration for the methylene groups in the SAIL method can be illustrated for the δ -methylene signals of the six Arg residues in SAIL- and UL-CaM (Fig. 2.4g–i). Since the δ_3 methylene proton is stereo-specifically deuterated in SAIL-CaM, the six observed δ -CH signals can be unambiguously assigned to δ_2 -CH (Fig. 2.4g), while in UL-CaM, most of the Arg δ -CH₂ signals are completely overlapped for both the ^1H and ^{13}C dimensions, except for Arg-37 and 106, for which no facile stereo-specific assignment methods could be envisaged (Fig. 2.4h). Since similar problems are ubiquitously encountered in UL proteins for the methylene groups in other amino acid residues, only SAIL proteins provide accurate structural information for the side-chain moieties. The complete spectral analyses of SAIL proteins are usually quite straightforward, by using the standard triple-resonance pulse sequences with slight modifications [15].

The analyses of aromatic ring signals in proteins are generally very cumbersome or infeasible for UL proteins, due to their complex ^{13}C , ^1H spin systems. The situation becomes even more complicated when the flipping rates of aromatic rings influence the signal shapes. Nevertheless, the information about the bulky aromatic rings, which tend to be embedded in the hydrophobic cores, is extremely important for structural studies of proteins. Therefore, we expended a great deal of effort toward optimizing the isotope-labeling patterns of aromatic amino acids, to obtain all of the structural information for aromatic rings in straightforward manners [26–28]. Figure 2.5 illustrates the application of SAIL Phe with various labeling patterns, aiming toward analyses of the ring ^1H - ^{13}C signals, as exemplified by the assignment of the 12 Phe residues in *E. coli* peptidyl prolyl cis-trans isomerase b (EPPiB) [27]. The aromatic ^1H - ^{13}C region of the ct-HSQC spectrum for EPPiB labeled with UL Phe (Fig. 2.5d) gives virtually no detailed information, due to the overcomplicated spin systems. In contrast, each of the HSQC spectra for EPPiBs, selectively labeled with δ -, ϵ -, or ζ -SAIL Phe, shows clean, well-resolved signals (Fig. 2.5e–g). The δ -, ϵ -, and ζ -CH could be readily assigned by the NOE

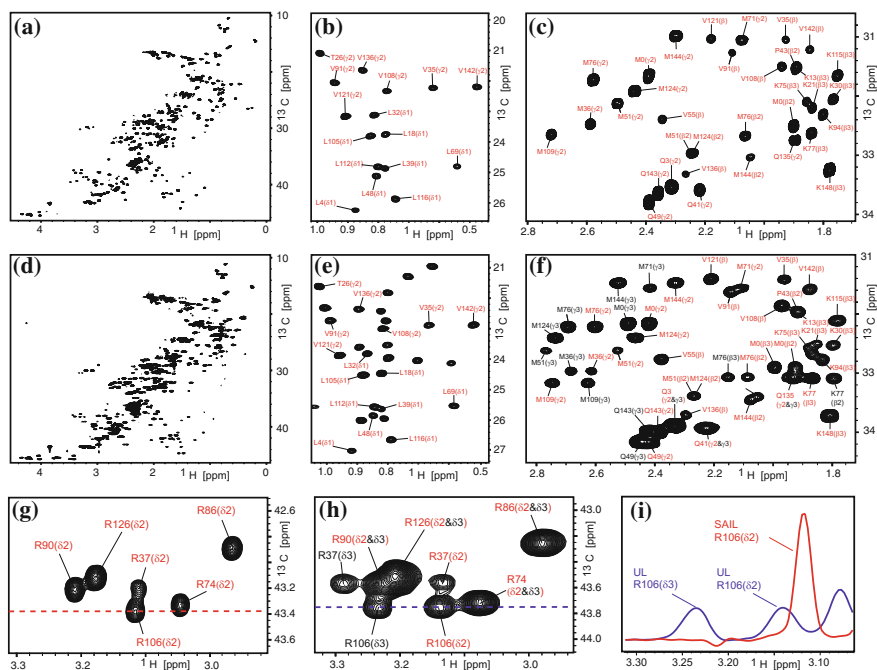


Fig. 2.4 800 MHz ^1H - ^{13}C ct-HSQC spectra of calmodulin (CaM). **a** SAIL-CaM, aliphatic region; **b** SAIL-CaM, methyl region; **c** SAIL-CaM, methylene region; **d** UL-CaM, aliphatic region; **e** UL-CaM, methyl region; **f** UL-CaM, methylene region; **g** SAIL-CaM, Arg δ region; **h** UL-CaM, Arg δ region; **i** Cross sections from (**g**) and (**h**). The spectra for SAIL-CaM and UL-CaM were recorded under identical conditions and scaled for equal noise levels [20]

correlations between the δ - ^1H and β_3 - ^1H , HBCB(CG)HE, and HBCB(CGZ)HZ sequences, respectively (Fig. 2.5a-c). The most significant advantage of the SAIL Phe can be realized by comparing the spectra between the ζ -SAIL-labeled (Fig. 2.5g) and the δ -, ε -SAIL-labeled EPPIb (Fig. 2.5e, f). The ζ -CH signals in ζ -SAIL Phe-labeled EPPIb have line shapes that are invariant against the flipping rates, because the ζ -CH bond lies along the C_β - C_γ bond, and thus they appeared as 12 discrete signals with almost the same intensities (Fig. 2.5g). In contrast, in the δ - and ε -SAIL Phe-labeled EPPIb, the ^1H - ^{13}C HSQC spectra show very weak or no visible signals for F27, F110, and F123 (Fig. 2.5e, f). The results clearly indicated that the ring-flipping rates for these three residues are relatively slow, as compared to the chemical shift differences for the two δ - and ε -CH atoms on the opposite side of the ring. As the aromatic ring-flipping motions are perfectly degenerated conformational exchange phenomena, which can only be manifested by the NMR line-shape analysis, these SAIL aromatic amino acids are extremely valuable for studying large-amplitude dynamics in proteins (see Sect. 2.3.2).

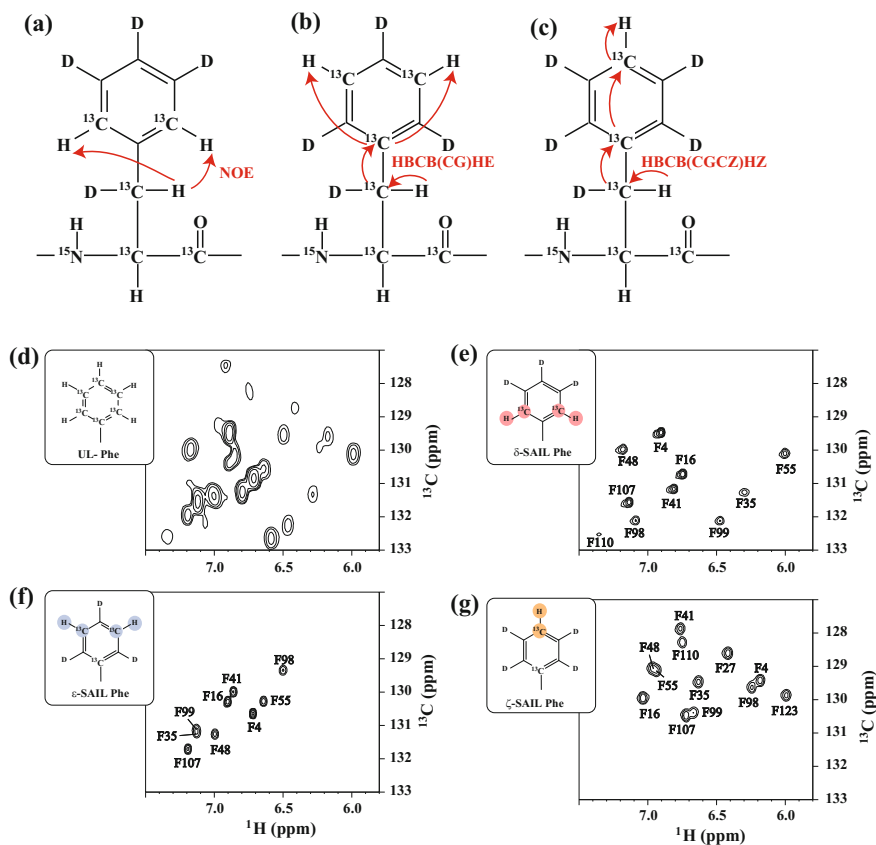


Fig. 2.5 Assignment of aromatic ring CH signals of the 12 Phe residues in EPPiB using various SAIL Phe. Structures of **a** δ -SAIL Phe, **b** ϵ -SAIL Phe and **c** ζ -SAIL Phe. The *arrows* indicate the magnetization transfer pathways for making the ring CH signal assignments, starting from the β -proton that can be assigned by the conventional sequential assignment protocols. 600 MHz ^1H - ^{13}C HSQC spectra with the assignments of EPPiBs specifically labeled with **d** [U- ^{13}C]Phe; **e** δ -SAIL Phe; **f** ϵ -SAIL Phe; and **g** ζ -SAIL Phe [27]

2.2.2 Structural Determination of SAIL Proteins

As described above, SAIL proteins provide unmatched chemical shift data quality for the backbone and aliphatic/aromatic side chains, including the full stereo-specific assignments for prochiral groups. Therefore, SAIL proteins facilitate the automated assignment of the ^{13}C - and ^{15}N -edited NOESY spectra, with the concurrent structure refinement using the CYANA program adapted to SAIL proteins (SAIL-CYANA). It is important to mention that the substantially reduced

proton density in SAIL proteins allows the use of a longer mixing time for collecting long-range NOE constraints, without spin diffusion problems. Taking advantage of the improved distance constraints and the unequivocal stereo-specific assignments for prochiral groups, the structures determined for SAIL proteins are accurate, as illustrated in Fig. 2.6 for some of the structures determined by the SAIL-CYANA method [20, 21].

The marked improvement in the overall quality of the NMR spectra obtained for SAIL proteins further encouraged us to use the FLYA program for automated backbone and side-chain resonance assignments. The chemical shift data obtained by FLYA can then be used as the input data for the NOESY spectral analysis and the structure calculation by CYANA. This two-step automated structure determination using the FLYA-CYANA program works well for small SAIL proteins, without additional human participation [29, 30]. We also tried to develop a fully automated structure determination method exclusively based on NOESY data, obviating the need to measure any other spectra than those necessary for the

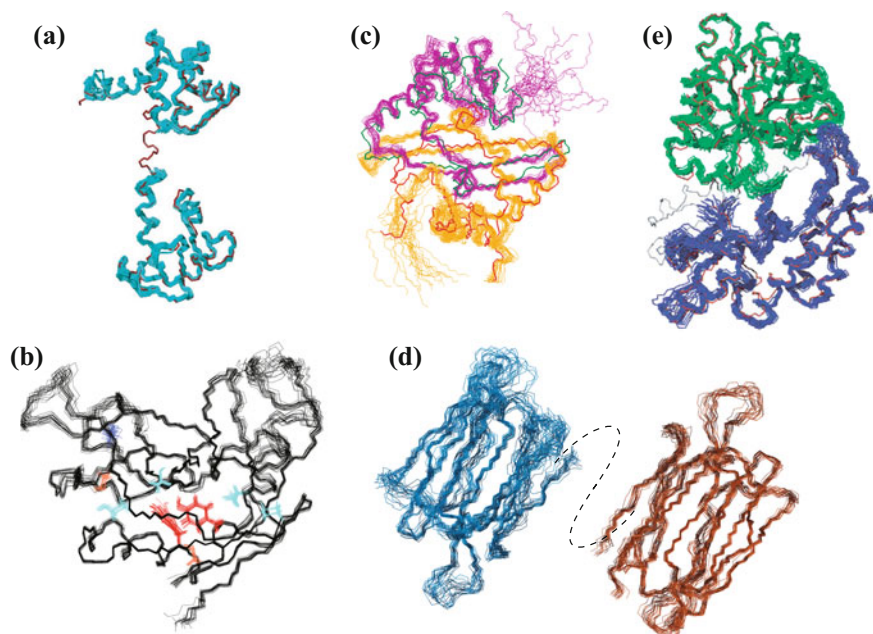


Fig. 2.6 Backbone structures of proteins determined by the SAIL-NMR method. **a** Calmodulin in cyan, overlaid with the X-ray structure in red [20]; **b** EPP1b in black, which was refined using additional NOEs involving slowly exchanging side-chain polar groups: the side chains of Ser, Thr in cyan, Cys in pink and Tyr in red [48]; **c** C-terminal dimerization domain of SARS coronavirus nucleocapsid protein. The two subunits of the homodimer are shown in pink and yellow, and the overlaid X-ray structures are shown in green and red, respectively [51]; **d** the putative protein At3g16450.1, encoded by the *Arabidopsis thaliana* gene, composed of the N-terminal domain in cyan and the C-terminal domain in brown [52]; **e** maltose-binding protein (MBP): the N-domain in cyan and C-domain in blue, overlaid with the X-ray structure in red [20]

resonance assignment. This ambitious automated approach, which would be useful for determining a large number of structures as efficiently as possible, was actually utilized for two small SAIL proteins and yielded well-defined structures that coincide closely with those determined by the conventional method [31].

2.3 Recent Trends in the Isotope-Aided NMR Methods for Studying Proteins

We described above our early studies on isotope-aided NMR techniques and then introduced a recent achievement, the SAIL method. However, the prospective roles of NMR in structural biology are rapidly changing, especially because other methods, such as X-ray crystallography and cryo-electron microscopy, have been overwhelmingly employed for the structural determinations of biologically important proteins, such as membrane proteins or extraordinarily large protein complexes. Obviously, NMR cannot be a competitive structural determination tool for those targets. Instead, a variety of alternative applications are envisaged for NMR spectroscopy, to bridge the gap between protein structures and their biological functions. In principle, NMR could afford unique information for this purpose, even if the proteins are too large for structure determination by NMR. Actually, in many cases, one could start with the three-dimensional structures previously determined by the other methods and focus on the structures and dynamics of the selected regions of interest, which could be precisely manifested by NMR. For that purpose, it is necessary to develop a method to observe and assign the NMR signals for any regions of the selected amino acid residues in such proteins. Recently, there have been major advances in NMR signal observations for larger proteins. Wuethrich et al. exploited transverse relaxation-optimized spectroscopy (TROSY) for observing the backbone amide $^1\text{H}^{15}\text{N}$ signals in deuterated proteins and the aromatic ring $^1\text{H}^{13}\text{C}$ signals for uniformly ^{13}C -labeled proteins [32, 33]. Kay et al. developed a method to observe the Ile, Leu, and Val (ILV) $^{13}\text{CH}_3$ signals, utilizing them as NMR probes for studying protein structures and dynamics [34]. The $^1\text{H}^{15}\text{N}$ TROSY and methyl observation methods can be applied for proteins as large as 1 MDa and are routinely used for studying larger proteins in solution [35].

Since the backbone amides and the side-chain methyl groups ($^{13}\text{CH}_3$) cover considerable portions of larger proteins, their NMR signals provide valuable structural information. However, it may not be sufficient for analyzing the precise side-chain conformations and dynamics for the selected residues in order to understand the molecular basis of biological functions, which we expect to obtain by solution NMR. This is where sophisticated isotope-aided methods such as SAIL come in. We have exploited methods to observe NMR signals for any parts of

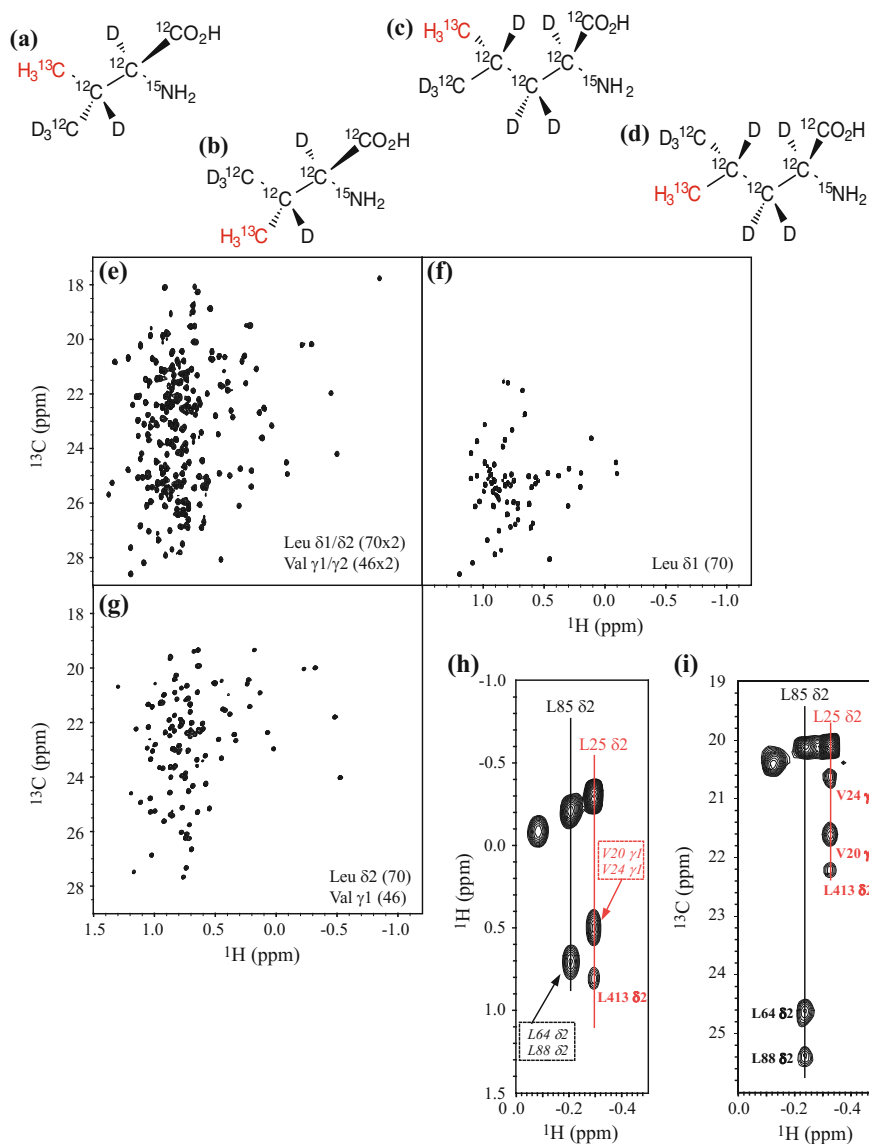
aliphatic and aromatic side chains in a protein, by further optimizing the isotope-labeling patterns of the original SAIL design concepts. In the following, we describe some of our recent work along this line, in order to provide an outline of our current research endeavors.

2.3.1 Residue- and Stereo-Specific Labeling Method: The Case for Leu and Val Methyl Labeling of Larger Proteins

Larger proteins have numerous methyl groups in their Ala, Thr, Met, Ile, Leu, and Val (ATMILV) residues, which are widely distributed on their surface and interior regions. Therefore, the methyl signals are valuable probes for studying the structures and dynamics of proteins and protein complexes, if the individual methyl signals could be observed. However, this is not trivial especially for larger proteins, since they have so many ATMILV residues. For example, the 82-kDa protein malate synthase G (MSG) has 289 ATMILV residues that comprise approximately 40% of its 723 residues. Among them, the 160 ILV residues are especially useful as NMR probes, since the total of 317 ILV methyls accounts for as many as $\sim 70\%$ of the 446 methyl signals. Therefore, extensive efforts have been exerted to develop robust methods to observe and assign ILV methyl signals in larger proteins. Most of them employ region-specifically ^{13}C , D-labeled precursors, such as [4- ^{13}C ;3,3- D_2]- α -ketobutyrate and [3- $^{13}\text{CH}_3$;3,4,4,4- D_4]- α -ketoisovalerate (α -KIV), for preparing fully deuterated proteins except for the Ile (δ_1), Val (γ_1/γ_2), and Leu (δ_1/δ_2) methyls, which are to be labeled with $^{13}\text{CH}_3$ [36, 37]. However, since the racemic α -KIV precursor labels both of the prochiral methyls in Leu and Val residues, the observable number of methyl signals cannot be decreased, and the labeling rates are 50% or less. Therefore, the signal congestion for Leu/Val methyls could not be improved and the methyl–methyl NOEs are significantly reduced. In order to compensate for the drawbacks of this Leu/Val precursor, a few new precursors for the stereo-specific methyl labeling of Leu/Val residues have been developed [38]. However, it is difficult to use labeled precursors to prepare any desired combinations for either one of the prochiral methyls in Val and Leu, since Val is converted to Leu biosynthetically, as shown in the metabolic map.

In the SAIL method, we synthesized SAIL Leu/Val, in which one of the prochiral methyls is stereo-specifically labeled with ^{13}CHD and the other one with CD_3 . The SAIL Val/Leu can be incorporated into proteins by the *in vitro* expression system, using the *E. coli* cell-free extract (vide supra). Using the same protocols, we synthesized all four of the Leu and Val residues, in which either one of δ_1/δ_2 or γ_1/γ_2 in Leu and Val, respectively, is stereo-specifically labeled with $^{13}\text{CH}_3$, and all of the other protons are fully deuterated, as shown in Fig. 2.7 [39]. These four

stereo-specifically methyl-labeled Val and Leu, designated as γ_1 -Val (a), γ_2 -Val (b), δ_1 -Leu (c), and δ_2 -Leu (d), could be incorporated into proteins in any combinations by the conventional cellular expression using the *E. coli* BL21 (DE3) strain. The incorporation rate for the labeled Leu into MSG at a 20 mg/L concentration was $\sim 90\%$, but that for the labeled Val at 100 mg/L was only close to 80%. In order to increase the incorporation rates for Leu and Val at lower amino acid concentrations, we have recently developed an auxotrophic *E. coli* BL21 (DE3) strain [40]. The mutant was derived from the BL21 (DE3) strain by deleting the *ilvD* and *leuB* genes encoding dihydroxy acid dehydratase and β -isopropylmalate dehydrogenase, respectively; therefore, it cannot survive in the absence of Ile, Leu, and Val. Using this auxotrophic *E. coli* strain, the incorporation rates of isotope-labeled Leu, Val, and also Ile into MSG were found to be higher than 95% even at a 10 mg/L concentration of each of the stereo-specifically methyl-labeled Leu or Val, or the region-specifically methyl-labeled Ile, without any observable scrambling [41]. The usefulness of the stereo-specifically methyl-labeled Leu and Val is clearly illustrated in Fig. 2.7e–g. Figure 2.7e shows the methyl region of the ^1H - ^{13}C HMQC spectrum of deuterated MSG labeled with a conventional precursor, $[3\text{-}^{13}\text{CH}_3; 3,4,4,4\text{-D}_4]\text{-}\alpha\text{-ketoisovalerate}$. The spectrum showed 232 considerably overlapped methyl signals, including the δ_1 , δ_2 methyls of 70 Leu residues (including one extra Leu in the His tag at the C-terminus) and the γ_1 , γ_2 methyls of 46 Val residues. In contrast, the spectra in Fig. 2.7f, g for deuterated MSG labeled by δ_1 -Leu or δ_2 -Leu + γ_1 -Val, respectively, show almost no signal overlap. It may not be apparent from the spectra, but the methyl signals observed for the MSGs labeled by δ_1 -Leu, δ_2 -Leu + γ_1 -Val showed increased sensitivities, as compared to the MSG labeled by the α -KIV precursor, by a factor of 2. It is important to mention that any single-residue-labeled MSGs and also any combinatorially dual-residue-labeled MSGs can be prepared by using the four different stereo-specifically methyl-labeled Leus and Vals. The combinatorial methyl-labeling method using stereo-specifically isotope-labeled Leu and Val is especially useful for collecting the inter-residue methyl–methyl NOEs at higher sensitivities, by a factor of 4, as compared to the precursor method. Actually, the Val γ_1 and Leu δ_2 combinatorially labeled MSG gave highly sensitive inter-residue methyl–methyl NOE signals. Even if the signal overlap remains in the 2D ^1H - ^1H plane of the 3D ^{13}C -edited NOESY-HMQC spectrum for a combinatorially labeled protein, as illustrated in Fig. 2.7h, the 3D HMQC-NOESY-HMQC should have better resolution, as shown for the 2D ^1H - ^{13}C plane, making good use of the wider dispersion in the ^{13}C dimension (Fig. 2.7i). Obviously, further extensions of the combinatorial labeling method involving multiple regio- and stereo-specifically isotope-labeled amino acid residues would eventually approach the concept of the SAIL method.



2.3.2 Large-Amplitude Dynamics of Proteins as Probed by Aromatic Ring-Flipping Motions—The Case for the Interface Between FKBP and Drug Complexes

Nowadays, it is generally accepted that folded proteins occasionally undergo large-amplitude slow-breathing motions (LASBMs) under physiological

◀**Fig. 2.7** Residue- and stereo-specific isotope labeling for the Leu and Val methyl groups in the 82-kDa protein malate synthase G (MSG) using stereo-specifically methyl-labeled amino acids. Structures of the stereo-specifically ^{13}C -labeled, otherwise uniformly deuterated, valines and leucines: **a** [γ_1 - ^{13}C ; α - ^{15}N ; D_5]-valine, “ γ_1 -Val””; **b** [γ_2 - ^{13}C ; α - ^{15}N ; D_5]-valine, “ γ_2 -Val””; **c** [δ_1 - ^{13}C ; α - ^{15}N ; D_7]-leucine, “ δ_1 -Leu””; **d** [δ_2 - ^{13}C ; α - ^{15}N ; D_7]-leucine, “ δ_2 -Leu.” 900 MHz 2D ^1H - ^{13}C HMQC spectra of labeled MSGs: **e** Leu/Val selectively labeled MSG expressed by *E. coli* BL21 (DE3) using the precursor [3 - ^{13}C ;3,4,4,4- D_4]- α -ketoisovalerate; **f** Leu-specific, δ_1 -methyl-specific-labeled MSG prepared by the *ΔilvD/ΔleuB E. coli* mutant using “ δ_1 -Leu” and deuterated Ile/Val; **g** γ_1/δ_2 - ^{13}C -labeled MSG prepared by the *ΔilvD/ΔleuB E. coli* mutant using “ γ_1 -Val,” “ δ_2 -Leu” and deuterated Ile. Shown in **(h)** and **(i)** are the 2D planes at the ^{13}C chemical shift of 20.1 ppm, which corresponds to the δ_2 methyls of L25 and L85, of the 3D ^{13}C -edited NOESY-HMQC and 3D HMQC-NOESY-HMQC spectra, respectively, measured for a 0.2-mM solution of the γ_1/δ_2 - ^{13}C specifically labeled, otherwise fully deuterated MSG. Ambiguous NOE signal assignments are labeled in *italics* [40]

conditions. Since the frequency of LASBMs is within the millisecond to second region, such motions have attracted the interests of biophysicists in the context of biological functions and protein dynamics. The LASBM was initially implicated by the intriguing observations that the δ - and ϵ -protons for the Phe and Tyr aromatic rings in proteins showed time-averaged NMR signals. It was quite surprising, especially for most of the crystallographers back in the 1970s, that such bulky aromatic rings flip about the C_β - C_γ axis so frequently, since they are often deeply embedded in the hydrophobic core, which was thought to be the most solid part of a protein [41, 42]. Ironically, until recently there have only been few cases in which the aromatic rings show discrete signals for the δ - and ϵ -nuclei of Phe and Tyr, due to slow ring-flipping rates [43]. Theoretically, it might be possible that the ^1H and ^{13}C nuclei at the δ - and ϵ -positions incidentally have identical, or nearly identical, chemical shifts and thus appear as a time-averaged single peak regardless of the flipping rates, although it is quite unlikely that such situations happen very often. Taking advantage of the simplified spin systems of the SAIL Phe and Tyr, we revisited the ring-flipping phenomena and found that there are actually many more cases showing flipping rate-dependent aromatic ring signals. Apparently, the aromatic rings in conventional protein samples have such complicated spin networks that such cases are rarely identified. Therefore, proteins selectively labeled by δ -, ϵ -SAIL Phe and Tyr would provide unprecedented opportunities to investigate LASBMs through the widely distributed aromatic rings in the hydrophobic interior and on the ligand-binding surface. In the following, we illustrate the application of the aromatic ring-flipping phenomena for characterizing the LASBMs within the binding interface in FKBP12–ligand complexes [44].

The tight complexes FKBP12 forms with immunosuppressive drugs, such as FK506 and rapamycin, have long been used as models for developing various approaches to structure-based drug design. The regions of rapamycin and FK506 that bind FKBP are very similar to each other, but the opposite sides of them, which are referred to as the effector regions, are entirely different for the two drugs. Rapamycin and FK506 can bind to their targets, mTOR and calcineurin, only if they are complexed with FKBP12. Therefore, it is very interesting to understand the

molecular mechanism by which FKBP12 activates these drugs to trigger their diversified biological functions by forming the ternary complexes. The aromatic ring cluster in FKBP12 forms an extremely hydrophobic, concave binding pocket, composed of Trp59, Tyr26, Phe99, Phe46, and Phe48, for the immunosuppressive drugs, such as rapamycin or FK506, with high affinities (Fig. 2.8). Although the interfaces between FKBP12 and these drugs are well defined structurally and are almost identical in the crystallographic structures of various complexes, our NMR studies have clearly revealed the existence of substantial large-amplitude motions in the FKBP12–ligand interfaces that strongly depend on the nature of the drug. We have monitored these motions by measuring the rates of Tyr and Phe aromatic ring flips, and hydroxyl proton exchange for Tyr residues clustered within the FKBP12–ligand interface. To do so, we prepared FKBP12 proteins selectively labeled by δ -, ϵ -, ζ -SAIL Tyr and by δ -, ϵ -SAIL Phe. Free in solution, all of the Phe and Tyr residues in the ligand-binding pocket of FKBP12 show time-averaged signals for their δ - and ϵ -CHs, due to their rapid ring-flipping rates. In contrast, in the ligand bound states, Tyr26 and Phe99 give two separated signals for their δ - and ϵ -CHs, due to the slow ring-flipping rates of these residues. In addition, significant decreases in the ring-flipping rates were observed for all of the other Tyr and Phe residues in the binding pocket, namely, Tyr82, Phe36, Phe46, and Phe48, as illustrated for the FKBP12–rapamycin complex (Fig. 2.8). The rates of hydroxyl proton exchange were also measured for ζ -SAIL Tyr-labeled FKBP12 in the drug complexes, using the method described in the next paragraph. Pairwise comparisons between FKBP12 complexed with rapamycin and FK506 revealed that the hydroxyl proton exchange and the ring-flipping rates for Tyr26 are much slower in

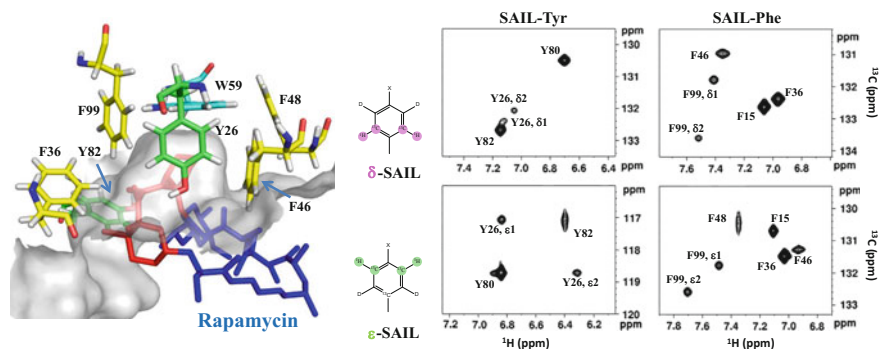


Fig. 2.8 Ligand-binding interface structure of the FKBP12–rapamycin complex in the crystalline state (PDB entry 2DG3), in which the FKBP binding motif of the rapamycin backbone is shown red, and the aromatic ring ^1H – ^{13}C correlation signals in the 600 MHz HSQC spectra at 20 °C for the rapamycin bound FKBP12s, which are residue specifically labeled with δ -, ϵ -SAIL Phe and δ -, ϵ -SAIL Tyr, respectively. The ring-flipping rates of the aromatic rings of Phe and Tyr residues in the primary binding concave surface, namely Trp59, Phe46, Phe48, Phe99 and Tyr26, were significantly slowed down, as shown by the line shapes of either or both δ - and ϵ -CH. Similarly, the ring-flipping rates of Phe48, in juxtaposition with Phe46, and Tyr82, which forms a hydrogen bond with the carbonyl oxygen at C8 of rapamycin, were also slowed down [44]

the FK506 complex than in the rapamycin complex, whereas the ring-flipping rates for Phe48 and Phe99 are significantly faster in the FK506 complex than in the rapamycin complex. The apparent rate differences observed for the interfacial aromatic residues in the two complexes confirm that these dynamic processes occur without ligand dissociation. We attribute the differential interface dynamics for these complexes to a single hydrogen bond between the ζ -hydrogen of Phe46 and the C32 carbonyl oxygen of rapamycin, which is not present in the FK506 complex. This newly identified Phe46 ζ -hydrogen bond in the rapamycin complex imposes motional restriction on the surrounding hydrophobic cluster and subsequently regulates the dynamics within the protein–ligand interface. Such information concerning large-amplitude dynamics at drug–target interfaces has the potential to provide novel clues for drug design [44].

2.3.3 Deuterium-Induced Isotope Shifts for Measuring Hydrogen Exchange Rates of Polar Side-Chain Groups in Proteins: Facile Screening of the Polar Groups Involved in Hydrogen Bond Networks

The hydrogen exchange phenomena of the backbone amides in aqueous solutions are among the most intensively studied protein dynamics by NMR spectroscopy. The exchange rates are usually estimated by a time course of the amide proton signal intensity changes for a protein freshly dissolved in D₂O. The information has made crucial contributions toward understanding the backbone dynamics and the folding–unfolding processes of proteins in solution. In contrast, the hydrogen exchange rates for the polar side-chain groups, such as hydroxyl (OH) or sulfhydryl (SH), have not been studied extensively, because they are usually too rapid to be measured by the method used for the backbone amides. We have exploited an alternative approach for the facile screening of the slowly exchanging polar side-chain groups and the estimation of their hydrogen exchange rates with the surrounding water. We adapted our previous method for detecting slowly exchanging backbone amide hydrogens by the steady-state line shapes of the amide carbonyl ¹³C signals, in a protein dissolved in a 1:1 mixture of H₂O and D₂O [7, 8, 45]. In such an environment, the line shape of the amide carbonyl of the *i*th residue depends on the isotope shift values induced by deuteration for both of the (*i* + 1)th and *i*th amides and also on their hydrogen–deuterium exchange rates [45].

Since the ¹³C chemical shift differences for the carbons directly bonded to side-chain OH or SH groups measured in H₂O and D₂O are usually a little greater than 0.1 ppm, we could use these relatively large isotope shifts for the facile screening of the slowly exchanging polar groups. To do so, we prepared proteins selectively labeled by ζ -SAIL Tyr, [3-¹³C; 3,3'-D₂]-Ser, [3-¹³C; 3-D]-Thr, or [3-¹³C; 3,3'-D₂]-Cys. These labeled proteins gave extremely sharp 1D ¹³C-NMR signals for the C $_{\zeta}$ or C $_{\beta}$ under deuterium decoupling and thus were quite useful for

estimating the isotope shifts on the C_ζ or C_β and then the hydrogen exchange rates for the slowly exchanging polar groups by the EXSY experiment. We found that quite a few Tyr, Ser, Thr, and Cys residues in various proteins, dissolved in a 1:1 H_2O – D_2O mixture, actually exhibit slow hydrogen–deuterium exchanging rates for their side-chain hydroxyl or sulfhydryl groups. Interestingly, all of the polar groups identified by this method as having very slow hydrogen exchange rates form hydrogen bonds and give 1H -NMR signals in H_2O [46–49]. Therefore, this approach is useful for screening slowly exchanging polar functional groups that are likely to play important structural roles in proteins. A typical example of a search for the Tyr residues in a protein, which might have slowly exchanging hydroxyl groups, is illustrated as follows.

In order to search for the Tyr residue with a slowly exchanging hydroxyl group, if any exist, we prepared proteins selectively labeled by ζ -SAIL Tyr (Fig. 2.9a). This particular type of SAIL Tyr has the optimal labeling pattern for observing the ^{13}C signals of ζ -carbons most efficiently and is also convenient for making sequential assignments through the NOE and HSQC correlations between $H_{\beta 3}$ and C_ζ , as shown by the red and blue arrows. The three Tyr residues in the 18.2-kDa protein EPP1b labeled by ζ -SAIL Tyr gave three sharp signals in H_2O and D_2O solutions with the sequential assignment (Fig. 2.9b, top and bottom). The chemical shifts for the $^{13}C_\zeta$ signals in H_2O appeared ~ 0.1 ppm downfield as compared to those in D_2O , due to the two-bond isotope shift induced by the deuteration of hydroxyl groups in D_2O . In a 1:1 H_2O – D_2O mixture, the $^{13}C_\zeta$ of Tyr36 and Tyr30 appeared as two separate signals with equal intensities, corresponding to the Tyr residues with a protonated and a deuterated hydroxyl group, respectively. In

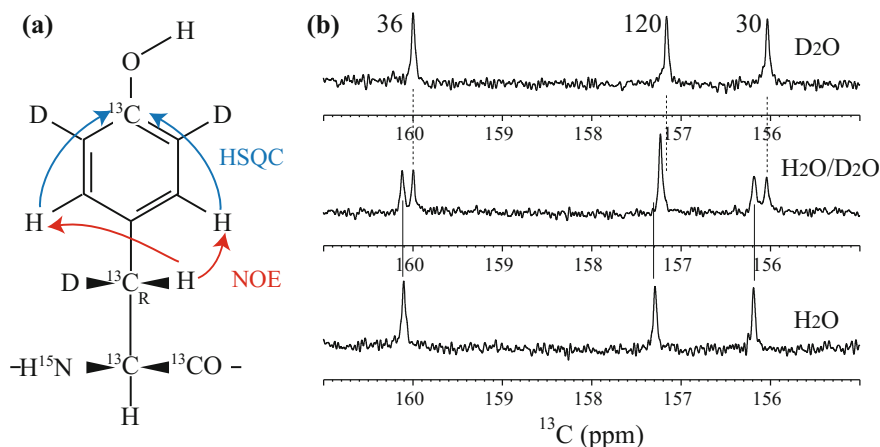


Fig. 2.9 Deuterium isotope effect on the $^{13}C_\zeta$ chemical shifts of Tyr residues in EPP1b. **a** Structure of ζ -SAIL Tyr and the magnetization transfer pathway to make the sequential assignments of the $^{13}C_\zeta$ signals. **b** Proton-decoupled ^{13}C -NMR (125 MHz, 40 °C) spectra of EPP1b selectively labeled with ζ -SAIL Tyr under conditions of 100% D_2O (top), 1:1 H_2O – D_2O mixture (middle) and 100% H_2O (bottom) [46]

contrast, the $^{13}\text{C}_\zeta$ of Tyr120 showed a single peak just in the middle of the chemical shifts observed in H_2O and D_2O solutions (Fig. 2.9b, middle). The results clearly show that the hydrogen exchange rates of the hydroxyl groups in Tyr30 and Ty36 are much slower than the isotope shift difference, while Tyr120 has a rapidly exchanging hydroxyl group. It is interesting to mention that we identified the hydrogen bonds involving the hydroxyl groups of Tyr30 and Tyr36 by NOEs, but Tyr120 is on the surface of EPPIb. We could also observe the hydroxyl proton signals for Tyr30 and Tyr36 in H_2O at the chemical shifts identified by the NOE experiment [46].

2.4 Future Perspectives of the Isotope-Aided NMR Method

The SAIL method has become well known worldwide as a state-of-the-art isotope-aided NMR technology. However, it is clear that further efforts are required to make it a standard practice among the international biological NMR communities. The substantially high cost of the SAIL amino acids is certainly one of the obstacles, but they will be more affordable if the SAIL method becomes routinely used. Cell-free protein expression, which is necessary to prepare proteins exclusively composed of SAIL amino acids, seems to be another barrier for most NMR laboratories with no such experience. However, the *E. coli* cell-free kit for preparing isotope-labeled protein samples for NMR is now available commercially at a moderate cost. Therefore, there are no major hurdles to trying out the SAIL method for structure studies of proteins. The SAIL method could be extended for solving precise structures of proteins as large as 100 kDa or even more, by further optimization of the relaxation properties of SAIL amino acids (Miyanoiri et al., unpublished). The applications of the SAIL method to solid-state NMR spectroscopy are also interesting, but they have just started [50].

Meanwhile, the expected role of NMR spectroscopy in structural biology seems to be rapidly shifting from structure determinations to dynamics studies of biologically interesting targets, such as membrane proteins and larger protein complexes. The dynamic aspects of protein–protein and protein–ligand interactions are closely related to their biological functions and can be efficiently studied by using proteins residue selectively labeled with amino acids bearing optimized labeling patterns, prepared with conventional cellular expression systems. It is therefore quite important for the NMR community to explore the stable isotope-labeling technology to its full potentials. We are absolutely confident that biological NMR spectroscopy will be continually developing with further innovations of isotope-labeling methods in the coming era, featuring ultrahigh field spectrometers beyond 1 GHz.

Acknowledgements None of our work cited in this review could have been accomplished without the seminal contributions of numerous students and collaborators over the last

half-century. We sincerely regret that we could not mention most of their important accomplishments, simply because of the space limitation. Needless to mention, our research activities were not possible without the generous, persistent financial support from various government organizations and funding agencies, although we cannot mention them individually.

References

1. Saunders, M., Wishnia, A., Kirkwood, J.G.: The nuclear magnetic resonance spectrum of ribonuclease. *J. Am. Chem. Soc.* **79**, 3289–3290 (1957)
2. Jardetzky, O.: An approach to the determination of the active site of an enzyme by nuclear magnetic resonance spectroscopy. In: *Proceeding of International Conference on Magnetic Resonance*, Tokyo, N3-14 (1965)
3. Markley, J.L., Putter, I., Jardetzky, O.: High-resolution nuclear magnetic resonance spectra of selectively deuterated staphylococcal nuclease. *Science* **161**, 1249–1251 (1968)
4. Kainosho, M.: Developing stable-isotope-aided NMR technologies for biological systems. In: *Encyclopedia of Magnetic Resonance*, pp. 1–4. doi:10.1002/9780470034590.emrhp1026 (2012)
5. Kainosho, M., Ajisaka, K., Kamisaku, M., Murai, A., Kyogoku, Y.: Conformational analysis of amino acids and peptides using specific isotope substitution. I. Conformation of L-phenylalanylglycine. *Biochem. Biophys. Res. Commun.* **64**, 425–432 (1975)
6. Kainosho, M., Ajisaka, K.: Conformational analysis of amino acids and peptides using specific isotope substitution. II. Conformation of serine, tyrosine, phenylalanine, aspartic acid, asparagine, aspartic acids β -methyl ester in various ionization states. *J. Am. Chem. Soc.* **97**, 5630–5631 (1975)
7. Kainosho, M., Tsuji, T.: Assignment of the three methionyl carbonyl carbon resonances in *Streptomyces* subtilisin inhibitor by a ^{13}C and ^{15}N double labeling technique. A new strategy for structural studies of proteins in solution. *Biochemistry* **24**, 6273–6279 (1982)
8. Kainosho, M., Tsuji, T.: ^{13}C -NMR Study of the SSI-subtilisin BPN' complex. In: Hiromi, K., Akasaka, K., Mitsui, Y., Tonomura, B., Murao, S. (eds.) *Protein Proteinase Inhibitor—The Case of Streptomyces Subtilisin Inhibitor (SSI)*, pp. 322–331. Elsevier, Amsterdam (1985)
9. Miyake, Y., Kainosho, M.: Nuclear magnetic resonance studies on *Streptomyces* subtilisin inhibitor and its complexes with proteinases. In: Fukui, T., Soda, K. (eds.) *Molecular Aspects of Enzyme Catalysis*, pp. 37–53. Kodansha-VCH, Tokyo (1994)
10. Westler, W.M., Kainosho, M., Nagao, H., Tomonaga, N., Markley, J.L.: Two-dimensional NMR strategies for carbon-carbon correlations and sequence-specific assignments in ^{13}C labeled proteins. *J. Am. Chem. Soc.* **110**, 4093–4095 (1988)
11. Westler, W.M., Stockman, B.J., Hosoya, Y., Miyake, Y., Kainosho, M., Markley, J.L.: Correlation of ^{13}C and ^{15}N chemical shifts in selectively and uniformly labeled proteins by heteronuclear two-dimensional NMR spectroscopy. *J. Am. Chem. Soc.* **110**, 6256–6258 (1988)
12. Kainosho, M., Nagao, H., Imamura, Y., Nakamura, Y., Tomonaga, N., Uchida, K., Tsuji, T.: A new strategy for the prediction of solution structure of proteins using specific isotope labeling techniques. In: Govil, G., Khetrpal, C.L., Saran, A. (eds.) *Magnetic Resonance in Biology and Medicine*, pp. 27–34. Tata-McGraw Hill, New York (1985)
13. Terauchi, T., Kobayashi, K., Okuma, K., Oba, M., Nishiyama, K., Kainosho, M.: Stereoselective synthesis of triply isotope-labeled Ser, Cys, and Ala: Amino acids for stereo-array isotope labeling (SAIL) technology. *Org. Lett.* **10**, 2785–2787 (2008)
14. Okuma, K., Ono, A.M., Tsuchiya, S., Oba, M., Nishiyama, K., Kainosho, M.: Asymmetric synthesis of (2*S*,3*R*)- and (2*S*,3*S*)-[2- ^{13}C ;3- ^2H]glutamic acid. *Tetrahedron Lett.* **50**, 1482–1484 (2009)

15. Terauchi, T., Kamikawai, T., Vinogradov, M.G., Starodubtseva, E.V., Takeda, M., Kainosho, M.: Synthesis of stereo-array isotope labeled (SAIL)-lysine *via* the head-to-tail conversion of SAIL-glutamic acid. *Org. Lett.* **13**, 161–163 (2011)
16. Oki, S.-Y., Kainosho, M.: Recent developments in stable-isotope-aided methods for protein NMR spectroscopy. In: Webb, G.A. (ed.) *Modern Magnetic Resonance Part 1*, pp. 211–218. Springer, Dordrecht (2006)
17. Oki, S.-Y., Kainosho, M.: Stable isotope labeling methods for protein NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **53**, 208–226 (2008)
18. Miyanoiri, Y., Takeda, M., Kainosho, M.: Stereo-array isotope labeling method for studying protein structure and dynamics. In: Atreya, H.S. (ed.) *Isotope Labeling Methods for Biomolecular NMR Spectroscopy. Advances in Experimental Medicine and Biology*, vol. 992, pp. 83–93. Springer, Berlin (2012)
19. Miyanoiri, Y., Takeda, M., Kainosho, M.: Stable-isotope-aided NMR spectroscopy. In: Webb, G.A. (ed.), *Modern Magnetic Resonance*, 2nd edn. Springer, Berlin. doi:[10.1007/978-3-319-28275-6](https://doi.org/10.1007/978-3-319-28275-6) (2017)
20. Kainosho, M., Torizawa, T., Iwashita, Y., Terauchi, T., Ono, A.M., Güntert, P.: Optimal isotope labelling for NMR protein structure determinations. *Nature* **440**, 52–57 (2006)
21. Kainosho, M., Güntert, P.: SAIL-Stereo-array isotope labeling. *Q. Rev. Biophys.* **42**, 247–300 (2009)
22. Kalbitzer, H.R., Leberman, R., Wittnighofer, A.: ^1H -NMR spectroscopy on elongation factor Tu from *Escherichia coli*. Resolution enhancement by perdeuteration. *FEBS Lett.* **180**, 40–42 (1985)
23. Torizawa, T., Shimizu, M., Taoka, M., Miyano, H., Kainosho, M.: Efficient production of isotopically labeled proteins by cell-free synthesis: a practical protocol. *J. Biomol. NMR* **30**, 311–325 (2004)
24. Takeda, M., Kainosho, M.: Cell-free protein production for NMR studies. In: Shekhtman, A., Burz, D.S. (eds.) *Protein NMR Techniques*, 3rd edn. *Methods in Molecular Biology*, vol. 831, pp. 71–84. Humana Press, Springer Protocols, New York (2012)
25. Takeda, M., Kainosho, M.: Cell-free protein synthesis using *E. coli* cell extract for NMR studies. In: Atreya, H.S. (ed.) *Isotope Labeling Methods for Biomolecular NMR Spectroscopy. Advances in Experimental Medicine and Biology*, vol. 992, pp. 167–177. Springer, Berlin (2012)
26. Torizawa, T., Ono, A.M., Terauchi, T., Kainosho, M.: NMR assignment methods for the aromatic ring resonances of phenylalanine and tyrosine residues in proteins. *J. Am. Chem. Soc.* **127**, 12620–12626 (2005)
27. Takeda, M., Ono, A.M., Terauchi, T., Kainosho, M.: Application of SAIL phenylalanine and tyrosine with alternative isotope-labeling patterns for protein structure determination. *J. Biomol. NMR* **46**, 45–49 (2010)
28. Miyanoiri, Y., Takeda, M., Jee, J.-G., Ono, A.M., Okuma, K., Terauchi, T., Kainosho, M.: Alternative SAIL-Tip for robust aromatic signal assignment and determination of the $\chi(2)$ conformation by intra-residue NOEs. *J. Biomol. NMR* **51**, 425–435 (2011)
29. Takeda, M., Ikeya, T., Güntert, P., Kainosho, M.: Automated structure determination of proteins with the SAIL-FLYA NMR method. *Nat. Protoc.* **2**, 2896–2902 (2007)
30. Ikeya, T., Takeda, M., Yoshida, H., Terauchi, T., Jee, J.-G., Kainosho, M., Güntert, P.: Automated NMR structure determination of stereo-array isotope labeled ubiquitin from minimal sets of spectra using the SAIL-FLYA system. *J. Biomol. NMR* **44**, 261–272 (2009)
31. Ikeya, T., Jee, J.-G., Shigemitsu, Y., Hamatsu, J., Mishima, M., Ito, Y., Kainosho, M., Güntert, P.: Exclusively NOESY-based automated NMR assignment and structure determination of proteins. *J. Biomol. NMR* **50**, 137–146 (2011)
32. Pervushin, K., Riek, R., Wider, G., Wüthrich, K.: Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of large biological macromolecules in solution. *Proc. Natl. Acad. Sci. USA* **94**, 12366–12371 (1997)

33. Pervushin, K., Riek, R., Wider, G., Wüthrich, K.: Transverse relaxation-optimized spectroscopy (TROSY) for NMR studies of aromatic spin systems in ^{13}C -labeled proteins. *J. Am. Chem. Soc.* **120**, 6394–6400 (1998)
34. Tugarinov, V., Kay, L.E.: 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.* **125**, 13868–13878 (2003)
35. Ruschak, A.M., Kay, L.E.: Proteasome allostery as a population shift between interchanging conformers. *Proc. Natl. Acad. Sci. USA* **109**, E3454–E3462 (2012)
36. Gardner, K.H., Kay, L.E.: Production and incorporation of ^{15}N , ^{13}C , ^2H (^1H - δ_1 methyl) isoleucine into proteins for multidimensional NMR studies. *J. Am. Chem. Soc.* **119**, 7599–7600 (1997)
37. Goto, N.K., Gardner, K.H., Mueller, G.A., Willis, R.C., Kay, L.E.: A robust and cost-effective method for the production of Val, Leu, Ile (δ_1) methyl-protonated ^{15}N -, ^{13}C -, ^2H -labeled proteins. *J. Biomol. NMR* **13**, 369–374 (1999)
38. Gans, P., Hamelin, O., Sounier, R., Ayala, I., Durá, M.A., Amero, C.D., Noirclerc-Savoye, M., Franzetti, B., Plevin, M.J., Boibouvier, J.: Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. *Angew. Chem. Int. Ed. Engl.* **49**, 1958–1962 (2010)
39. Miyanoi, Y., Takeda, M., Okuma, K., Ono, A.M., 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* **57**, 237–249 (2013)
40. Miyanoi, Y., Ishida, Y., Takeda, M., Terauchi, T., Inouye, M., Kainosho, M.: Highly efficient residue-selective labeling with isotope-labeled Ile, Leu, and Val using a new auxotrophic *E. coli* strain. *J. Biomol. NMR* **65**, 109–119 (2016)
41. Wagner, G., Wüthrich, K.: Dynamic model of globular protein conformations based on NMR studies in solution. *Nature* **275**, 247–248 (1978)
42. Wagner, G.: Characterization of the distribution of internal motions in the basic pancreatic trypsin inhibitor using a large number of internal NMR probes. *Q. Rev. Biophys.* **16**, 1–57 (1983)
43. Skalicky, J.J., Mills, J.L., Sharma, S., Szypersky, T.: Aromatic ring-flipping in supercooled water: implications for NMR-based structural biology of proteins. *J. Am. Chem. Soc.* **123**, 388–397 (2001)
44. Yang, C.-J., Takeda, M., Terauchi, T., Jee, J.-G., Kainosho, M.: Differential large-amplitude breathing motions in the interface of FKBP12-Drug complexes. *Biochemistry* **54**, 6983–6995 (2015)
45. Markley, J.L., Kainosho, M.: Stable isotope labelling and resonance assignments in larger proteins. In: Roberts, G.C.K. (ed.) *NMR of Macromolecules. A Practical Approach*, pp. 101–152. Oxford University Press, Oxford (1993)
46. Takeda, M., Jee, J.-G., Ono, A.M., Terauchi, T., Kainosho, M.: Hydrogen exchange rate of tyrosine hydroxyl groups in proteins as studied by the deuterium isotope effect on C_α chemical shifts. *J. Am. Chem. Soc.* **131**, 18226–18562 (2009)
47. Takeda, M., Jee, J.-G., Terauchi, T., Kainosho, M.: Detection of the sulfhydryl groups in proteins with slow hydrogen exchange rates and determination of their proton/deuteron fractionation factors using the deuterium-induced effects on the $^{13}\text{C}_\beta$ NMR signals. *J. Am. Chem. Soc.* **132**, 6254–6260 (2010)
48. Takeda, M., Jee, J.-G., Ono, A.M., Terauchi, T., Yang, C.-J., Kainosho, M.: Hydrogen exchange study on the hydroxyl groups of serine and threonine residues in proteins and structure refinement using NOE restraints with polar side-chain groups. *J. Am. Chem. Soc.* **133**, 17420–17427 (2011)
49. Takeda, M., Miyanoi, Y., Terauchi, T., Kainosho, M.: Use of H/D isotope effects to gather information about hydrogen bonding and hydrogen exchange rates. *J. Magn. Reson.* **241**, 148–154 (2014)
50. Wang, S., Parthasarathy, S., Nishiyama, Y., Endo, Y., Nemoto, T., Yamauchi, K., Asakura, T., Takeda, M., Terauchi, T., Kainosho, M., Ishii, Y.: Nano-mole scale sequential signal

- assignment by ^1H -detected protein solid-state NMR by ultra-fast magic-angle spinning and stereo-array isotope labeling. *Chem. Commun.* **51**, 15055–15058 (2015)
51. Takeda, M., Chang, C.-K., Ikeya, T., Güntert, P., Chang, Y.-H., Hsu, Y.-L., Huang, T.-H., Kainosho, M.: Solution structure of the C-terminal dimerization domain of SARS coronavirus nucleocapsid protein solved by the SAIL-NMR method. *J. Mol. Biol.* **380**, 608–622 (2008)
 52. Takeda, M., Sugimori, N., Torizawa, T., Terauchi, T., Ono, A.M., Yagi, H., Yamaguchi, Y., Kato, K., Ikeya, T., Jee, J.-G., Güntert, P., Aceti, D.J., Markley, J.L., Kainosho, M.: Structure of the putative 32 kDa myosinase binding protein from *Arabidopsis* (At3g16450.1) determined by SAIL-NMR. *FEBS J.* **275**, 5873–5884 (2008)