Chapter 3
Isotope Labeling for Solution and Solid-State NMR Spectroscopy of Membrane Proteins

Raffaello Verardi, Nathaniel J. Traaseth, Larry R. Masterson, Vitaly V. Vostrikov, and Gianluigi Veglia

Abstract In this chapter, we summarize the isotopic labeling strategies used to obtain high-quality solution and solid-state NMR spectra of biological samples, with emphasis on integral membrane proteins (IMPs). While solution NMR is used to study IMPs under fast tumbling conditions, such as in the presence of detergent micelles or isotropic bicelles, solid-state NMR is used to study the structure and orientation of IMPs in lipid vesicles and bilayers. In spite of the tremendous progress in biomolecular NMR spectroscopy, the homogeneity and overall quality of the sample is still a substantial obstacle to overcome. Isotopic labeling is a major avenue to simplify *overlapped spectra* by either diluting the NMR active nuclei or allowing the resonances to be separated in multiple dimensions. In the following we will discuss isotopic labeling approaches that have been successfully used in the study of IMPs by solution and solid-state NMR spectroscopy.

Abbreviations

IMP	Integral Membrane Protein
SSNMR	Solid-State NMR
O-SSNMR	Oriented SSNMR
MAS-SSNMR	Magic-Angle-Spinning SSNMR
PISEMA	Polarization Inversion Spin Exchange at Magic Angle

R. Verardi
Department of Biochemistry, Molecular Biology, and Biophysics,
University of Minnesota, Minneapolis, MN 55455, USA

N.J. Traaseth
Department of Chemistry, New York University, New York, NY 10003, USA

L.R. Masterson • V.V. Vostrikov
Department of Chemistry, University of Minnesota, 6-155 Jackson Hall, 321 Church St SE,
Minneapolis, MN 55455, USA

G. Veglia (✉)
Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota,
Minneapolis, MN 55455, USA
vegli001@umn.edu

Department of Chemistry, University of Minnesota,
6-155 Jackson Hall, 321 Church St SE, Minneapolis, MN 55455, USA

H.S. Atreya (ed.), *Isotope labeling in Biomolecular NMR*, Advances in Experimental Medicine and Biology 992,
DOI 10.1007/978-94-007-4954-2_3, © Springer Science+Business Media Dordrecht 2012

3.1 Introduction

Isotopic enrichment has been an integral part of the advancements made by nuclear magnetic resonance (NMR) spectroscopy for the characterization of biomacromolecules at atomic resolution. The first pioneering studies on isotopically labeled proteins were carried out in the late 1960s, resulting in the production of isotopically labeled proteins extracted from organisms (bacteria and plants) cultured in media containing isotopically labeled nutrients [1–4]. In the past few years, there has been a true explosion of labeling schemes and production techniques that has enabled NMR spectroscopic studies of proteins and protein complexes larger than 100 kDa [5–7].

While most of the structural biology has been focusing on soluble proteins, outstanding progress is being made both in liquid and solid-state NMR for the structural analysis of membrane-bound proteins. In fact, an estimated 30% of all proteins synthesized in most organisms are integral membrane proteins [8, 9], which necessitate lipid environments to properly fold and function. IMPs are involved in signal transduction, transport of molecules across the membrane, conduction of ions and many other vital cellular processes [10–13]. Despite their importance, only 308 IMPs (http://blanco.biomol.uci.edu/mpstruc/listAll/list) have been deposited in the protein data bank (PDB) as of 2011, which is a rather exiguous number compared to the thousands of high-resolution structures determined for their soluble counterparts. There are several reasons for the paucity of high-resolution IMP structures. First of all, IMPs are difficult to express and purify in large amounts (tens of milligrams) and with the proper folding. Second, IMPs need lipids or detergents for structural and functional studies. The membrane mimetic environments coat the proteins forming large and slowly tumbling complexes that complicate NMR analysis. In recent years however, improvements in protein production systems, NMR hardware, pulse sequences and isotopic labeling strategies have made possible a number of successes in the study of IMPs [6, 14].

This chapter highlights the recent progress in isotopic labeling technologies to aid solution and solid-state NMR studies of IMPs. Although only four isotopes (^{1}H, ^{15}N, ^{13}C, ^{2}H) are routinely used in biomolecular NMR, there are several ways for introducing them along the amino acid sequence (see Fig. 3.1). We focus on the recent progress from our laboratory and other research groups in the production of isotopically labeled IMPs for both liquid- and solid-state NMR studies. In addition, we review how isotopic labeling schemes can be exploited for studying protein-protein interactions in micelles and lipid vesicles. Finally, we will discuss some of the most common techniques to engineer spin-labels and isotopically labeled chemical groups to image large mammalian membrane proteins.

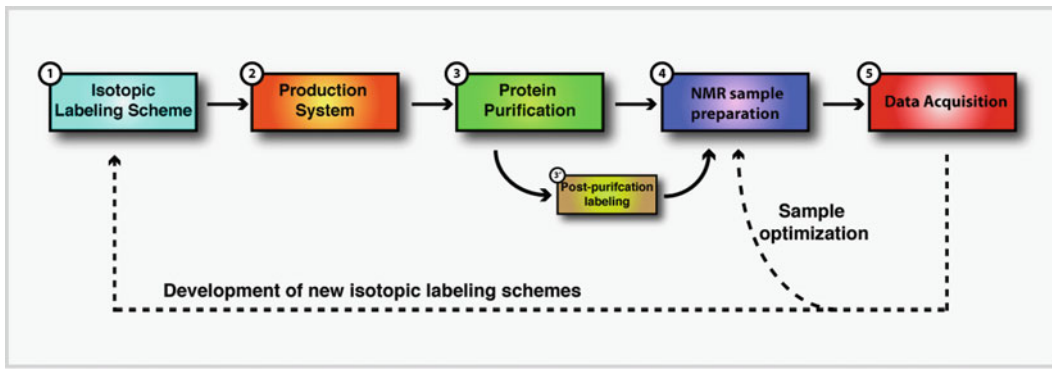

Fig. 3.1 Production of isotopically labeled membrane proteins for NMR spectroscopy

3.2 Recent Advances in the Production of IMPs

The main isotopes routinely used in protein NMR spectroscopy are ^{1}H, ^{2}H, ^{13}C and ^{15}N, with a more sparse use of ^{31}P, ^{19}F and ^{17}O. Among the main isotopes, only ^{1}H is found naturally at high abundance (>99.9%), whereas the others must be artificially introduced in proteins. Isotopic labeling schemes can be divided into two broad categories: uniform and selective labeling. In the first category, we list all methods that produce a protein with uniform incorporation of NMR active isotope (i.e., uniformly ^{13}C labeled or U-^{13}C). Conversely, if a protein is enriched with an isotope only at particular sites, the protein will be defined as selective labeled.

Because of the inherent insensitivity of NMR, it is generally necessary to have an production system capable of yielding milligram amounts of IMPs properly folded and biologically active. There are three well-established approaches: (1) heterologous overexpression, (2) total chemical synthesis and (3) cell-free expression. Depending on the protein under investigation each one of these approaches can be a viable choice. However, each system has advantages and drawbacks that need to be evaluated on a case-by-case basis.

3.2.1 Heterologous Overexpression Systems for Membrane Proteins

Heterologous overexpression consists of the use of living cells to synthesize proteins. It involves manipulation of the host DNA in such a way that the foreign gene is transcribed and translated at high levels. There are several heterologous systems for the expression and purification of IMPs [15–17], but the most widely used for isotopic labeling are: bacteria, yeasts, and insect cells. Each system has its own advantages and drawbacks, nonetheless a number of IMPs have been successfully produced for NMR studies [6]. When choosing an expression system, there are at least three important parameters to consider and eventually optimize: (1) the amount of final product (pure protein) per liter of growth medium, (2) whether the protein is properly folded and (3) whether biological activity of the expressed protein is retained.

3.2.1.1 Bacteria

The use of bacteria (especially *Escherichia coli* strains) for heterologous expression of proteins was established in the 1980s when molecular cloning techniques became widely available [18]. Bacteria offer a number of advantages over other expression systems: they can grow at high densities in a variety of synthetic media, foreign genes can be inserted in their genome using simple molecular cloning techniques, and growth rates are fast (doubling time is on the order of 30 min). *E. coli* strains can be grown in fermenter vessels, where important parameters such as pH, temperature and dissolved oxygen are monitored to increase biomass and protein expression levels. Several strategies for efficient isotopic labeling of recombinant proteins in *E. coli* have been proposed [4, 19–21]. All these methods focus on obtaining high cell densities using inexpensive unlabeled media and subsequent transfer in labeled medium immediately before expression. High expression levels for IMPs have also been obtained using a clever manipulation of the common T7 expression system, which cause autoinduction of the recombinant gene [22].

A promising new strategy for the efficient expression of labeled proteins in *E. coli* is the single protein production system [23, 24]. By expression of an mRNA interferase (MazF) that cleaves RNA at ACA nucleotide sequences, it is possible to stop cellular growth. If the mRNA of the gene of interest is engineered so that no ACA sequences are present, MazF will not cleave it and translation will continue undisturbed. By using this expression system, it has been estimated that up to 30% of total

cellular content is comprised of the recombinant protein, making it possible to acquire NMR spectra without substantial purification. When such a system is used for the production of isotopically labeled poteins the savings in terms of materials could be substantial. Indeed its success has been demonstrated by producing several IMPs [23, 24].

Although *E. coli* is a robust and reliable host cell, it presents a number of problems for the expression of IMPs. Overexpression of IMPs is often toxic to the cell, thereby decreasing the viability of the cell itself. When IMPs are expressed at high levels they often tend to aggregate into inclusion bodies [25] which require unfolding and refolding strategies in order to extract the target protein. Although these problems can be circumvented by expressing the IMPs at lower temperature, or using soluble fusion tags, the IMPs might not be in an active form since *E. coli* bacteria do not possess post-translational modification machinery.

In addition to *E. coli* bacteria, other prokaryotes have been investigated for the overexpression of IMPs. The two most promising organisms are *Pseudomonas Aeruginosa* and *Lactococcus lactis*. *P. Aeruginosa* is a gram-negative bacterium that breaks down glucose using the Entner-Doudoroff pathway rather than glycolysis, producing alternative labeling patterns. McDermott and coworkers produced Pf1 coat protein labeled with ^{13}C only at the carbonyl position by feeding *P. aeruginosa* with 1-^{13}C-glucose [26]. Although this labeling scheme was used for solid-state NMR investigation of Pf1, it has great potential for studying the dynamics of IMPs by solution NMR as well.

The second promising prokaryote for the production of IMPs is *L. lactis*. This gram-positive bacterium offers several attractive features: (1) it has a single cellular membrane, which facilitates the insertion of heterologous IMPs and reduces the formation of inclusion bodies, (2) it can grow at high cell densities in the absence of oxygen and (3) it possesses a tightly regulated inducible expression system that uses the peptide nisin for induction [27]. IMPs have been successfully produced using this system [28], although the use for isotopic labeling in NMR studies has yet to be demonstrated.

3.2.1.2 Yeasts

The inability to introduce complex post-translational modifications and obtain properly folded and functional proteins are among the most significant drawbacks for the expression of IMPs in bacteria. A solution to these problems is to use more sophisticated expression systems, such as eukaryotic cells. The simplest and most studied eukaryotic system for the expression of IMPs are the yeasts *Saccharomyces cerevisiae* and *Pichia pastoris*.

Both systems have been used to produce many IMPs for NMR and X-ray studies [29, 30]. As for *E. coli*, yeast can be cultured in completely defined media composed of simple sugars and salts. Moreover, molecular biology techniques for the recombinant expression of foreign genes are available and readily applicable for the isotopic labeling of IMPs.

3.2.1.3 Higher Eukaryotes

Other eukaryotic organisms have been used for the production of IMPs. The major advantage of using higher eukaryotes over simpler systems is the presence of more complex folding machinery and post-translational patterns. Some of the most promising systems for the isotopic labeling of IMPs are baculovirus-infected insect cells and transfected mammalian cells. Recently, a simple and inexpensive protocol for the selective isotopic labeling of proteins in insect cells has been proposed [31]. Despite their utility, insect cells suffer from some important drawbacks: (1) cost of labeled media can

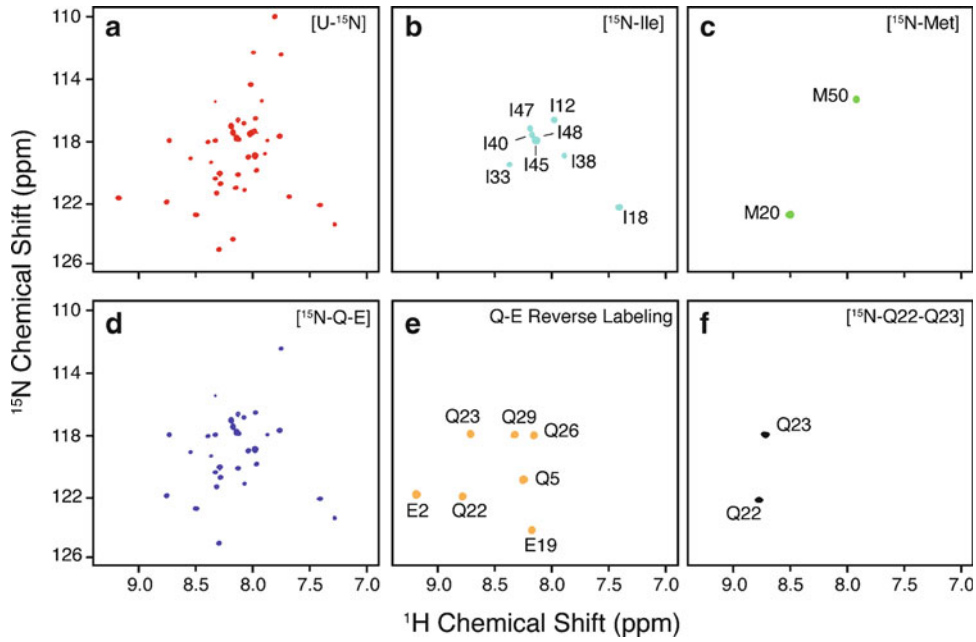

Fig. 3.2 Examples of ^{15}N uniform and selective labeling of the membrane protein PLN. (**a**) ^{15}N-^{1}H HSQC of [U-^{15}N] recombinant PLN in 300 mM DPC. (**b–c**) Selective ^{15}N-Ile and ^{15}N Met labeled recombinant PLN. Notice the absence of isotopic scrambling. (**d**) An attempt to label PLN at Gln and Glu residues using ^{15}N-Gln and ^{15}N-Glu labeled amino acids in the growth medium resulted in significant isotopic scrambling. (**e**) Labeling of Glu and Gln in PLN using the reverse labeling approach. No isotopic scrambling is present. (**f**) PLN selective labeled at Q22-Q23 produced by peptide synthesis

be prohibitive, (2) deuteration has not yet been reported and (3) the yield of pure protein can be substantially lower than other systems.

Transfected mammalian cells are another useful system to express active and properly folded IMPs. Isotopically labeled IMPs have been produced with CHO and HEK293 cells at levels comparable to simpler systems [32]. Moreover, growth media for the incorporation of ^{15}N and ^{13}C are commercially available.

3.2.2 Total Chemical Synthesis

All the production systems described so far involve the use of living cells from different organisms. There are, however, chemical methods for the synthesis of proteins of up to 100 amino acids, which can be easily adapted for isotopic labeling purposes. Chemical synthesis is usually carried out using the standard solid-phase peptide synthesis (SPPS) developed by Merrifield and coworkers [33]. SPPS uses solid resins composed of amino acid covalently linked to polystyrene beads. Protected amino acids are added to the reaction vessel where they form peptide bonds through a series of couplings and deprotection reactions. Thanks to microwave-assisted technologies which increase yields during difficult couplings and make more efficient use of isotopically labeled reagents during synthesis it is now possible to routinely produce IMPs isotopically labeled at single sites in the primary sequence (Figs. 3.2f and 3.4c).

3.2.3 Cell-Free Expression Systems

Cell-free systems are *in vitro* transcription/translation systems extracted from a variety of cells (bacteria, wheat germ, insect cells etc.) [34–36]. For cell-free systems to work, a mixture of all the 20 amino acids must be added in the reaction vessel. Because of the absence of other enzymes other than those necessary for transcription and translation, isotopic scrambling is nearly eliminated for most amino acids. In addition, this approach provides an alternative avenue to obtain IMPs that may be toxic to host cells during overexpression.

Cell-free systems can be used not only to produce residue-type selectively labeled proteins, but also for some ingenious applications such as combinatorial labeling [37–40] and stereo array isotopic labeling (SAIL) [41].

3.2.4 Membrane Protein Purification

So far, we reviewed biological and chemical systems to introduce isotopes in different positions of a protein. However, once the protein has been recombinantly expressed or chemically synthesized it must be purified to high levels (generally more than 90% purity) before NMR experiments can be undertaken. For solid-phase peptide synthesis, purification involves cleavage of the peptide from the resin and subsequent precipitation of the peptide in organic solvents. A final step of reverse-phase chromatography usually yields pure protein suitable for structural studies.

For heterologous expression of IMPs, the purification process is more involved and usually requires the use of fusion tags [42, 43].

A fusion tag is a protein or short peptide included in the same reading frame as the gene of the target protein. When the gene is transcribed and translated, the final protein will be fused to the tag through a peptide bond. Fusion tags are engineered either at the C-terminus or N-terminus and are usually separated from the protein of interest by a flexible loop.

Two important classes of fusion tags in this context are: (1) solubility tags and (2) affinity tags. To the first category belong all those tags that are used to improve solubility of the target protein. The most widely used solubility tags are: maltose binding protein (MBP), glutathione S-transferase (GST), N-utilization substance A (NusA), and Thioredoxin [43, 44].

Affinity tags are used to aid the purification of the target protein. The most common affinity tags for IMPs are: hexahistidine, GST, biotin acceptor peptide, MISTIC (acronym for membrane-integrating sequence for translation of IM protein constructs), and streptavidin binding peptide. Affinity tags bind strongly to solid supports (usually resins or gels) together with their fusion partners. The bound fusion protein can be subsequently eluted off the resin and the affinity tag removed by chemical or proteolytic cleavage [44].

Removal of the fusion tag by proteases requires the presence of specific recognition sequences that must be engineered in the gene. Tobacco etch virus (TEV) protease, factor Xa, thrombin, and enterokinase are the most commonly used enzymes to cleave off fusion tags from the target protein [45–47]. Factor Xa has a four amino acid recognition sequence (IEGR), while TEV has a more stringent seven amino acid recognition sequence (ENLYFQ/G). TEV, however, leaves one amino acid at the C-terminal side of the cleavage site that in most cases can be constructed to coincide with a native N-terminal residue in the protein sequence [48].

Some fusion tags such as MBP and GST act as both solubility and affinity tags. The MBP system is one of the most versatile systems for the expression and purification of IMPs.

In the commercially available pMal plasmid (New England BioLabs Inc.), the gene of interest is inserted upstream of the MBP gene. A recognition sequence for TEV or Factor Xa proteases can also be engineered between the two fusion partners. The plasmid is transformed into *E. coli* BL21(DE3)

competent cells and the protein is expressed under the control of the inducible Ptac promoter. Upon expression, the cells are lysed and loaded onto an amylose resin [49], which binds MBP at high affinity. After washing the resin with buffer, the fusion protein is eluted off the resin by addition of maltose, which competes with amylose to bind MBP. Purified fusion protein is cleaved using TEV protease. Following cleavage, the target protein can be separated by reverse-phase HPLC or gel filtration to the desired purity. Alternatively, solvent extraction has been successfully used in some cases [50].

3.3 Labeling Strategies in Solution State NMR

3.3.1 Uniform Isotopic Labeling

Uniform isotopic labeling consists of replacing all nuclei of a certain element with its respective isotope. As of today, the only cost-effective way to produce uniformly labeled proteins is to make use of recombinant expression in heterologous systems (see previous section). The isotope of interest is incorporated into the polypeptide by providing the organism with labeled substrates, which are then converted to labeled amino acids in the metabolic pathways [51, 52]

In the 1980s and 1990s, the development of multidimensional NMR techniques for structure and dynamics studies required proteins to be uniformly enriched in ^{15}N and/or ^{13}C. In general, ^{15}N and ^{13}C are easily introduced in the polypeptide by growing cells in minimum media containing ^{15}N ammonium salts and ^{13}C glucose as the sole nitrogen and carbon sources, respectively [51]. New media containing algal lysate have been recently used to produce uniformly labeled proteins in bacteria, achieving higher yields at lower costs [53, 54]. ^{15}N uniform labeling has become a standard strategy to enable NMR studies. Figure 3.2a shows an example of well-dispersed and homogenous correlation spectrum for a uniformly ^{15}N labeled membrane protein.

For large IMPs, the strong ^{1}H-^{1}H dipolar and heteronuclear (^{1}H-^{13}C or ^{1}H-^{15}N) relaxation pathways introduced with uniform ^{13}C and ^{15}N labeling, becomes a source of sensitivity loss. To circumvent this problem, partial and complete deuteration of proteins has been introduced [55–57]. Deuterium is a quadrupolar nucleus with a significantly lower gyromagnetic ratio compared to proton, therefore the previous relaxation pathways are largely eliminated [56]. Triple labeled proteins (U-^{2}H-^{13}C-^{15}N) are now routinely produced and used for resonance assignment purposes [57]. However, complete deuteration has some inconveniences. First, the absence of ^{1}H sites does not allow the detection of the structurally important ^{1}H-^{1}H NOE connectivities. Second, most pulse sequences terminate with detection of the proton resonances to increase sensitivity; therefore they would be useless with a completely deuterated protein. Fortunately, amide deuterons are readily exchanged with water protons and for most soluble proteins ^{1}H amide exchange is achieved during the purification steps. However, for IMPs the back exchange of amides might be more difficult due to the reduced accessibility and strong hydrogen bonding of the hydrophobic domains buried in the interior of the detergent micelle [58, 59]. In such cases, the protein must be unfolded and refolded in the presence of protonated buffers, which may generate misfolded proteins [60]. For the detection of short-range NOE contacts in large proteins, deuteration can still be useful if it is carried out at lower levels (60–70%). It has been demonstrated that partial deuteration can improve resolution and sensitivity, while enabling the detection of NOE contacts with the remaining protons [56].

As for the other isotopes, uniform deuteration is accomplished by growing cells in media containing only deuterated water as solvent and deuterated carbon sources [1]. Historically, the first isotopic labeling strategy used in protein NMR was selective deuteration in order to simplify the spectra (by dilution of the natural abundance ^{1}H signals) and decrease the linewidths (by removing the broadening effect of dipolar spin relaxation) [2, 4]. Proteins were enriched in ^{2}H by growing cells in media containing deuterated carbon sources (^{2}H-amino acid mixtures derived from algae grown in deuterated water or

^{2}H glucose) and deuterated water [2, 4]. Crespi and coworkers demonstrated how completely deuterated organisms were still able to survive and reproduce, although plant and mammalian cells could only be enriched at 20–60% with ^{2}H [61]. However, extensive deuteration can alter the structure and activity of proteins [62, 63]. Although uniform isotopic labeling still represents the first step for most protein NMR studies, this strategy does not provide the same gain for very large helical IMPs. The main obstacle when using uniform isotopic labeling of IMPs is spectral overlap, which is caused by different factors: (1) increase in the rotational correlation times, which causes line broadening, (2) degenerate chemical shifts due to the presence of only a small number of residue types (mostly Ile, Leu, Val) in transmembrane regions and (3) high occurrence of α-helical secondary structures, which decrease the breath of chemical shifts. These problems can be alleviated by using selective isotopic labeling schemes.

3.3.2 Selective Isotopic Labeling

By selective isotopic labeling, we indicate any labeling strategy that results in the incorporation of isotopes at specific sites along the polypeptide sequence. This results in NMR spectra of particular residue types in a protein sequence. An alternative approach, introduced by Oschkinat and co-workers [64], involves spectroscopic identification of individual or groups of residue types such as Gly, Ala, Thr, Val, Ile, Asn, and Gln. This approach is based on the clever use of INEPT transfer steps. However, the easiest and most widespread approach is the isotopic labeling of specific residue types using ^{15}N (and more recently ^{13}C) labeled amino acids. Traditionally, the ^{15}N and/or ^{13}C labeled amino acids are included in the growth media along with all the other "unlabeled" (^{14}N/^{12}C) amino acids. Residue-type selective labeling is extensively used to simplify spectra for assignment purposes. Not all 20 amino acids can be labeled using this strategy. In fact, the use of some amino acids results in isotopic dilution or scrambling [65]. Scrambling occurs for those amino acids that serve as precursors for the synthesis of other amino acids and results in isotopic dilution and/or distribution of the labels among other amino acids. A classic example is the amino acid glutamate, which is a central precursor for most of the other residues [66]. If ^{15}N-glutamate is used in the growth medium, the protein synthesized will have most of the other residues labeled as well. In the case of ^{15}N-labeling in heterologous expression systems, there are two ways to overcome this problem: (1) use of mutated strains (auxotrophs) and (2) reverse labeling. In the first case, libraries of *E. coli* bacteria strains have been engineered so that the metabolic pathways leading to the synthesis of each amino acid are altered through mutations [66–68]. For the amino acids Arg, Cys, Gln, Gly, His, Ile, Lys, Met, Pro and Thr, a single lesion is sufficient to eliminate isotopic scrambling [66]. This is because all of these amino acids (except Thr and Ile) are located at the end of metabolic pathways and are not used as precursors for other residues [52]. For the other amino acids, more than one genetic deletion is necessary [66]. An alternative approach is reverse labeling, which does not require mutant strains of *E. coli*. With this approach, all of the amino acids are included in the growth medium in the unlabeled (^{14}N) form, whereas the amino acid(s) of interest is omitted. ^{15}N-ammonium chloride is also included in the medium [69]. When cells grow, they will use the unlabeled amino acids for protein synthesis, but they will use ^{15}N-ammonium chloride to make up the missing amino acid(s). The result will be identical to the traditional method, but isotope scrambling can be significantly reduced. Figure 3.2 shows the comparison between an attempt to label Glu and Gln in a membrane protein using the traditional selective labeling method, resulting in severe isotopic scrambling, (Fig. 3.2d) and the reverse labeling method (Fig. 3.2e).

The use of cell-free expression systems has also been applied to a number of membrane proteins, alleviating the scrambling encountered in protein expression with bacterial host cells. In this manner, high resolution spectra of membrane proteins have been obtained from *in vitro* protein synthesis [36, 70]. A number of labeling strategies, including combinatorial, sequence-optimized, or SAIL approaches, have been used in cell-free protein synthesis to aid in resonance assignment and improve

the spectral quality of membrane proteins [71–73]. These approaches are different variations of selective-labeling of amino acids into a target protein sequence during cell-free protein expression. However, since *in vitro* expression is not complicated by various catabolic and metabolic pathways, unique protein labeling patterns can be obtained.

Another promising approach for studying large proteins is to incorporate isotopically labeled unnatural amino acids such as p-methoxy-phenylalanine (p-OMePhe), o-nitrobenzyl-tyrosine (oNBTyr), 2-amino-3-(4-(trifluoromethoxy)phenyl)propanoic acid (OCF_3Phe), trifluoromethyl-l-phenylalanine [74–76] into specific single positions along the primary sequence of a protein. This is possible by using orthogonal tRNA/tRNA synthetase pairs, which generates tRNA charged with the unnatural amino acid [75, 77, 78]. The validity of this approach was demonstrated by incorporating three unnatural amino acid in the 33 kDa thioesterase domain of human fatty acid synthase without perturbation of the protein structure [74].

Fluorine can also be selectively introduced in proteins by using fluorinated tryptophan, tyrosine or phenylalanine amino acids in *E. coli* strains auxotrophic for those amino acids [79]. Fluorine labeled amino acids have been used extensively to study protein folding, ligand binding, dynamics [79, 80], membrane immersion depth [81] and more recently solvent accessibility [82].

Finally, a new method for the labeling of specific domains of proteins has been proposed with the name "segmental labeling". This method exploits the post-translational modification, known as splicing, performed by inteins [78]. For a detailed description of this technique see previous reviews [83]. The main point of this approach is that it is possible to label (with ^{15}N and/or ^{13}C) only specific domains, while the rest of the protein remains unlabeled. This has important consequences in NMR, since the spectra are considerably simplified while retaining important inter-residue information for the labeled domain. Although useful, this technique has not been extensively applied for the production of IMPs.

3.3.3 Methyl Labeling

In highly deuterated proteins, it is advantageous to reintroduce some of the protons at specific positions [84]. For the methyl groups of isoleucine, leucine and valine, this is achieved by adding protonated precursors to the deuterated growth medium just before induction [84]. The most common of these precursors are α-ketobutyrate (yielding isoleucine) and α-ketosovalerate (yielding leucine and valine) (Fig. 3.3a, b). Due to the high degree of sensitivity via TROSY NMR of deuterated, methyl labeled proteins, a number of commercially available precursors with specific labeling patterns have been developed. For the methyl labeling of methionine, alfa-oxomethionine is added as precursor in the presence of glucose (Fig. 3.3c), whereas labeling of the methyl group of threonine can be achieved by growing cells in a medium containing a mixture of 2-^{13}C-glycerol and $NaH^{13}CO_3$ [7] (Fig. 3.3d). Slightly more involving is the ^{13}C labeling of alanine, which requires the addition of ^{13}C-labeled alanine supplemented with unlabeled succinate, α-ketoisovalerate and isoleucine to reduce isotopic scrambling (Fig. 3.3e) [7].

Methyl group labeling has proven to be a very useful strategy for membrane proteins since hydrophobic amino acids Ile, Leu and Val occur at high frequency in transmembrane domains and they are often involved in the packing of those domains [85–87]. Selective methyl labeling has been successfully applied to the study of several IMPs by solution NMR in the past few years [88–92].

3.4 Labeling Strategies in Solid State NMR

Unlike in solution NMR where rapid reorientation leads to isotropic chemical shifts and averaging of dipolar interactions, SSNMR spectra are dominated by anisotropic interactions such as anisotropic chemical shifts, quadrupolar, and dipolar couplings. The two primary classes of SSNMR methodology

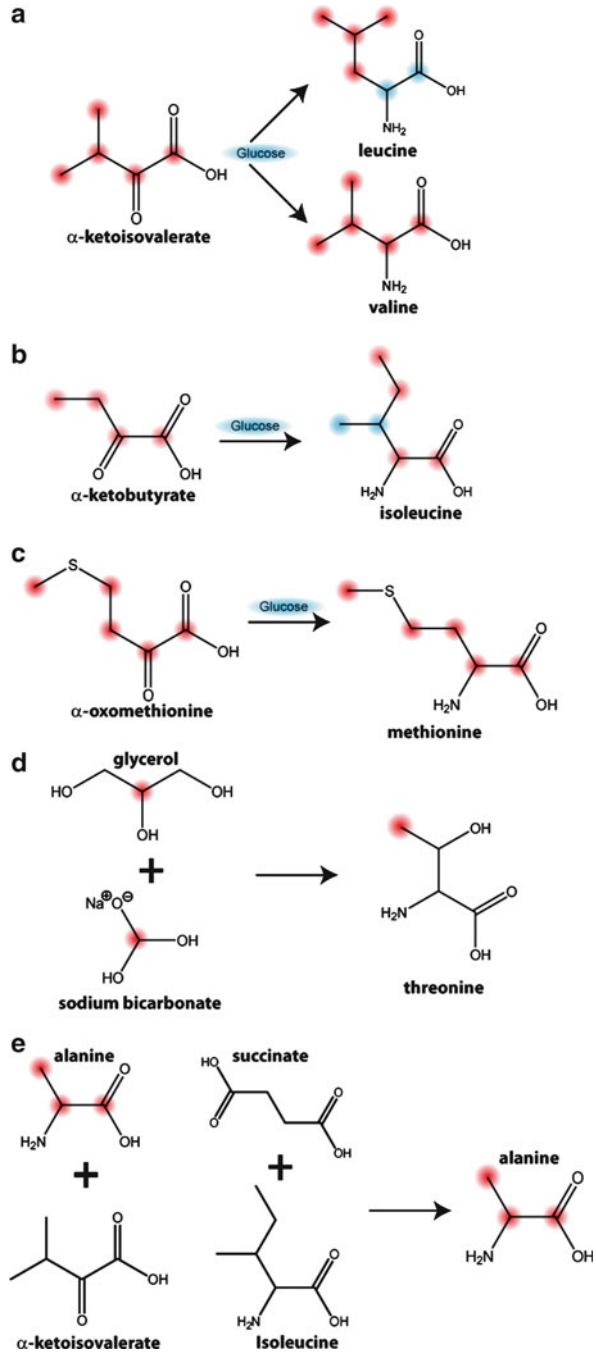

Fig. 3.3 Selective ^{13}C enrichment of methyl containing amino acids using different precursors in the presence of glucose. Carbons derived from the precursors are indicated in *red*. Note that these precursors lead to very high ^{13}C incorporation for all sites (>90%). We did not include other carbon sources (such as ^{13}C-pyruvate) that lead to lower enrichment levels at the methyl sites

are oriented (static) and magic angle spinning (MAS) experiments. MAS experiments most commonly result in solution-like isotropic spectra, whereas oriented solid-state NMR (O-SSNMR) gives orientation dependent parameters, which can be used to determine the orientation of membrane proteins in lipid bilayers or single/liquid crystals such as bicelles. Highly anisotropic systems for MAS or O-SSNMR have primarily utilized detection on ^{15}N or ^{13}C, since ^{1}H observation is hindered due to strong ^{1}H-^{1}H dipolar couplings that give rise to severe line-broadening. Techniques such as fast MAS (>60 kHz) in combination with ^{2}H labeling have made proton detection feasible in biological samples [93]. In addition, stroboscopic detection allows for the detection of signals while *simultaneously* decoupling them in a windowed-fashion [94]. Both windowed PMLG in MAS and PISEMO in O-SSNMR have benefited from these approaches. Advancements in these techniques will play an important role in the future of SSNMR due to the significant gains in sensitivity.

The following section will be broken down into labeling approaches in (1) O-SSNMR and (2) MAS-SSNMR. Subcategories of isotopic labeling strategies will be discussed that (a) reduce spectral complexity and (b) decrease the linewidth of the resonances. These two approaches are often used synergistically for optimal spectral quality.

3.4.1 Labeling Strategies in Magic-Angle-Spinning (MAS)

3.4.1.1 Uniform Isotopic Labeling

While SSNMR lines of the best-behaving samples can approach the quality of solution NMR spectra, the majority of proteins give substantially broader spectra. As an example, consider the following typical backbone ^{15}N and ^{13}C linewidths of the 6 kDa transmembrane protein phospholamban monomer (PLN) at a magnetic field of 14.1 T (600 MHz ^{1}H frequency): (a) solution NMR in detergent micelles ~0.25–0.35 ppm, (b) MAS-SSNMR in lipids ~0.75–1.5 ppm, (c) O-SSNMR in lipid bicelles ~3–6 ppm, and (d) O-SSNMR in mechanically aligned lipid bilayers ~5–10 ppm. As expected from these linewidths, the ability to resolve peaks is substantially reduced in the case of MAS and O-SSNMR. An MAS N-CA 2D correlation spectrum of uniformly labeled ^{13}C, ^{15}N spectra, [U-^{13}C,^{15}N] PLN is shown in Fig. 3.4a. From the known labeling in the sample, 52 peaks are expected. One alternative is to use 3D experiments to improve the resolution by carrying out experiments such as N-CA-C′, N-C′-CX, CA-N-C′, and other *sequential experiments* in SSNMR. However, for redundant primary sequences and helical structures such as membrane proteins, 3D experiments alone are not sufficient to resolve all the peaks. The ^{15}N dimension typically has only ~5–10 ppm in resolution (not including glycine residues). In addition, the sensitivity of multiple magnetization transfers considerably attenuates signal-to-noise, further complicating the scenario. For these reasons, reduction of spectral complexity is needed for unambiguous assignment purposes.

Similar to solution state NMR, deuteration of protein MAS samples eliminates the dipolar interactions involving protons, thus reducing the linewidths of the detected nuclei [95]. A portion of the dipolar network can be reintroduced by back-exchanging the amide protons, while the magnetization transfer to non-exchangeable side chains is achieved by expressing the proteins in the media containing minor amounts of protonated substrates [96, 97]. Since the majority of MAS pulse sequences have cross polarization as an essential block for boosting the sensitivity of low γ nuclei, deuterated samples require either direct polarization of heteronuclei (long T1 values and therefore costly from the experimental time standpoint), but can be shortened by paramagnetic doping [98]. Protein deuteration has also been observed to be beneficial in dynamic nuclear polarization experiments, yielding higher sensitivity relative to the protonated samples [99]. Furthermore, aside from providing line-narrowing of heteronuclear lineshapes (*vide supra*), deuterium itself can be employed for assignment purposes. Utility of ^{2}H in triple uniformly labeled proteins has been demonstrated for the assignment of spin

systems in ^{13}C edited spectra [100]. We note that the acquisition of such experiments can be facilitated with the help of DUMAS approach [101].

3.4.1.2 Synthetic Labeling

The simplest strategy that yields the most unambiguous assignment is to label a single residue. In this case, the assignment problem is reduced (or eliminated), and a single broad line does not cause the same resolution problems as when several signals are present. For ^{2}H or ^{17}O quadrupoles, the inherent linewidths in the spectra are on the order of ~50–100 kHz, with mosaic spread and IMP dynamics further increasing the linewidths, requiring the use of single labeled samples [102, 103]. Interpretation of quadrupolar splitting can give orientation as well as dynamics of peptides and proteins (see Sect. 3.4.2.2) [104]. This approach is very similar to EPR spectroscopy that also utilizes site-specific labeling, often with the methanesulfonothioate (MTSL) spin label if samples are made by single cysteine mutants, or 2,2,6,6-tetramethyl-piperidine-1-oxyl-4-amino-4-carboxylic acid (TOAC), prepared by SPPS.

An extension of single site-specific labeling strategy is the incorporation of two nuclear probes in which distance and dynamics information can be obtained. This is the foundation for a number of rotational-echo double-resonance (REDOR) experiments which have been used extensively in the SSNMR studies of peptides and proteins [105–109].

A further step is to selectively label stretches of residues in the primary sequence in a contiguous fashion. Such an approach has been successfully implemented by a number of MAS research groups for studying fibrils. For example, Jaroniec et al. [110] relied on three samples to assign the chemical shifts from a fragment of transthyretin (residues 105–115) fibrils. In each case the spectra are substantially simplified, since one can avoid overlap from unlike amino acids by carefully choosing the stretches of amino acids to label. Also due to the limited labeling, 2D spectra are usually sufficient to assign the spectra, without the need for longer 3D sequences that can take several weeks to acquire. Many other research groups have used this strategy in the study of amyloid fibrils, where broad lines similar to membrane proteins are present [111, 112]. We recently implemented this strategy for membrane proteins to understand the complicated folding pathways of amphipathic helices at the membrane interface [113]. Figure 3.4c shows an example of the simplification that is expected when solid-phase peptide synthesis is used to introduce a limited number of labeled residues. The main disadvantages of this technique are (a) limited applicability for large proteins (>50–75 residues in length), (b) high costs associated with purchasing some of the isotopically labeled and protected amino acids, and (c) difficulty in measuring long-range distances, since only a limited number of labeled sites are present. Nevertheless, if the protein of interest can be synthesized using SPPS, spectral quality and the ability to unambiguously assign peaks is improved.

3.4.1.3 Residue-Type Labeling

Another potential way to reduce spectral complexity and overlap is to incorporate isotopically labeled amino acids into the growth media. Unfortunately for IMPs this does not reduce a primary problem in the [U-^{15}N,^{13}C] spectra: overlap of peaks of the same residue-type (Fig. 3.4b). However, when multiple [U-^{13}C,^{15}N] amino acids are labeled at the same time, pairwise-selective labeling can be obtained. For example, consider the stretch of six residues Val^{1}-Ala^{2}-Ile^{3}-Ile^{4}-Asn^{5}-Ala^{6}. If all the residues were labeled [U-^{13}C,^{15}N], there would be five $^{13}C'$-^{15}N peptide bonds. Alternatively, residue-type selective labeling with [^{13}C,^{15}N]-Ile and [^{13}C,^{15}N]-Ala would give only two $^{13}C'$-^{15}N *pairwise* peptide bonds (Ala^{2}-Ile^{3} and Ile^{3}-Ile^{4}). A 2D N-CO MAS correlation experiment would give five cross-peaks

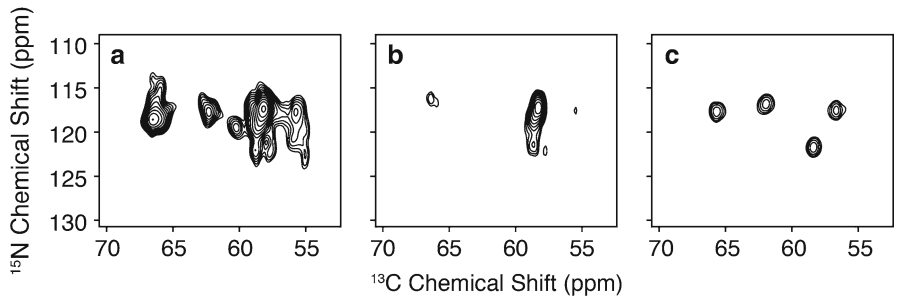

Fig. 3.4 MAS N-CA 2D correlation spectra of PLN in lipid vesicles. (**a**) uniformly labeled, [U-^{13}C,^{15}N] PLN. (**b**) Selective Leu and Val labeled PLN obtained by addition of [Val and Leu-^{13}C,^{15}N] to the growth medium. Notice the severe overlap in both dimensions. (**c**) PLN labeled with ^{13}C,^{15}N at residues Asn^{30}-Leu^{31}-Phe^{32}-Ile^{33} produced by peptide synthesis

for the [U-^{13}C,^{15}N] labeling pattern and only two for the selective labeling, thus improving unambiguous assignment.

Sensitivity of the experiment in connection with the labeling pattern can be improved with new pulse sequences. We recently implemented a complementary approach to the standard backbone experiments that increases the sensitivity of 2D correlation spectra by ~25–40%. Our filtering approach is similar to the spin-echo difference technique developed by Bax and co-workers for solution NMR [114]. This pulse sequence with a schematic and the results are shown in Fig. 3.6b. Broadly, we classify this approach as selective labeling with filtering blocks in pulse sequences to reduce the amount of peaks in the spectrum. This approach incorporates frequency selective REDOR with the N-CA selective CP of Baldus et al. [116]. Recently this approach has been extended to acquire multiple heteronuclear correlation datasets at the same time using afterglow magnetization from the cross-polarization experiment [117].

Residue-type labeling can also be employed in MAS SSNMR with selective amino acids that are not prone to scrambling. For instance, this approach has been utilized with 4-^{19}F-phenylalanine and 4-^{13}C-tyrosine to probe distances in the $\alpha_2\beta_2$ tetrameric enzyme tryptophan synthase using REDOR spectroscopy [118]. An extension of residue-type labeling is achieved using *reverse labeling* or *unlabeling*. These approaches utilize U-^{13}C glucose in the growth medium with isotopically unlabeled amino acids to produce a labeling pattern that labels those amino acids that were not supplied in the growth medium [119, 120]. This can be very advantageous, since several of these amino acids can be quite expensive to purchase, and would scramble in the growth as previously mentioned above.

3.4.1.4 Metabolic Labeling with Precursors in MAS SSNMR

An emerging approach for diluting the spin system in MAS SSNMR is the use of metabolic precursors. This method is beneficial when ^{13}C is the nucleus for direct observation. Since the presence of J-couplings (35-60 Hz) can cause line broadening, removing one-bond J-couplings can substantially improve ^{13}C spectra resolution [121]. For broader resonances > 1 ppm, only minor improvement is expected. The most common ways of diluting the ^{13}C spins is by fractional labeling or use of specifically labeled precursors: glycerol (1,3-^{13}C-glycerol or 2-^{13}C-glycerol) (Fig. 3.5), glucose (1-^{13}C-glucose or 2-^{13}C-glucose), or pyruvate with bicarbonate labeling (Fig. 3.9). Note that there are many other precursors that can be used such as keto-acids (Fig. 3.3), but these labeling patterns are less common and primarily used for methyl group spectroscopy. In the following section, we will focus on obtaining the backbone labels, since these are the foremost challenge to assign crowded SSNMR spectra.

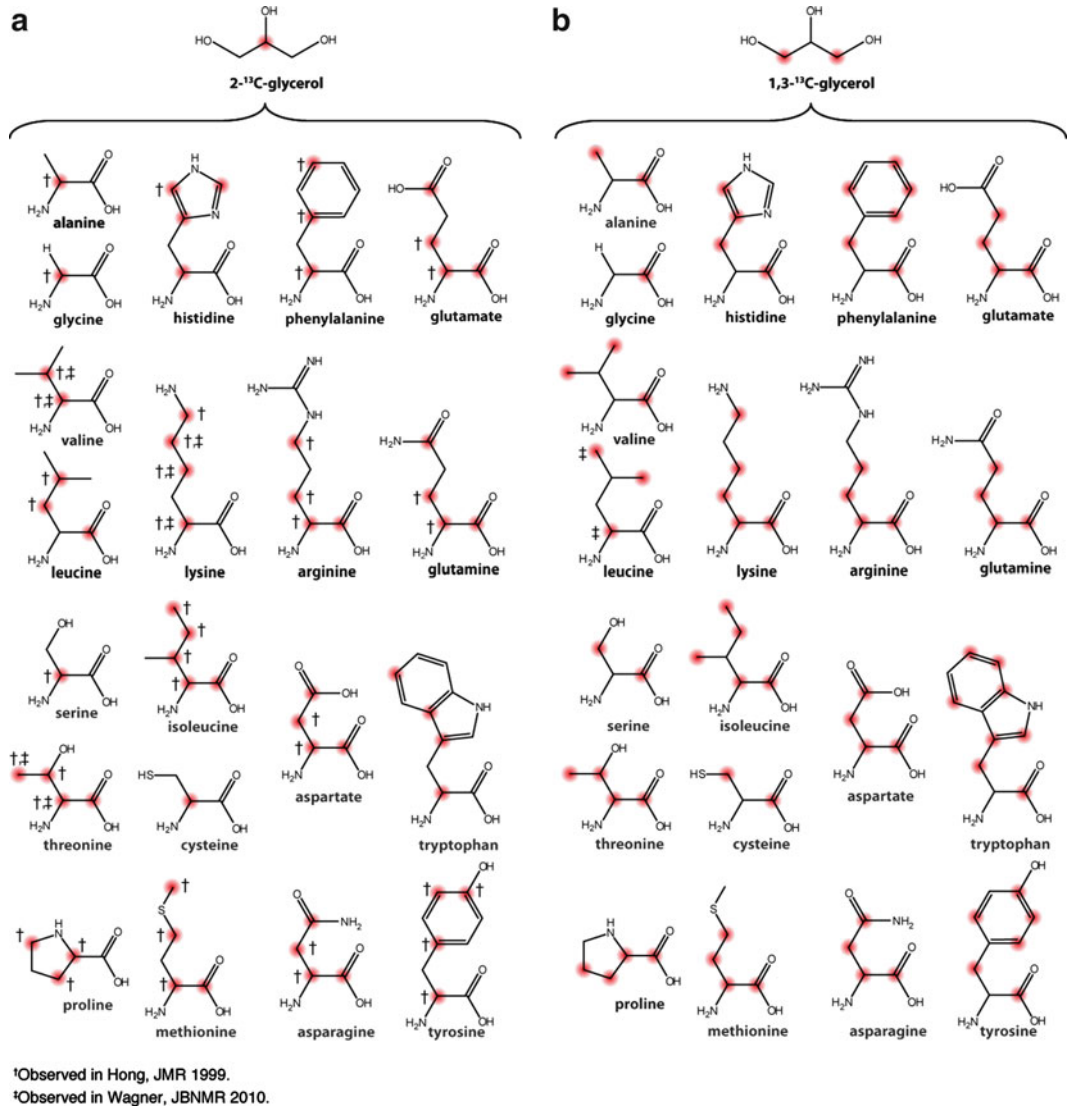

Fig. 3.5 Expected ^{13}C distribution using **a**) 2-^{13}C-glycerol or **b**) 1,3-^{13}C-glycerol as the sole carbon source and *E. coli* BL21(DE3) strain. ^{13}C labeled carbons are indicated in red two studies [121, 122] reported different results using 2-^{13}C-glycerol therefore both are indicated in the labeling pattern for each amino acid

The original approach to dilute the spin system was simply to fractionally label the protein by using a mixture of unlabeled and labeled carbon source [124]. With this approach, the labels are distributed in a stochastic manner. A significant disadvantage is the lack of pairwise labeling to assign the simplified spectra. To overcome these problems, Hong and Jakes introduced the TEASE approach (ten-amino acid selective and extensive labeling), which utilizes 2-^{13}C-glycerol, $^{15}NH_4Cl$ isotopic sources and ten unlabeled amino acids (Asp, Asn, Arg, Gln, Glu, Ile, Lys, Met, Pro and Thr) [124]. This labeling scheme results in 100% $^{13}C'$ for Gly, Ala, Ser, Cys, Phe, Tyr, Trp, His, Val and 100% incorporation at $^{13}C\alpha$ for Leu. To avoid or limit the fractional ^{13}C or ^{15}N labeling of these ten amino acids, they are added at natural abundance. Due to the use of unlabeled amino acids such as glutamine and glutamate, a two-fold dilution of ^{15}N is obtained by this method. Likewise, the 1,3-^{13}C-glycerol, gives 100% incorporation for nine amino acids (Gly, Ala, Ser, Cys, Phe, Tyr, Trp, His, Val) at the $^{13}C\alpha$

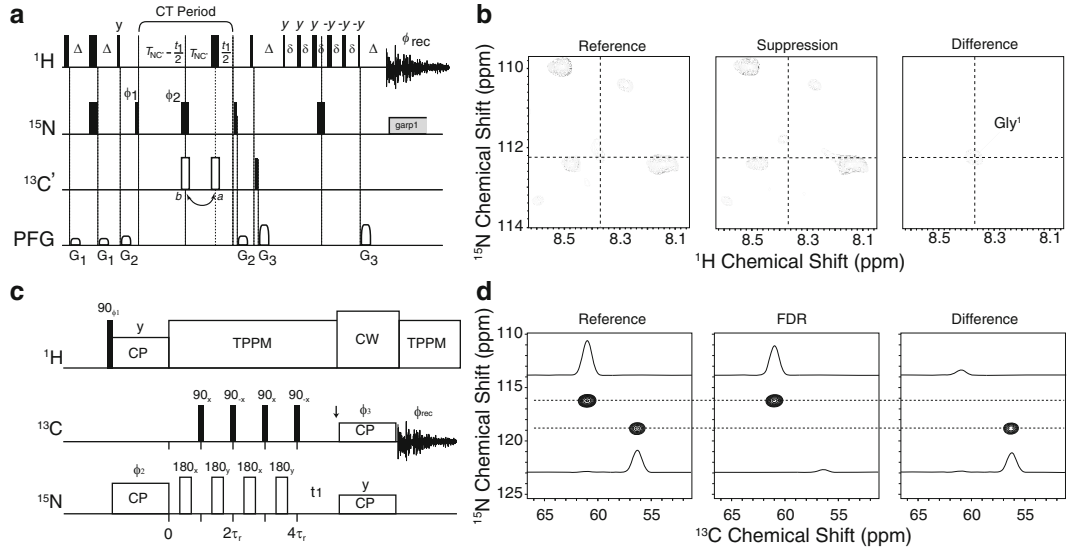

Fig. 3.6 (a–b) CCLS-HSQC. (**a**) Schematic of the CCLS-HSQC pulse sequence. (**b**) The reference spectrum is obtained by executing the pulse sequence with the 180° ^{13}C' pulse (open rectangle) at position a; the ^{13}C' suppressed spectrum is obtained with this pulse at position b. (c–d) Frequency-selective heteronuclear dephasing and selective carbonyl labeling to deconvolute crowded spectra of membrane proteins by magic angle spinning NMR. (**c**) Pulse sequence used to obtain 2D FDR-^{15}N–^{13}Cα. (**d**) FDR-^{15}N–^{13}Cα spectra for N-acetyl-valyl-leucine. Spectra were acquired with (FDR – red spectrum) and without ^{13}C 90° pulses (reference – black spectrum) (Reproduced with permission from Traaseth and Veglia [115])

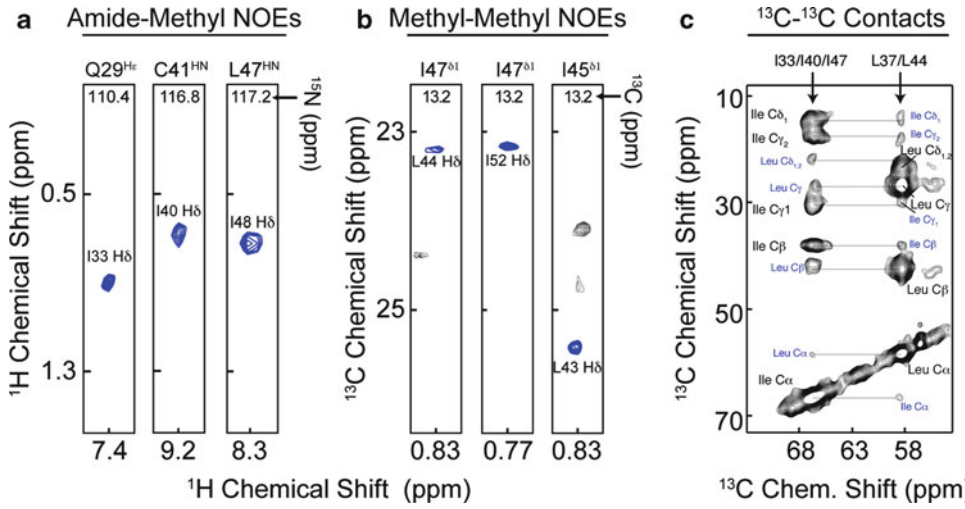

Fig. 3.7 Asymmetric labeling scheme for the detection of inter-protomer contacts in homo-oligomeric membrane proteins using solution and solid-state NMR. (**a**) 2D planes from 3D [^{1}H, ^{1}H, ^{15}N]-NOESY-HSQC (400 ms mixing time) on a mixed PLN sample with 1:1 ratio of [^{2}H-^{15}N] and [^{2}H-^{14}N-^{13}CH$_3$-Ile$^{\delta1}$] PLN. (**b**) 2D planes from 3D [^{1}H, ^{13}C, ^{13}C]-HSQC-NOESY-HSQC experiment performed on a sample containing 1:1 ratio [^{2}H-^{14}N-^{13}CH$_3$-Ile$^{\delta1}$] and [^{2}H-^{14}N-^{13}CH$_3$-Leu$^{\delta1}$/Val$^{\gamma1}$] PLN. (**c**) 2D-DARR experiments (200 ms mixing time) on a 50% [U^{13}C]-Leu/ 50% [U^{13}C]-Ile PLN sample. Intra-residue and interprotomer cross-peaks are labeled in black and blue, respectively (Reproduced with permission from Verardi et al. [90])

site. If unlabeled amino acids are not provided as in the TEASE approach, the other ten amino acids are fractionally labeled [126]. In addition to direct J-coupling removal, diluting the ^{13}C spins can also reduce cross-relaxation between ^{13}C nuclei leading to both increased resolution and sensitivity [127].

While 1,3-^{13}C-glycerol and 2-^{13}C-glycerol labeling patterns are not ideal for backbone walk due to non-contiguous ^{13}C labels, improved ^{13}C-^{13}C correlations in spin diffusion experiments have been observed due to the reduction in dipolar truncation effects. Additionally, the 1,3-^{13}C-glycerol labeling scheme is useful to reduce spectral overlap in N-CO correlation spectra (Hiller M. et al. Application note 22, Cambridge Isotope Laboratory, Inc.).

To obtain isolated ^{13}C spins (i.e., non-bonded ^{13}C-^{13}C pairs in the backbone or side chain carbon atoms), it is possible to use 2-^{13}C-glucose. This method generates 20–45% enrichment at the ^{13}Cα position, with virtually no labeling at the ^{13}C′ for all residues with the exception of Leu, which is labeled at the ^{13}C′ position. In addition, all residues are devoid of $^{13}C_{\beta}$ labeling with the exception of Leu, Val, and Ile residues. It is also possible to use 1-^{13}C-glucose. This labeling scheme enables the introduction of ^{13}C at the α-position of Leu and Ile, which are very abundant in membrane protein sequences. This labeling scheme also gives stretches of ^{13}C atoms such as ^{13}Cα-^{13}Cβ-^{13}Cγ for many residues, which can be useful for side chain detection. For a detailed summary of labeled atoms in 1-^{13}C-glucose and 2-^{13}C-glucose, see Figure 3 in Lundstrom et al. [128]. A combination of fractional labeling with selectively labeled precursors has also been used to achieve isolated spin systems. Wand et al. [124] used 15% [1-^{13}C-acetate], 15% [2-^{13}C-acetate], and 70% [1,2-^{12}C-acetate] to achieve isolated ^{13}C spins for relaxation experiments on ubiquitin.

3.4.2 Labeling Strategies for Oriented Solid-State NMR (O-SSNMR) Studies

While MAS has been used to study membrane proteins, fibrils, amorphous proteins, and crystalline proteins, O-SSNMR has been primarily used to study membrane protein structure and orientation [129–131]. Complete membrane protein structure determination requires characterization of the orientation of the membrane protein with respect to the lipid bilayer, i.e. topology. Since the energetic penalty for distorting the hydrogen bonding network is high in the lipid bilayer environment with low dielectric permeability [132], the O-SSNMR data has been often successfully interpreted assuming an idealized α-helical environment. Alternatively, O-SSNMR data can be incorporated in a total potential for structure minimization, restraining both protein topology and geometry [133–136]. Furthermore, the analysis of OSS NMR data from multiple isotopes can yield whole body dynamics of the transmembrane segments as well [137]. The O-SSNMR signal is dependent upon the angle θ between the interaction tensor components and the applied magnetic field according to the second order Legendre polynomial, $\frac{1}{2}\left(3\cos^2\theta - 1\right)$. The essential requirement for interpreting the θ angle in scope of the transmembrane domain orientation is that the NMR-active label must be rigidly attached to the polypeptide backbone. Below we discuss three different approaches in O-SSNMR, based on ^{15}N, ^{2}H and ^{19}F labeling.

3.4.2.1 Nitrogen Labeling in O-SSNMR

The most common way to determine the topology of a membrane protein is through separated local field experiments (SLF) such as PISEMA [138]. The PISEMA spectrum is considered the *fingerprint* for oriented membrane proteins, and is the most popular of the SLF class. The PISEMA experiment measures the anisotropic chemical shift of spin S and correlates it to the corresponding I-S dipolar coupling. Typically, the S spin is ^{15}N (although applicability of ^{13}C PISEMA has been illustrated [138]) and the dipolar coupling is ^{1}H-^{15}N, and correspondingly membrane proteins are either uniformly or selectively labeled with ^{15}N.

The PISEMA spectra result in periodic spectral patterns called PISA wheels [140–142]. From these wheels, it is possible to immediately obtain the tilt angles of helices or sheets with respect to the lipid bilayer normal, while determination of the rotation angle requires the assignment of the PISEMA spectrum.

As initial step, the macroscopic alignment of the protein is verified by acquiring a PISEMA spectrum using a U-^{15}N labeled protein. Often, small adjustments to the lipid composition, buffer, and temperature are necessary to find the best (homogenous) alignment. Once conditions are optimal, a high quality U-^{15}N PISEMA can be obtained that can be fit to obtain the global angle of orientation of the helices [140, 141]. One significant challenge that arises is how to assign a labeled PISEMA spectrum. There are several ways this can be done: (1) spin diffusion experiments with a single [U-^{15}N] sample [143, 144]; (2) assignment of isotropic ^{1}H and ^{15}N chemical shifts from solution NMR or MAS SSNMR in conjunction with a pair of flipped and unflipped aligned bicelle SLF spectra, requiring selective labeling [145]; (3) use of periodic assignment algorithms (based on PISA wheel) with uniform and/or selective labeled samples ("shotgun" approach) [146, 147].

Since chemical shifts are anisotropic in O-SSNMR, the orientation of the internuclear NH vector with respect to the magnetic field rather than residue-type or secondary structure determines the resonance position. This is a significant help to resolve spectral overlap in highly degenerate transmembrane helical segments. Nevertheless, uniformly labeled samples can still present severe spectral overlap and are often difficult to assign with selective labeling represents a reliable source for completing the assignments. Fortunately, the majority of the transmembrane helices are enriched with amino acids that have aliphatic side chains, which are not prone to isotopic scrambling. By labeling a protein sample with U-^{15}N-Leu or U-^{15}N-Ile, one can substantially decrease the complexity of the spectra. One can also use residue-specific labeling to determine accessibility as in H/D accessibility or proximity to a spin-label as is commonly done in solution NMR for membrane proteins [148].

Pairwise labeling utilized in solution NMR has not been extensively tested in O-SSNMR. This labeling scheme will be useful to resolve backbone resonances, when triple-resonance experiments will become routine for membrane proteins aligned in bicelles or mechanically aligned bilayers. In addition, isotopic dilution will reduce strong dipolar couplings and enable the acquisition of high quality spectra.

3.4.2.2 Deuterium Labeling in O-SSNMR

While the SLF experiments provide an initial picture of the IMP topology in lipid bilayers, they suffer from an intrinsically low sensitivity due to the orientation of the internuclear ^{15}N-H bond vectors, and in many cases where more precision is required it is often advantageous to employ isotopic labels which axes of interactions are positioned close to the magic angle relative to the helix axis. The combination of Φ-Ψ dihedral angles in a regular α-helix along with the tetrahedral geometry of the C_α carbon dictates that the C_α-C_α and C_α-H_α bond vectors form angles close to the magic angle with the helix axis (59.4° and 122.0° respectively) thus providing the maximum sensitivity for the interactions which are directed along these bonds. Alanine with a deuterated methyl group is therefore a natural choice for determining the topology of IMPs. Initial proof of concept has been carried out by labeling only a few residues at a time [149, 150] and the first systematic study was performed utilizing model Ala-rich peptides in a variety of lipid bilayers [151]. Since then deuterium NMR of methyl groups has been extensively applied for the investigation of antimicrobial peptides [152], IMPs [153], numerous model systems [154] and peptaibols [155].

Since deuterium NMR is recorded in a one-dimensional fashion typically employing a quadrupolar echo experiment [156] or quadrupolar CPMG [157], the spectral resolution precludes labeling of multiple sites, typically limiting the IMP to one or two labeled alanines. Unlike ^{1}H-^{15}N dipolar couplings, which retain a constant sign for transmembrane segments of IMPs, quadrupolar couplings oscillate

between positive and negative values, but the sign typically cannot be determined experimentally, unless it exceeds ¾ of the quadrupolar coupling constant (i.e. >37 kHz for the methyl groups) in which case the sign must be positive. Such sign ambiguity necessitates employing multiple labels, or combining the methyl restraints with other O-SSNMR labels. Limited resolution that can be achieved in 1D experiment along with the complexity of the metabolic pathways limits deuterium NMR to the synthetic sequences.

The deuteron at an α-carbon presents an appealing supplement to the alanine methyl groups, since it is present in each of the canonical amino acids and its quadrupolar coupling undergoes major changes upon the transmembrane domain tilt or rotation. The attempts to employ $^{2}H_{\alpha}$ O-SSNMR has had limited success so far. In multiple single-span IMPs the backbone deuteron either could not be detected, or observed with extremely low sensitivity [151, 158]. Interestingly in several cases a significant increase in $^{2}H_{\alpha}$ signal intensity has been observed, which potentially relates to the peptide plane and/or whole body dynamics of IMPs [159].

These examples by no means cover all the uses of deuterium in oriented solid-state NMR of IMPs (for its use in solution and MAS NMR see above). Other applications include probing the aliphatic side chain dynamics [160–162], orientation of the Trp indoles [163], IMPs oligomerization [164], mobility of the lipidated IMPs [165, 166] as well as a multitude of studies of lipid bilayer membranes – IMPs hosts.

3.4.2.3 Fluorine Labeling in OSS NMR

For detailed considerations of ^{19}F O-SSNMR the reader is referred to the excellent recent review by Ulrich and co-workers [167], while we present a brief overview below. Fluorine is a highly appealing nucleus in biological NMR. High gyromagnetic ratio, 100% natural abundance of the NMR-active ^{19}F isotope and the lack of natural background leads to high sensitivity [168]. Care must be taken to exclude fluorinated solvents (e.g. trifluoroacetic acid, a frequent ion pairing additive) as well as fluorinated polymers from the probe assembly. Close Larmor frequencies of fluorine and hydrogen exert stringent requirements on the NMR hardware. Since biomolecules do not contain fluorine, unnatural amino acids, usually based on Phe, Pro or Aib, need to be introduced in the sequence synthetically, although promising results have been achieved with the genetic incorporation [169].

3.5 Isotopic Labeling for Protein-Protein Interaction Studies

A very useful application of methyl labeling (see Sect. 3.3.3) and uniform isotopic labeling (see Sect. 3.4.1.1) is found in the study of homo-oligomeric membrane proteins by NMR. Because of the symmetry of such molecules, the NMR signals are chemically equivalent; therefore only one set of resonances is observed. In order to obtain structural information about symmetric oligomers, asymmetric labeling strategies have been developed [91, 170, 171]. The objective of these strategies is to introduce “isotopic asymmetry” in the complexes. This can be done by labeling one of the protomers with a certain isotopic scheme and the other with a different scheme. Upon formation of the complex or oligomer, the intermolecular contacts can be unambiguously assigned. Pulse sequences can be designed to detect the dipolar contacts between the protomers [90, 170, 172].

We recently proposed two asymmetric labeling schemes to measure inter-protomer contacts in the pentameric phospholamban (PLN) for solution and solid-state NMR [90, 170]. PLN is homo-pentamer composed of five identical protomers (52 residues each). The transmembrane portion of each protomer consists of mainly hydrophobic amino acids Ile, Leu and Val, which are involved in keeping the oligomer together thorough hydrophobic interactions. The first labeling scheme was devised in order

to probe inter-protomer contacts in detergent micelles by solution NMR. In this scheme, half of the protomers were labeled [U-^{2}H, ^{12}C, ^{14}N] and $^{13}CH_3$ at the Ile$^{\delta1}$ (using 2-ketobutyric acid-4-^{13}C,3,3-d_2 as precursor), whereas the other half was labeled [U-^{2}H, ^{12}C, ^{14}N] and $^{13}CH_3$ at the Leu$^{\delta1/2}$/Val$^{\gamma1/}$ (using 2-keto-3-(methyl-d_3)-butyric acid-4-^{13}C as precursor). Using a methyl-methyl NOESY pulse sequence, it was possible to identify and unambiguously assign inter-protomer contacts, which were used for structure calculations (Fig. 3.7b). This I-LV methyl labeling scheme is very powerful since Ile$^{\delta1}$ resonates at significantly different frequencies compared to Leu$^{\delta1/2}$/Val$^{\gamma1/2}$. Therefore the presence of inter-protomer contacts is straightforward to identify and correctly assign. This scheme can easily be extended to measure inter-protomer contacts between methyls and backbone amides, where half of the protein is uniformly (or selectively) labeled with ^{15}N at the amide groups in a deuterated background and half of the protein is methyl labeled at either Ile, or Leu/Val (Fig. 3.7a) [90]. A similar approach was used to identify inter-protomer contacts in lipid vesicles using MAS-NMR. In this case half of the protein was selectively labeled with ^{13}C using [^{13}C-Ile] amino acid and the other half was labeled with ^{13}C using [^{13}C-Leu] amino acid. The inter-protomer contacts were detected using a dipolar assisted rotational resonance (DARR) pulse sequence (Fig. 3.7c).

3.6 Post-expression Labeling

3.6.1 Post-expression Isotopic Labeling

There are several chemical methods to modify reactive amino acid side-chain groups after protein expression and purification [173]. By using isotopically labeled reagents, it is possible to selectively enrich amino acids with molecules containing NMR active isotopes. The most common residues whose side-chains can be chemically modified for NMR studies are cysteines, tyrosines and lysines.

The sulfhydryl group (–SH) of free cysteine in a protein can easily react in mild conditions with different chemical groups. Two applications that make use of the high nucleophilicity of free thiol groups in cysteines are the introduction of fluorine atoms and site directed methyl group substitution. In the first case, the NMR active ^{19}F is attached to cysteine by reaction of the free thiol with trifluoromethyl derivatives such as: 3-bromo-1,1,1-trifluoroacetone (BTFA) [174], trifluoroethylthio group (TET) [174], S-ethyl-trifluorothioacetate (SETFA) [176] and trifluoroacetamidosuccinican-hydride (TFASAN) [177]. This labeling approach has been successfully applied to the study of several proteins such as: citrate synthase [178], G-actin [179, 180], Myosin S-1/F-actin complex [181], SH3 domain [182], rhodopsin [175] and β2-Adrenergic Receptor [183]. Recently, Kay and co-workers introduced isotopically labeled methyl groups in cysteine side chains using methyl methanethiosulfonate to form ^{13}C-S-methylthiocysteine [184]. This labeling is very promising considering the advantages of observing methyl resonances by NMR and the fact that S-methylthiocysteine is very similar to a methionine residue, therefore it should not substantially alter the secondary structure of the protein. We have recently applied this approach to the selective methyl labeling of accessible cysteines in the 110 kDa integral membrane protein SERCA (sarcoplasmic reticulum Ca^{2+} ATPase) and obtained high-resolution solid and solution state NMR spectra (Fig. 3.8).

Another residue whose side chain can be modified is tyrosine. Richards et al. have proposed an electrochemical method for the nitration of the tyrosine ring at positions 3 in different proteins [185]. Tyrosine can also be mono-fluorinated by electrophilic substitution using acetyl hypofluorite in mild conditions and high yields (50–65%) [186].

Reductive methylation of lysine side chain has been used in many solution NMR studies to detect protein-protein interactions and ligand binding. The reaction occurs by addition of ^{13}C labeled formaldehyde to the protein solution in reducing condition [187]. If sufficient formaldehyde is present, the

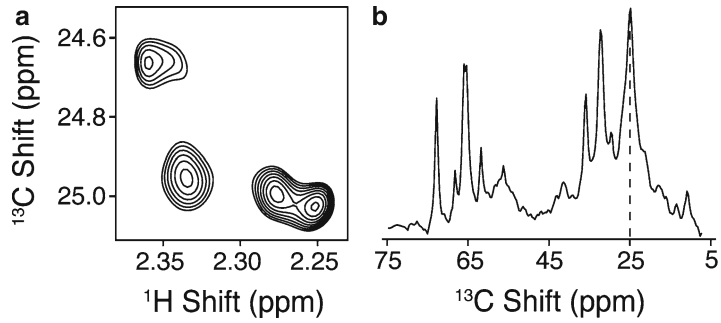

Fig. 3.8 Cysteine methylation of SERCA1a by methyl methanethiosulfonate (MMTS) reaction. (**a**) ^{1}H–^{13}C HSQC spectrum of ^{13}C methylthiocysteine in 100 mM ^{2}H dodecylphosphocholine acquired at 14.1 T field strength. (b) MAS one-dimensional cross-polarization of ^{13}C methylthiocysteine labeled SERCA1a in ^{2}H DMPC lipid vesicles run at −20 °C and spinning rate of 8,000 Hz acquired at 14.1 T field strength. Dashed lines indicate the peak corresponding to the labeled cysteines

side-chain of lysine residues will form a tertiary amine with two methyl-group substitutions [188]. This approach has been successfully applied by Kobilka and coworkers for the solution NMR study of the β2-Adrenergic Receptor [189].

3.6.2 Spin Labeling in NMR

Spin labeling refers to the covalent attachment of molecules with one or more unpaired electrons to proteins. Traditionally spin labeling has been used to study polypeptides by electron spin resonance; however, the effects of unpaired electron on the relaxation of nuclei is becoming routine in protein NMR studies [93, 190, 191]. Paramagnetic-based distance restraints have been used for the refinement of membrane protein structures [148] and for the positioning of membrane proteins in the lipid bilayers or detergent micelle [93].

Spin labeling is usually achieved post-translationally by *in vitro* chemical reactions involving cysteines through disulfide formation [192, 193] or lysines [173].

All these chemical methods must be used with caution to ensure that the reaction does not jeopardize the structural integrity or function of the protein. Furthermore, if the residues to be labeled are found buried in the core of soluble proteins or in transmembrane segments of membrane proteins, they might not be accessible to the labeling reagent.

3.7 Conclusions

The investigation of membrane proteins by NMR is a complex endeavor, but thanks to the development of improved instrumentation and production methods it is becoming increasingly feasible. New pulse sequences are continuously being devised that require specific labeling schemes, such as those described in this chapter. At the same time methods for the production of larger and more complex membrane proteins are also being actively developed.

Taken together, these accomplishments will permit an increasing number of medically relevant membrane proteins and protein complexes to be studied.

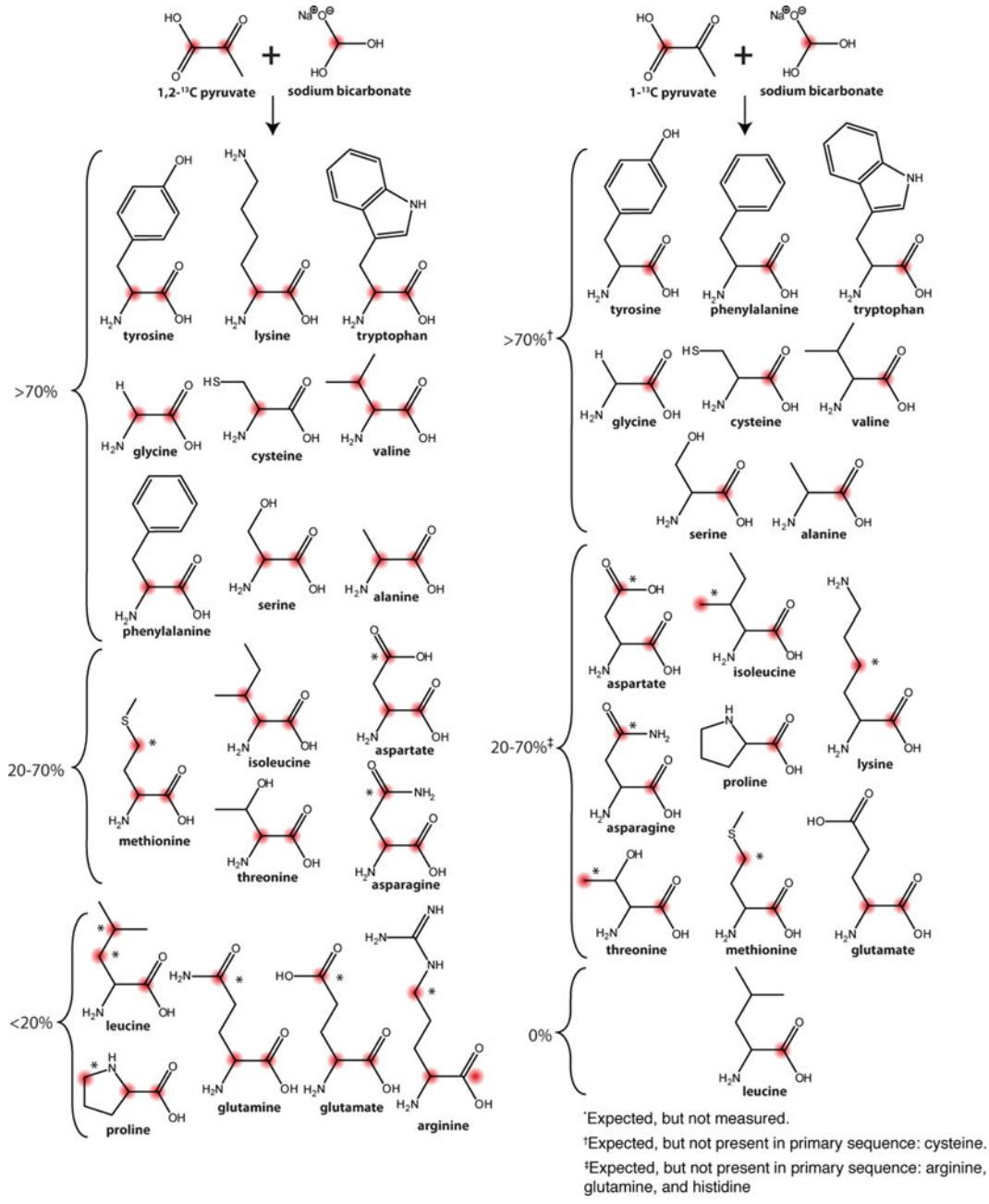

Fig. 3.9 Expected ^{13}C distribution using pyruvate and sodium bicarbonate as the sole carbon sources in E. *coli* BL21(DE3). (a) 1,2-^{13}C-pyruvate and $NaH^{13}CO_3$ and (b) 1-^{13}C-pyruvate and $NaH^{13}CO_3$

Finally, we should point out that this chapter is not exhaustive of this field, which is in continuous evolution. Most of the examples reported are based on our own experience with membrane protein structural biology. The inevitable gaps present in this Chapter are filled in the other chapters of this book by outstanding scientist in the field of structural biology.

References

1. Crespi HL, Katz JJ (1969) High resolution proton magnetic resonance studies of fully deuterated and isotope hybrid proteins. Nature 224:560–562
2. Crespi HL, Rosenberg RM, Katz JJ (1968) Proton magnetic resonance of proteins fully deuterated except for 1H-leucine side chains. Science 161:795–796
3. Putter I, Barreto A, Markley JL, Jardetzky O (1969) Nuclear magnetic resonance studies of the structure and binding sites of enzymes. X. Preparation of selectively deuterated analogs of staphylococcal nuclease. Proc Natl Acad Sci USA 64:1396–1403
4. Markley JL, Putter I, Jardetzky O (1968) High-resolution nuclear magnetic resonance spectra of selectively deuterated staphylococcal nuclease. Science 161:1249–1251
5. Ohki S, Kainosho M (2008) Stable isotope labeling methods for protein NMR spectroscopy. Prog Nucl Magn Reson Spectrosc 53:208–226
6. Kim HJ, Howell SC, Van Horn WD, Jeon YH, Sanders CR (2009) Recent advances in the application of solution NMR spectroscopy to multi-span integral membrane proteins. Prog Nucl Magn Reson Spectrosc 55:335–360
7. Ruschak AM, Kay LE (2010) Methyl groups as probes of supra-molecular structure, dynamics and function. J Biomol NMR 46:75–87
8. Wallin E, von Heijne G (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. Protein Sci 7:1029–1038
9. Ahram M, Litou ZI, Fang R, Al-Tawallbeh G (2006) Estimation of membrane proteins in the human proteome. In Silico Biol 6:379–386
10. Gilman AG (1987) G proteins: transducers of receptor-generated signals. Annu Rev Biochem 56:615–649
11. Wettschureck N, Offermanns S (2005) Mammalian G proteins and their cell type specific functions. Physiol Rev 85:1159–1204
12. Hille B (2001) Ion channels of excitable membranes. Sinauer Associates, Sunderland, pp 814, [8]
13. Brini M, Carafoli E (2009) Calcium pumps in health and disease. Physiol Rev 89:1341–1378
14. Traaseth NJ et al (2008) Structural and dynamic basis of phospholamban and sarcolipin inhibition of ca(2+)-ATPase. Biochemistry 47:3–13
15. Page RC et al (2006) Comprehensive evaluation of solution nuclear magnetic resonance spectroscopy sample preparation for helical integral membrane proteins. J Struct Funct Genomics 7:51–64
16. Eshaghi S et al (2005) An efficient strategy for high-throughput expression screening of recombinant integral membrane proteins. Protein Sci 14:676–683
17. Tate CG (2001) Overexpression of mammalian integral membrane proteins for structural studies. FEBS Lett 504:94–98
18. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor
19. Ross A et al (2004) Optimised fermentation strategy for 13C/15N recombinant protein labelling in Escherichia coli for NMR-structure analysis. J Biotechnol 108:31–39
20. Cai M et al (1998) An efficient and cost-effective isotope labeling protocol for proteins expressed in Escherichia coli. J Biomol NMR 11:97–102
21. Marley J, Lu M, Bracken C (2001) A method for efficient isotopic labeling of recombinant proteins. J Biomol NMR 20:71–75
22. Studier FW (2005) Protein production by auto-induction in high density shaking cultures. Protein Expr Purif 41:207–234
23. Suzuki M, Mao L, Inouye M (2007) Single protein production (SPP) system in Escherichia coli. Nat Protoc 2:1802–1810
24. Schneider WM et al (2010) Efficient condensed-phase production of perdeuterated soluble and membrane proteins. J Struct Funct Genomics 11:143–154
25. Baneyx F, Mujacic M (2004) Recombinant protein folding and misfolding in Escherichia coli. Nat Biotechnol 22:1399–1408
26. Goldbourt A, Day LA, McDermott AE (2007) Assignment of congested NMR spectra: carbonyl backbone enrichment via the entner-doudoroff pathway. J Magn Reson 189:157–165
27. Kunji ER, Slotboom DJ, Poolman B (2003) Lactococcus lactis as host for overproduction of functional membrane proteins. Biochim Biophys Acta 1610:97–108
28. Janvilisri T, Shahi S, Venter H, Balakrishnan L, van Veen HW (2005) Arginine-482 is not essential for transport of antibiotics, primary bile acids and unconjugated sterols by the human breast cancer resistance protein (ABCG2). Biochem J 385:419–426
29. Koth CM, Payandeh J (2009) Strategies for the cloning and expression of membrane proteins. Adv Protein Chem Struct Biol 76:43–86

30. Lin-Cereghino J, Lin-Cereghino GP (2007) Vectors and strains for expression. Methods Mol Biol 389:11–26
31. Gossert AD et al (2011) A simple protocol for amino acid type selective isotope labeling in insect cells with improved yields and high reproducibility. J Biomol NMR 51(4):449–456
32. Werner K, Richter C, Klein-Seetharaman J, Schwalbe H (2008) Isotope labeling of mammalian GPCRs in HEK293 cells and characterization of the C-terminus of bovine rhodopsin by high resolution liquid NMR spectroscopy. J Biomol NMR 40:49–53
33. Stewart JM, Young JD (1984) Solid phase peptide synthesis. Pierce Chemical Co, Rockford, p 176
34. Klammt C et al (2007) Cell-free production of G protein-coupled receptors for functional and structural studies. J Struct Biol 158:482–493
35. Klammt C et al (2006) Cell-free expression as an emerging technique for the large scale production of integral membrane protein. FEBS J 273:4141–4153
36. Klammt C et al (2004) High level cell-free expression and specific labeling of integral membrane proteins. Eur J Biochem 271:568–580
37. Wu PS et al (2006) Amino-acid type identification in 15N-HSQC spectra by combinatorial selective 15N-labelling. J Biomol NMR 34:13–21
38. Ozawa K, Wu PS, Dixon NE, Otting G (2006) N-labelled proteins by cell-free protein synthesis. strategies for high-throughput NMR studies of proteins and protein-ligand complexes. FEBS J 273:4154–4159
39. Jeremy Craven C, Al-Owais M, Parker MJ (2007) A systematic analysis of backbone amide assignments achieved via combinatorial selective labelling of amino acids. J Biomol NMR 38:151–159
40. Parker MJ, Aulton-Jones M, Hounslow AM, Craven CJ (2004) A combinatorial selective labeling method for the assignment of backbone amide NMR resonances. J Am Chem Soc 126:5020–5021
41. Kainosho M, Guntert P (2009) SAIL – stereo-array isotope labeling. Q Rev Biophys 42:247–300
42. Xie H, Guo XM, Chen H (2009) Making the most of fusion tags technology in structural characterization of membrane proteins. Mol Biotechnol 42:135–145
43. Waugh DS (2005) Making the most of affinity tags. Trends Biotechnol 23:316–320
44. Arnau J, Lauritzen C, Petersen GE, Pedersen J (2006) Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. Protein Expr Purif 48:1–13
45. Kapust RB, Waugh DS (2000) Controlled intracellular processing of fusion proteins by TEV protease. Protein Expr Purif 19:312–318
46. Abdullah N, Chase HA (2005) Removal of poly-histidine fusion tags from recombinant proteins purified by expanded bed adsorption. Biotechnol Bioeng 92:501–513
47. Jenny RJ, Mann KG, Lundblad RL (2003) A critical review of the methods for cleavage of fusion proteins with thrombin and factor xa. Protein Expr Purif 31:1–11
48. Kapust RB, Tozser J, Copeland TD, Waugh DS (2002) The P1′ specificity of tobacco etch virus protease. Biochem Biophys Res Commun 294:949–955
49. Buck B et al (2003) Overexpression, purification, and characterization of recombinant ca-ATPase regulators for high-resolution solution and solid-state NMR studies. Protein Expr Purif 30:253–261
50. Hu J et al (2007) Structural biology of transmembrane domains: efficient production and characterization of transmembrane peptides by NMR. Protein Sci 16:2153–2165
51. McIntosh LP, Dahlquist FW (1990) Biosynthetic incorporation of 15N and 13C for assignment and interpretation of nuclear magnetic resonance spectra of proteins. Q Rev Biophys 23:1–38
52. Hoogstraten CG, Johnson JE (2008) Metabolic labeling: taking advantage of bacterial pathways to prepare spectroscopically useful isotope patterns in proteins and nucleic acids. Concepts Magn Reson A 32A:34–55
53. Fiaux J, Bertelsen EB, Horwich AL, Wuthrich K (2004) Uniform and residue-specific 15N-labeling of proteins on a highly deuterated background. J Biomol NMR 29:289–297
54. Suzuki H et al (2005) Isotopic labeling of proteins by utilizing photosynthetic bacteria. Anal Biochem 347:324–326
55. LeMaster DM, LaIuppa JC, Kushlan DM (1994) Differential deuterium isotope shifts and one-bond 1H-13C scalar couplings in the conformational analysis of protein glycine residues. J Biomol NMR 4:863–870
56. Grzesiek S, Anglister J, Ren H, Bax A (1993) Carbon-13 line narrowing by deuterium decoupling in deuterium/carbon-13/nitrogen-15 enriched proteins. Application to triple resonance 4D J connectivity of sequential amides. J Am Chem Soc 115:4369–4370
57. Gardner KH, Kay LE (1998) The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins. Annu Rev Biophys Biomol Struct 27:357–406
58. White SH, Wimley WC (1999) Membrane protein folding and stability: physical principles. Annu Rev Biophys Biomol Struct 28:319–365
59. Veglia G, Zeri AC, Ma C, Opella SJ (2002) Deuterium/hydrogen exchange factors measured by solution nuclear magnetic resonance spectroscopy as indicators of the structure and topology of membrane proteins. Biophys J 82:2176–2183
60. Oxenoid K, Kim HJ, Jacob J, Sonnichsen FD, Sanders CR (2004) NMR assignments for a helical 40 kDa membrane protein. J Am Chem Soc 126:5048–5049

61. Katz JJ, Crespi HL (1966) Deuterated organisms: cultivation and uses. Science 151:1187–1194
62. Meilleur F, Contzen J, Myles DA, Jung C (2004) Structural stability and dynamics of hydrogenated and perdeuterated cytochrome P450cam (CYP101). Biochemistry 43:8744–8753
63. Brockwell D et al (2001) Physicochemical consequences of the perdeuteriation of glutathione S-transferase from S. japonicum. Protein Sci 10:572–580
64. Schubert M, Smalla M, Schmieder P, Oschkinat H (1999) MUSIC in triple-resonance experiments: amino acid type-selective (1)H-(15)N correlations. J Magn Reson 141:34–43
65. Muchmore DC, McIntosh LP, Russell CB, Anderson DE, Dahlquist FW (1989) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. Methods Enzymol 177:44–73
66. Waugh DS (1996) Genetic tools for selective labeling of proteins with alpha-15N-amino acids. J Biomol NMR 8:184–192
67. LeMaster DM, Kushlan DM (1996) Dynamical mapping of *E. coli* thioredoxin via 13C NMR relaxation analysis. J Am Chem Soc 118:9255–9264
68. Lin MT et al (2011) A rapid and robust method for selective isotope labeling of proteins. Methods 55:370–378
69. Vance CK, Kang YM, Miller AF (1997) Selective 15N labeling and direct observation by NMR of the active-site glutamine of fe-containing superoxide dismutase. J Biomol NMR 9:201–206
70. Maslennikov I et al (2010) Membrane domain structures of three classes of histidine kinase receptors by cell-free expression and rapid NMR analysis. Proc Natl Acad Sci USA 107:10902–10907
71. Sobhanifar S et al (2010) Cell-free expression and stable isotope labelling strategies for membrane proteins. J Biomol NMR 46:33–43
72. Makino S, Goren MA, Fox BG, Markley JL (2010) Cell-free protein synthesis technology in NMR high-throughput structure determination. Methods Mol Biol 607:127–147
73. Reckel S et al (2008) Transmembrane segment enhanced labeling as a tool for the backbone assignment of alpha-helical membrane proteins. Proc Natl Acad Sci USA 105:8262–8267
74. Cellitti SE et al (2008) In vivo incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy. J Am Chem Soc 130:9268–9281
75. Jones DH et al (2010) Site-specific labeling of proteins with NMR-active unnatural amino acids. J Biomol NMR 46:89–100
76. Jackson JC, Hammill JT, Mehl RA (2007) Site-specific incorporation of a (19)F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity. J Am Chem Soc 129:1160–1166
77. Xie J, Schultz PG (2005) Adding amino acids to the genetic repertoire. Curr Opin Chem Biol 9:548–554
78. Xie J, Schultz PG (2005) An expanding genetic code. Methods 36:227–238
79. Gerig JT (1994) Fluorine NMR of proteins. Prog Nucl Magn Reson Spectrosc 26(Part 4):293–370
80. Danielson MA, Falke JJ (1996) Use of 19F NMR to probe protein structure and conformational changes. Annu Rev Biophys Biomol Struct 25:163–195
81. Prosser RS, Luchette PA, Westerman PW (2000) Using O2 to probe membrane immersion depth by 19F NMR. Proc Natl Acad Sci USA 97:9967–9971
82. Kitevski-LeBlanc JL, Evanics F, Prosser RS (2009) Approaches for the measurement of solvent exposure in proteins by 19F NMR. J Biomol NMR 45:255–264
83. Skrisovska L, Schubert M, Allain FH (2010) Recent advances in segmental isotope labeling of proteins: NMR applications to large proteins and glycoproteins. J Biomol NMR 46:51–65
84. Goto NK, Gardner KH, Mueller GA, Willis RC, Kay LE (1999) A robust and cost-effective method for the production of val, leu, ile (delta 1) methyl-protonated 15N-, 13C-, 2H-labeled proteins. J Biomol NMR 13:369–374
85. Janin J, Miller S, Chothia C (1988) Surface, subunit interfaces and interior of oligomeric proteins. J Mol Biol 204:155–164
86. Miller S, Janin J, Lesk AM, Chothia C (1987) Interior and surface of monomeric proteins. J Mol Biol 196: 641–656
87. Miller S, Lesk AM, Janin J, Chothia C (1987) The accessible surface area and stability of oligomeric proteins. Nature 328:834–836
88. Imai S, Osawa M, Takeuchi K, Shimada I (2010) Structural basis underlying the dual gate properties of KcsA. Proc Natl Acad Sci USA 107:6216–6221
89. Hiller S et al (2008) Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. Science 321:1206–1210
90. Verardi R, Shi L, Traaseth NJ, Walsh N, Veglia G (2011) Structural topology of phospholamban pentamer in lipid bilayers by a hybrid solution and solid-state NMR method. Proc Natl Acad Sci USA 108:9101–9106
91. Oxenoid K, Chou JJ (2005) The structure of phospholamban pentamer reveals a channel-like architecture in membranes. Proc Natl Acad Sci USA 102:10870–10875
92. Zhou Y et al (2008) NMR solution structure of the integral membrane enzyme DsbB: functional insights into DsbB-catalyzed disulfide bond formation. Mol Cell 31:896–908

93. Zhou DH et al (2007) Solid-state protein-structure determination with proton-detected triple-resonance 3D magic-angle-spinning NMR spectroscopy. Angew Chem Int Ed Engl 46:8380–8383
94. Hologne M, Faelber K, Diehl A, Reif B (2005) Characterization of dynamics of perdeuterated proteins by MAS solid-state NMR. J Am Chem Soc 127:11208–11209
95. Paulson EK et al (2003) Sensitive high resolution inverse detection NMR spectroscopy of proteins in the solid state. J Am Chem Soc 125:15831–15836
96. Huang KY, Siemer AB, McDermott AE (2011) Homonuclear mixing sequences for perdeuterated proteins. J Magn Reson 208:122–127
97. Reif B et al (2012) Ultra-high resolution in MAS solid-state NMR of perdeuterated proteins: implications for structure and dynamics. J Magn Reson 216:1–12
98. Wickramasinghe NP, Kotecha M, Samoson A, Past J, Ishii Y (2007) Sensitivity enhancement in (13)C solid-state NMR of protein microcrystals by use of paramagnetic metal ions for optimizing (1)H T(1) relaxation. J Magn Reson 184:350–356
99. Akbey U et al (2010) Dynamic nuclear polarization of deuterated proteins. Angew Chem Int Ed Engl 49: 7803–7806
100. Lalli D et al (2011) Three-dimensional deuterium-carbon correlation experiments for high-resolution solid-state MAS NMR spectroscopy of large proteins. J Biomol NMR 51:477–485
101. Gopinath T, Veglia G (2012) Dual acquisition magic-angle spinning solid-state NMR-spectroscopy: simultaneous acquisition of multidimensional spectra of biomacromolecules. Angew Chem Int Ed Engl 51:2731–2735
102. Chekmenev EY et al (2006) Ion-binding study by 17O solid-state NMR spectroscopy in the model peptide gly-gly-gly at 19.6 T. J Am Chem Soc 128:9849–9855
103. Strandberg E et al (2004) Tilt angles of transmembrane model peptides in oriented and non-oriented lipid bilayers as determined by 2H solid-state NMR. Biophys J 86:3709–3721
104. Cady SD, Goodman C, Tatko CD, DeGrado WF, Hong M (2007) Determining the orientation of uniaxially rotating membrane proteins using unoriented samples: a 2H, 13C, AND 15N solid-state NMR investigation of the dynamics and orientation of a transmembrane helical bundle. J Am Chem Soc 129:5719–5729
105. Mani R et al (2006) Membrane-bound dimer structure of a beta-hairpin antimicrobial peptide from rotational-echo double-resonance solid-state NMR. Biochemistry 45:8341–8349
106. Buffy JJ, Waring AJ, Hong M (2005) Determination of peptide oligomerization in lipid bilayers using 19F spin diffusion NMR. J Am Chem Soc 127:4477–4483
107. Kandasamy SK et al (2009) Solid-state NMR and molecular dynamics simulations reveal the oligomeric ion-channels of TM2-GABA(A) stabilized by intermolecular hydrogen bonding. Biochim Biophys Acta 1788: 686–695
108. Liu W, Fei JZ, Kawakami T, Smith SO (2007) Structural constraints on the transmembrane and juxtamembrane regions of the phospholamban pentamer in membrane bilayers: Gln29 and Leu52. Biochim Biophys Acta 1768: 2971–2978
109. Jaroniec CP, MacPhee CE, Astrof NS, Dobson CM, Griffin RG (2002) Molecular conformation of a peptide fragment of transthyretin in an amyloid fibril. Proc Natl Acad Sci USA 99:16748–16753
110. Jaroniec CP et al (2004) High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy. Proc Natl Acad Sci USA 101:711–716
111. Tycko R (2006) Molecular structure of amyloid fibrils: insights from solid-state NMR. Q Rev Biophys 39:1–55
112. Doherty T, Su Y, Hong M (2010) High-resolution orientation and depth of insertion of the voltage-sensing S4 helix of a potassium channel in lipid bilayers. J Mol Biol 401:642–652
113. Gustavsson M, Traaseth NJ, Veglia G (2011) Probing ground and excited states of phospholamban in model and native lipid membranes by magic angle spinning NMR spectroscopy. Biochim Biophys Acta 1818:146–153
114. Vuister GW, Yamazaki T, Torchia DA, Bax A (1993) Measurement of two- and three-bond 13C-1H J couplings to the C delta carbons of leucine residues in staphylococcal nuclease. J Biomol NMR 3:297–306
115. Traaseth NJ, Veglia G (2011) Frequency-selective heteronuclear dephasing and selective carbonyl labeling to deconvolute crowded spectra of membrane proteins by magic angle spinning NMR. J Magn Reson 211:18–24
116. Baldus M, Petkova AT, Herzfeld J, Griffin RG (1998) Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. Mol Phys 95:1197–1207
117. Banigan and Traaseth (2012), J Phys Chem B, 116(24):7138-44
118. McDowell LM, Lee M, McKay RA, Anderson KS, Schaefer J (1996) Intersubunit communication in tryptophan synthase by carbon-13 and fluorine-19 REDOR NMR. Biochemistry 35:3328–3334
119. Krishnarjuna B, Jaipuria G, Thakur A, D'Silva P, Atreya HS (2011) Amino acid selective unlabeling for sequence specific resonance assignments in proteins. J Biomol NMR 49:39–51
120. Vuister GW, Kim S, Wu C, Bax A (1994) 2D and 3D NMR study of phenylalanine residues in proteins by reverse isotopic labeling. J Am Chem Soc 116:9206–9210
121. Bystrov VF (1976) Spinâ€"spin coupling and the conformational states of peptide systems. Prog Nucl Magn Reson Spectrosc 10:41–82

122. Hong M (1999) Resonance assignment of 13C/15N labeled solid proteins by two- and three-dimensional magic-angle-spinning NMR. J Biomol NMR 15:1–14
123. Takeuchi K, Gal M, Takahashi H, Shimada I, Wagner G (2011) HNCA-TOCSY-CANH experiments with alternate (13)C- (12)C labeling: a set of 3D experiment with unique supra-sequential information for mainchain resonance assignment. J Biomol NMR 49:17–26
124. Wand AJ, Bieber RJ, Urbauer JL, McEvoy RP, Gan Z (1995) Carbon relaxation in randomly fractionally 13C-enriched proteins. J Magn Reson B 108:173–175
125. Hong M, Jakes K (1999) Selective and extensive 13C labeling of a membrane protein for solid-state NMR investigations. J Biomol NMR 14:71–74
126. Higman VA et al (2009) Assigning large proteins in the solid state: a MAS NMR resonance assignment strategy using selectively and extensively 13C-labelled proteins. J Biomol NMR 44:245–260
127. Castellani F et al (2002) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. Nature 420:98–102
128. Lundstrom P et al (2007) Fractional 13C enrichment of isolated carbons using [1–13C]- or [2–13C]-glucose facilitates the accurate measurement of dynamics at backbone calpha and side-chain methyl positions in proteins. J Biomol NMR 38:199–212
129. McDermott A (2009) Structure and dynamics of membrane proteins by magic angle spinning solid-state NMR. Annu Rev Biophys 38:385–403
130. Naito A (2009) Structure elucidation of membrane-associated peptides and proteins in oriented bilayers by solid-state NMR spectroscopy. Solid State Nucl Magn Reson 36:67–76
131. Marassi FM et al (2011) Structure determination of membrane proteins in five easy pieces. Methods 55:363–369
132. Bowie JU (2011) Membrane protein folding: how important are hydrogen bonds? Curr Opin Struct Biol 21:42–49
133. Bertram R et al (2003) Atomic refinement with correlated solid-state NMR restraints. J Magn Reson 163:300–309
134. Traaseth NJ et al (2009) Structure and topology of monomeric phospholamban in lipid membranes determined by a hybrid solution and solid-state NMR approach. Proc Natl Acad Sci USA 106:10165–10170
135. Shi L et al (2009) A refinement protocol to determine structure, topology, and depth of insertion of membrane proteins using hybrid solution and solid-state NMR restraints. J Biomol NMR 44:195–205
136. Straus SK, Scott WR, Schwieters CD, Marvin DA (2011) Consensus structure of Pf1 filamentous bacteriophage from X-ray fibre diffraction and solid-state NMR. Eur Biophys J 40:221–234
137. Vostrikov VV, Grant CV, Opella SJ, Koeppe 2nd RE (2011) On the combined analysis of ^{2}H and ^{15}N/^{1}H solid-state NMR data for determination of transmembrane peptide orientation and dynamics, Biophys J 101:2939–2947
138. Wu CH, Ramamoorthy A, Opella SJ (1994) High-resolution heteronuclear dipolar solid-state NMR spectroscopy. J Magn Reson Ser A 109:270–272
139. Sinha N et al (2007) Tailoring 13C labeling for triple-resonance solid-state NMR experiments on aligned samples of proteins. Magn Reson Chem 45(Suppl 1):S107–S115
140. Marassi FM, Opella SJ (2000) A solid-state NMR index of helical membrane protein structure and topology. J Magn Reson 144:150–155
141. Wang J et al (2000) Imaging membrane protein helical wheels. J Magn Reson 144:162–167
142. Page RC, Kim S, Cross TA (2008) Transmembrane helix uniformity examined by spectral mapping of torsion angles. Structure 16:787–797
143. Mote KR et al (2011) Multidimensional oriented solid-state NMR experiments enable the sequential assignment of uniformly (15)N labeled integral membrane proteins in magnetically aligned lipid bilayers. J Biomol NMR 51:339–346
144. Knox RW, Lu GJ, Opella SJ, Nevzorov AA (2010) A resonance assignment method for oriented-sample solid-state NMR of proteins. J Am Chem Soc 132:8255–8257
145. Lu GJ, Son WS, Opella SJ (2011) A general assignment method for oriented sample (OS) solid-state NMR of proteins based on the correlation of resonances through heteronuclear dipolar couplings in samples aligned parallel and perpendicular to the magnetic field. J Magn Reson 209:195–206
146. Nevzorov AA, Opella SJ (2003) Structural fitting of PISEMA spectra of aligned proteins. J Magn Reson 160: 33–39
147. Asbury T et al (2006) PIPATH: an optimized algorithm for generating alpha-helical structures from PISEMA data. J Magn Reson 183:87–95
148. Shi L et al (2011) Paramagnetic-based NMR restraints lift residual dipolar coupling degeneracy in multidomain detergent-solubilized membrane proteins. J Am Chem Soc 133:2232–2241
149. Jones DH, Barber KR, VanDerLoo EW, Grant CW (1998) Epidermal growth factor receptor transmembrane domain: 2H NMR implications for orientation and motion in a bilayer environment. Biochemistry 37:16780–16787
150. Whiles JA et al (2001) Orientation and effects of mastoparan X on phospholipid bicelles. Biophys J 80:280–293

151. van der Wel PC, Strandberg E, Killian JA, Koeppe 2nd RE (2002) Geometry and intrinsic tilt of a tryptophan-anchored transmembrane alpha-helix determined by (2)H NMR. Biophys J 83:1479–1488
152. Strandberg E, Wadhwani P, Tremouilhac P, Durr UH, Ulrich AS (2006) Solid-state NMR analysis of the PGLa peptide orientation in DMPC bilayers: structural fidelity of 2H-labels versus high sensitivity of 19F-NMR. Biophys J 90:1676–1686
153. Resende JM et al (2009) Membrane structure and conformational changes of the antibiotic heterodimeric peptide distinctin by solid-state NMR spectroscopy. Proc Natl Acad Sci USA 106:16639–16644
154. Vostrikov VV, Hall BA, Greathouse DV, Koeppe 2nd RE, Sansom MS (2010) Changes in transmembrane helix alignment by arginine residues revealed by solid-state NMR experiments and coarse-grained MD simulations. J Am Chem Soc 132:5803–5811
155. Bertelsen K et al (2011) Long-term-stable ether-lipid vs conventional ester-lipid bicelles in oriented solid-state NMR: altered structural information in studies of antimicrobial peptides. J Phys Chem B 115:1767–1774
156. Davis JH, Maraviglia B, Weeks G, Godin DV (1979) Bilayer rigidity of the erythrocyte membrane2H-NMR of a perdeuterated palmitic acid probe. Biochim Biophys Acta 550:362–366
157. Larsen FH, Jakobsen HJ, Ellis PD, Nielsen NC (1998) QCPMG-MAS NMR of half-integer quadrupolar nuclei. J Magn Reson 131:144–147
158. Killian JA, Taylor MJ, Koeppe 2nd RE (1992) Orientation of the valine-1 side chain of the gramicidin transmembrane channel and implications for channel functioning. A ^{2}H NMR study. Biochemistry 31:11283–11290
159. Thomas R, Vostrikov VV, Greathouse DV, Koeppe 2nd RE (2009) Influence of proline upon the folding and geometry of the WALP19 transmembrane peptide. Biochemistry 48:11883–11891
160. Abu-Baker S et al (2007) Side chain and backbone dynamics of phospholamban in phospholipid bilayers utilizing 2H and 15N solid-state NMR spectroscopy. Biochemistry 46:11695–11706
161. Vold RL, Hoatson GL (2009) Effects of jump dynamics on solid state nuclear magnetic resonance line shapes and spin relaxation times. J Magn Reson 198:57–72
162. Vugmeyster L et al (2011) Slow motions in the hydrophobic core of chicken villin headpiece subdomain and their contributions to configurational entropy and heat capacity from solid-state deuteron NMR measurements. Biochemistry 50:10637–10646
163. van der Wel PC, Reed ND, Greathouse DV, Koeppe 2nd RE (2007) Orientation and motion of tryptophan interfacial anchors in membrane-spanning peptides. Biochemistry 46:7514–7524
164. Liu W, Crocker E, Siminovitch DJ, Smith SO (2003) Role of side-chain conformational entropy in transmembrane helix dimerization of glycophorin A. Biophys J 84:1263–1271
165. Struppe J, Komives EA, Taylor SS, Vold RR (1998) 2H NMR studies of a myristoylated peptide in neutral and acidic phospholipid bicelles. Biochemistry 37:15523–15527
166. Gaffarogullari EC et al (2011) A myristoyl/phosphoserine switch controls cAMP-dependent protein kinase association to membranes. J Mol Biol 411:823–836
167. Koch K, Afonin S, Ieronimo M, Berditsch M, Ulrich AS (2012) Solid-state (19)F-NMR of peptides in native membranes. Top Curr Chem 306:89–118
168. Luo W, Mani R, Hong M (2007) Side-chain conformation of the M2 transmembrane peptide proton channel of influenza a virus from 19F solid-state NMR. J Phys Chem B 111:10825–10832
169. Young TS, Schultz PG (2010) Beyond the canonical 20 amino acids: expanding the genetic lexicon. J Biol Chem 285:11039–11044
170. Traaseth NJ, Verardi R, Veglia G (2008) Asymmetric methyl group labeling as a probe of membrane protein homo-oligomers by NMR spectroscopy. J Am Chem Soc 130:2400–2401
171. Walters KJ et al (2001) Characterizing protein-protein complexes and oligomers by nuclear magnetic resonance spectroscopy. Methods Enzymol 339:238–258
172. Yang J, Tasayco ML, Polenova T (2008) Magic angle spinning NMR experiments for structural studies of differentially enriched protein interfaces and protein assemblies. J Am Chem Soc 130:5798–5807
173. Kosen PA (1989) Spin labeling of proteins. Methods Enzymol 177:86–121
174. Nelson DJ (1978) Fluorine-19 magnetic resonance of muscle calcium binding parvalbumin: PH dependency of resonance position and spin–lattice relaxation time. Inorg Chim Acta 27:L71–L74
175. Klein-Seetharaman J, Getmanova EV, Loewen MC, Reeves PJ, Khorana HG (1999) NMR spectroscopy in studies of light-induced structural changes in mammalian rhodopsin: applicability of solution (19)F NMR. Proc Natl Acad Sci USA 96:13744–13749
176. Adriaensens P et al (1988) Investigation of protein structure by means of 19F-NMR. A study of hen egg-white lysozyme. Eur J Biochem 177:383–394
177. Mehta VD, Kulkarni PV, Mason RP, Constantinescu A, Antich PP (1994) Fluorinated proteins as potential 19F magnetic resonance imaging and spectroscopy agents. Bioconjug Chem 5:257–261
178. Donald LJ, Crane BR, Anderson DH, Duckworth HW (1991) The role of cysteine 206 in allosteric inhibition of Escherichia coli citrate synthase. studies by chemical modification, site-directed mutagenesis, and 19F NMR. J Biol Chem 266:20709–20713

179. Phillips L, Separovic F, Cornell BA, Barden JA, dos Remedios CG (1991) Actin dynamics studied by solid-state NMR spectroscopy. Eur Biophys J 19:147–155
180. Brauer M, Sykes BD (1986) 19F nuclear magnetic resonance studies of selectively fluorinated derivatives of G- and F-actin. Biochemistry 25:2187–2191
181. Kay LE, Pascone JM, Sykes BD, Shriver JW (1987) 19F nuclear magnetic resonance as a probe of structural transitions and cooperative interactions in heavy meromyosin. J Biol Chem 262:1984–1988
182. Evanics F, Kitevski JL, Bezsonova I, Forman-Kay J, Prosser RS (2007) 19F NMR studies of solvent exposure and peptide binding to an SH3 domain. Biochim Biophys Acta 1770:221–230
183. Liu JJ, Horst R, Katritch V, Stevens RC, Wuthrich K (2012) Biased signaling pathways in beta2-adrenergic receptor characterized by 19F-NMR. Science 335:1106–1110
184. Religa TL, Ruschak AM, Rosenzweig R, Kay LE (2011) Site-directed methyl group labeling as an NMR probe of structure and dynamics in supramolecular protein systems: applications to the proteasome and to the ClpP protease. J Am Chem Soc 133:9063–9068
185. Richards PG, Coles B, Heptinstall J, Walton DJ (1994) Electrochemical modification of lysozyme: anodic reaction of tyrosine residues. Enzyme Microb Technol 16:795–801
186. Hebel D, Kirk KL, Cohen LA, Labroo VM (1990) First direct fluorination of tyrosine-containing biologically active peptides. Tetrahedron Lett 31:619–622
187. Abraham SJ, Hoheisel S, Gaponenko V (2008) Detection of protein-ligand interactions by NMR using reductive methylation of lysine residues. J Biomol NMR 42:143–148
188. Ivan R (1997) Macromolecular crystallography part A. In: Charles W, Carter J (eds) Methods in enzymology. Academic Press, New York, pp 171–179
189. Bokoch MP et al (2010) Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. Nature 463:108–112
190. Su XC, Otting G (2010) Paramagnetic labelling of proteins and oligonucleotides for NMR. J Biomol NMR 46:101–112
191. Berardi MJ, Shih WM, Harrison SC, Chou JJ (2011) Mitochondrial uncoupling protein 2 structure determined by NMR molecular fragment searching. Nature 476:109–113
192. Trad CH, James W, Bhardwaj A, Butterfield DA (1995) Selective labeling of membrane protein sulfhydryl groups with methanethiosulfonate spin label. J Biochem Biophys Methods 30:287–299
193. Hubbell WL, Gross A, Langen R, Lietzow MA (1998) Recent advances in site-directed spin labeling of proteins. Curr Opin Struct Biol 8:649–656