



Labeling of methyl groups: a streamlined protocol and guidance for the selection of ^2H precursors based on molecular weight

Alexandra Locke¹ · Kylee Guarino¹ · Gordon S. Rule¹

Received: 21 November 2023 / Accepted: 21 March 2024 / Published online: 24 May 2024
© The Author(s) 2024

Abstract

A streamlined one-day protocol is described to produce isotopically methyl-labeled protein with high levels of deuterium for NMR studies. Using this protocol, the D_2O and ^2H -glucose content of the media and protonation level of ILV labeling precursors (ketobutyrate and ketovalerate) were varied. The relaxation rate of the multiple-quantum (MQ) state that is present during the HMQC-TROSY pulse sequence was measured for different labeling schemes and this rate was used to predict upper limits of molecular weights for various labeling schemes. The use of deuterated solvents (D_2O) or deuterated glucose is not required to obtain ^1H – ^{13}C correlated NMR spectra of a 50 kDa homodimeric protein that are suitable for assignment by mutagenesis. High quality spectra of 100–150 kDa proteins, suitable for most applications, can be obtained without the use of deuterated glucose. The proton on the β -position of ketovalerate appears to undergo partial exchange with deuterium under the growth conditions used in this study.

Keywords Biotechnology · Biophysics · Bacterial expression · NMR · Protein structure

Introduction

Larger biomolecular systems have been accessible to NMR studies by the advent of deuteration, TROSY techniques, and in the case of proteins, methyl-specific labeling. Deuteration reduces spin–spin relaxation, allowing transfer of coherence over longer distances (Sattler and Fesik 1996). TROSY techniques select for the more slowly decaying excited state, enhancing resolution and sensitivity (Pervushin et al. 1997; Tugarinov and Kay 2004). In addition, the presence of three hydrogens on the methyl group, coupled with rapid methyl rotation, further enhances the signal from these groups. Methyl-specific labeling, typically Ile, Val, and Leu (ILV) residues, provides suitable spin probes to measure structure and dynamics at many sites in most proteins (Goto et al. 1999). Additional information can be obtained from methyl labeled Ala (Ayala et al. 2009) and Thr residues (Velyvis et al. 2012). High levels of deuteration of methyl-labeled samples increases sensitivity and resolution by reducing methyl proton and carbon relaxation rates (Goto et al. 1999).

An important requirement for the interpretation of NMR-derived properties, such as structural information, dynamics, and ligand induced chemical shifts, is the assignment of a resonance peak to an individual group in the protein. In the case of methyl resonances, several approaches have been successfully used (Pritišanac et al. 2020; Clay et al. 2022). NMR based approaches include: (i) spin–spin coherence transfer (Tugarinov and Kay 2003), (ii) through-space dipolar coupling (Nerli et al. 2021), (iii) paramagnetic enhancement (PRE) using nitroxides or lanthanides (Pritišanac et al. 2020), and (iv) pseudo contact shifts from lanthanides with anisotropic magnetic susceptibility tensors (John et al. 2007; Lescanne et al. 2017). Coherence transfer methods require mainchain assignments and becomes more difficult to apply for larger proteins (~50 kDa). Although dipolar coupling approaches become more efficient in larger systems, they rely on a known structure of the protein and methyl groups that are distant from other methyl groups cannot be assigned using this approach. Paramagnetic enhancement can provide useful assignment information; however, it can be difficult to estimate the position of the unpaired electron on nitroxides due to flexibility of the paramagnetic group. Lanthanide based pseudoccontact shifts provide very useful orientation information for assignments beyond PREs but require the creation of a lanthanide binding site.

✉ Gordon S. Rule
rule@andrew.cmu.edu

¹ Department of Biological Sciences, Carnegie Mellon University, 4400 5th Ave, Pittsburgh, PA 15213, USA

Another common approach for methyl assignments is site-directed mutagenesis (Amero et al. 2011; Crublet et al. 2014). In this case individual Ile residues are replaced by Val, Val by Ala, and Leu by Ala. Replacement of Ala and Thr residues by Gly and Ser respectively also provide an efficient way to assign methyl resonances from those residues. The mutation causes the signal from the mutated residue to either move to a different region of the spectrum (e.g. Ile to Val) or disappear entirely from the spectrum (e.g. Thr to Ser). This approach is very robust for Ile since the replacement of Ile by Val generally has little effect on protein structure (Trivedi and Nagarajaram 2020) and therefore little effect on the chemical shifts of the remaining Ile residues. This approach can also be very effective for Thr, Val, Leu, and Ala residues with the judicious selection of mutation sites and type of substitution, based on structure-based predictions (e.g. molecular dynamics, MD) of the effect of individual substitutions on the protein structure. For example, although Thr to Ser replacements are conservative, MD calculations may show that replacement of Thr by Val, Ala, or Gly may be less perturbing at some sites.

With the advent of expedient, automatable, and low-cost methods of gene synthesis as well as robust methods for chemical lysis of cells (Listwan et al. 2010), it is practical to generate a large panel of mutants for methyl assignment purposes that would include all Ile residues and a substantial number of Thr, Val, Leu, and Ala residues. Naturally, methyl labeled protein must be generated for each of these mutants. Consequently, it is advantageous to streamline the protein production pipeline and reduce the cost of labeled precursors.

Several approaches have been published that provide workflows to produce deuterated proteins (Crublet et al. 2014; Li and Byrd 2022; Cai et al. 2021, 2016). Although these workflows are reliable, they are somewhat time consuming to produce many samples, often requiring sequential adaptation to different media conditions with centrifugation. We present a more streamlined protocol for the adaptation of cells to deuterated media that allows adaptation and protein induction in one day, with the subsequent harvest of the cells the following day. Centrifugation of the culture during this process is not required unless very high levels (> 97.5%) of D₂O are required. This streamlined approach facilitates a high-throughput approach to the production of an array of mutant proteins for assignment or other mutational mapping purposes.

An important consideration in the generation of multiple samples is the efficient use of isotopes (D₂O, deuterated glucose, ILV methyl precursors) that are required to make highly deuterated samples. For example, Clore has shown that high levels of deuterated glucose can be used to obtain very dense cell cultures post-induction, reducing the deuterium requirements because of increased protein yield per

volume of culture (Cai et al. 2016). A simple way to reduce isotope costs is to reduce the deuterium content of the solvent (H₂O versus D₂O), glucose, and precursors. We have systematically explored the reduction of deuterium content of the solvent, glucose, and ketoacid precursors on the quality of ¹H–¹³C correlation spectra with the goal to reduce cost by establishing guidelines for the labeling of moderate-to-large sized proteins (e.g. 25–250 kDa). We show that it is possible to obtain 2D ¹H–¹³C spectra of a 50 kDa protein that are suitable for resonance assignment purposes by growing the cells in the presence of 100% H₂O and protonated glucose, with the only source of deuterium from highly deuterated ¹³C precursors. As the protein size increases, a higher percentage of D₂O and deuterated glucose are required. Guidelines are provided to select deuterium levels that are appropriate for different molecular weights. We show that the use of deuterated glucose is not required unless the molecular weight of the protein exceeds ~ 125 kDa; high quality HMQC spectra can be obtained from large proteins that are obtained from cultures grown in 90% D₂O and 100% ¹H-glucose.

Results and discussion

Streamlined process for cell growth

A rapid and reliable method of obtaining deuterated protein was developed by combining adaptation with a change in media from rich to minimal media. The overall strategy is shown in Fig. 1. Overnight growth of cells in 50% D₂O while still in 50% rich media (LB), followed by a short 4 h growth phase at increased D₂O/decreased LB levels the next morning allows the cells to adapt more quickly to growth at higher D₂O levels within the same day. The level of deuteration of the protein in the purified protein depends on the deuterium content of the solvent (D₂O) and whether protonated or deuterated glucose is used (Table 1). Higher levels of deuteration are achieved by increasing the level of D₂O as well as including deuterated glucose, at the expense of an increased time to protein induction. D₂O levels as high as 97.5% can be obtained without any centrifugation step. If higher levels of D₂O are required, it is necessary to centrifuge the cells and resuspend at the required D₂O level. The overall growth protocol is also suitable for Ala and Thr methyl labeling in combination with ILV labeling.

Effect of solvent and glucose deuterium levels

Table 1 shows the effect of D₂O level and protonation/deuteration state of glucose on the level of incorporated deuterium, growth time, and the MQ relaxation rate. The MQ relaxation rate was used as a proxy for overall sensitivity. This rate was measured when the magnetization

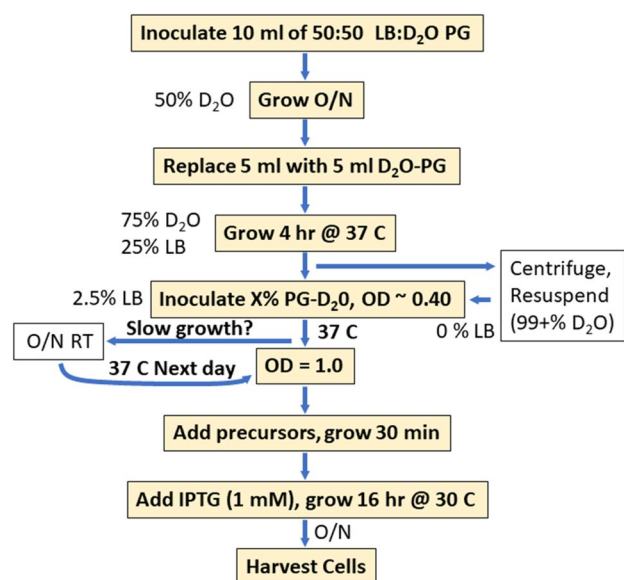


Fig. 1 Process for robust growth in high levels of D₂O. The flow chart indicates the key steps in the production of deuterated protein. The optional path on the right provides a higher level of deuteration but requires centrifugation and results in a longer growth period on the same day. The optional path (open box) on the left is for occasional slow growth that prohibits induction on the same day. O/N overnight

for both spins was transverse, i.e. essentially during the carbon evolution period (see Fig. S3). This permitted the direct measurement of the effect of surrounding protons and deuterons on the relaxation rate of the ¹³CH₃ methyl group. Analysis of the relaxation of CH₃ groups during a

HMQC sequence in large proteins by Ollerenshaw et al. (2003) showed that the signal intensity is:

$$I \propto e^{(-[4\tau+t_2]R_{2,H}^S - t_1R_{2,CH}^S)}$$

The proton relaxation rate, $R_{2,H}^S = R_{CH} + R_{ext}$, where R_{CH} is the relaxation rate due to intra-methyl proton-carbon dipolar coupling and R_{ext} is the rate due to coupling to surrounding protons and deuterons. In contrast, $R_{2,CH}^S$ (=MQ relaxation rate) only reflects relaxation from adjacent protons and deuterons due to complete cancellation of intra-methyl relaxation due to the TROSY effect (Ollerenshaw et al. 2003). Consequently, $R_{2,CH}^S$ is more reflective of the effect of residual protons on the overall sensitivity of the experiment because it directly determines the linewidth in ω_1 and it contributes to the linewidth in ω_2 and signal loss during magnetization transfer (4τ period), via R_{ext} .

We investigated how varying the deuterium level of both D₂O and glucose, from 0 to 90% (rows 1–5, Table 1), affected the quality of the spectra (Fig. 2) and the MQ relaxation rate. Figures 2, 7 and S4 show that samples grown in 0% D₂O and protonated glucose give spectra with adequate resolution and sensitivity to easily detect all the peaks in the spectra of the 50 kDa A1-1 protein used in this study. The sensitivity of detection could be improved by acquiring more scans over a shorter carbon evolution period or using SOFAST techniques (Schanda et al. 2005), but these approaches were not explored here. Increasing the D₂O content and ²H-glucose content increases the quality of the spectra (Fig. 2), as well as decreasing the MQ relaxation rate (Table 1), as expected.

Table 1 Effect of growth conditions on deuteration levels, growth time, and upper limit for molecular weight

Growth condition		² H level (%)	Growth time to OD 1 (h) ^a	MW kDa (upper limit)	MQ R ₂ (s ⁻¹)
D ₂ O%	Glucose				
0	¹ H/ ² H 100:0	0%	2.0	61	49.0
25	¹ H/ ² H 75:25	11%	N.D	75	40.0
50	¹ H/ ² H 50:50	30%	N.D	86	34.9
75	¹ H/ ² H 25:75	60%	N.D	102	29.4
90	¹ H/ ² H 10:90	81%	4.0	154	19.5
90 ^b	¹ H/ ² H 10:90(LB)	81%	4.0	146	20.5
90	¹ H 100%	71%	3.7	132	22.7
99.0 ^c	² H	96%	5.7	192	15.6
97.5 ^d	² H (LB)	93%	4.7	186	16.1
99.0 ^c	¹ H	83%	5.3	162	18.5

All samples were labeled using ketobuturate that was deuterated on the β-position and ketovalerate that was deuterated on the β-position and on one of the methyl groups

^aAssuming starting OD=0.4

^b2.5% LB remaining (no centrifugation)

^cCentrifugation required to attain level of D₂O content

^d2.5% LB remaining, highest D₂O level without centrifugation

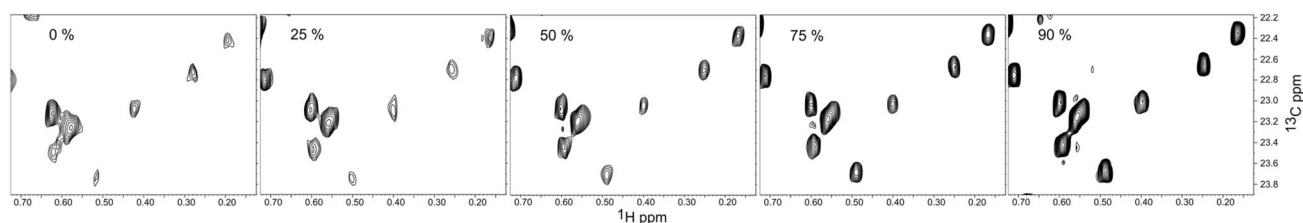


Fig. 2 Effect of deuteration level on the quality of the HMQC ILV spectra. A region of the spectra containing Leu/Val peaks is shown for proteins purified after expression in the indicated D₂O concentrations

The level of sensitivity and resolution that is obtained with growth in 100% H₂O is because the heavily deuterated precursors that were used place deuterium atoms in favorable locations to reduce spin–spin coupling to the methyl protons. The hydrogens on the same residue are among those closest to the labeled methyl and these have been replaced by deuterons. In addition, the hemi-deuterated methyl groups on Val and Leu reduce the overall proton density within the core of the protein. Calculations based on local proton density indicate that deuteration of the γ_1 -protons on Ile reduce the relaxation rate of the Ile δ methyl by approximately 30%. In the case of Val and Leu residues, the deuteration of the β/γ proton reduces the relaxation rate by 15% and deuteration of one methyl reduces the rate by an additional 30%. This analysis is consistent with results presented by Dubey et al. (2021) who showed that the local deuteration of leucine reduces the relaxation rate of leucine methyl groups in proteins. The principal application of Dubey's approach is for labeling with yeast and mammalian cells expression systems. The work presented here indicates that the synthesis of similarly labeled Val and Ile would also be advantageous in those systems.

The MQ-relaxation rates that were obtained under different labeling schemes were used to predict the approximate upper molecular weight limit for that scheme (see Fig. 8). We used a heuristic approach to relate the MQ relaxation rate for a particular labeling scheme to a suggested molecular weight limit. We assumed an average MQ-relaxation rate of $\sim 60 \text{ s}^{-1}$ would give spectra of suitable sensitivity and dispersion. This value was chosen because the relaxation rate for ILV resonances in A1-1 grown with ¹H glucose and H₂O showed a median relaxation rate of 49 s^{-1} for most residues while weak resonances in that sample had relaxation rates $> 70 \text{ s}^{-1}$.

To convert MQ relaxation rates to effective molecular weight we associated the τ_c for A1-1 in D₂O at 300 K ($\sim 30 \text{ ns}$), with the observed relaxation rate (R) for a particular labeling scheme. The τ_c that would result in a rate of 60 s^{-1} under those conditions was simply $30 \text{ ns} \times 60 \text{ s}^{-1}/R$. The τ_c was converted to molecular weight by dividing by 0.6 (Rossi et al. 2010). For example, the sample grown in ¹H glucose and H₂O gave a relaxation rate of 49 s^{-1} for

A1-1. Therefore, a relaxation rate of 60 s^{-1} corresponds to a τ_c of 36.7 ns ($= 30 \text{ ns} \times 60/49$), corresponding to a MW of 61 kDa. The slowest relaxation rate we measured, with growth in 99% D₂O and ²H-glucose was 15.6 s^{-1} . The τ_c to give a rate of 60 s^{-1} under that labeling condition is $30 \times 60/15.6 = 115 \text{ ns}$, corresponding to a MW of 192 kDa.

The process outlined in Fig. 1 will only require a centrifugation step if high levels of D₂O are required. We found that the residual LB media (2.5%) during the growth and induction phase has little effect on the deuterium level in the sample as well as the MQ relaxation rate. Figure 3 shows the MQ relaxation rates for Ile residues, without and with LB (first two bars for each residue). The relaxation rates of individual methyl groups are essentially the same for both conditions. A similar trend is also seen in 99% D₂O and for Val & Leu residues at both D₂O levels (not shown).

We also determined which source of deuterons (D₂O or glucose) had a larger impact on the quality of the spectra and found that D₂O was more important. The replacement of ²H-glucose with ¹H-glucose at 90% and 99% D₂O levels leads to minor increases in the MQ relaxation rate ($\sim 10\%$, Fig. 3, last bar of each group) and consequently minor decreases in the upper limit for molecular weight (see Table 1). For example, growth in 90% D₂O and ¹H-glucose should give acceptable spectra for proteins in the

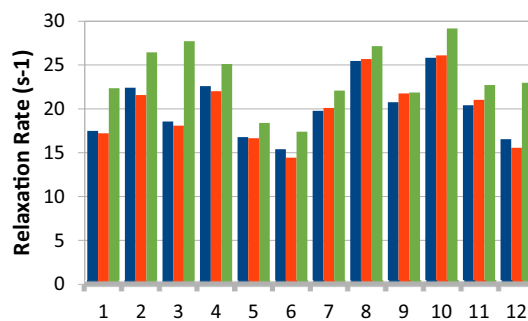


Fig. 3 Effect of Residual LB and Glucose source on MQ relaxation rates for individual Ile residues. Cells were grown in 90% D₂O in the presence of 90% ²H-glucose (blue, first bar), with residual 2.5% LB (red, second bar), or with 90% D₂O and 100% ¹H-glucose (green, third bar)

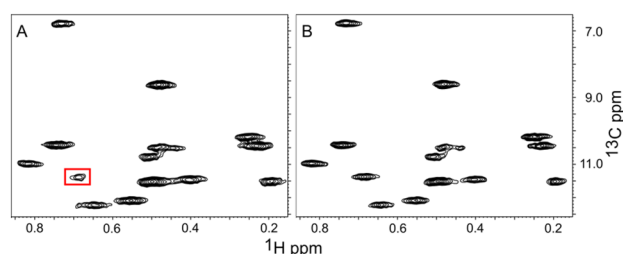


Fig. 4 Effect of Glucose Deuteration on Quality of spectra. Cells were grown in 90% D₂O in the presence of 100% ¹H-glucose (A) or 90% ²H-glucose (B). Spectra were measured in the presence of 20% sucrose at 285 K to increase the apparent molecular weight to ~150 kDa. The boxed resonance in panel A is from an Ile that relaxes faster due to a higher local ¹H density. Regardless, it is clearly observable in the spectra. Otherwise, the spectra and relaxation rates are similar

range of ~125 kDa. To demonstrate that this is the case we acquired HMQC spectra of A-11 in the presence of 20% sucrose at 285 K. These conditions should increase the rotational correlation time from ~28 ns to ~90 ns, increasing the apparent molecular weight to ~150 kDa. Figure 4 shows the Ile region of the HMQC spectra for samples grown in 90% D₂O in ¹H-glucose or ²H-glucose. The spectra show similar resolution and peak intensity. One Ile residue shows reduced intensity in the ¹H-glucose sample due to more rapid relaxation from a higher-than-average local proton density, but it is still clearly visible in the spectra.

The median MQ relaxation rates for the sample grown in ¹H-glucose were 41 s⁻¹ and 46 s⁻¹ for Val/Leu and Ile, respectively. Growth in ²H-glucose gave MQ relaxation rates of 36 s⁻¹ and 39 s⁻¹ for Val/Leu and Ile, respectively. The weaker peak in the sample obtained using ¹H-glucose (boxed in Fig. 4A) had a relaxation rate of 67 s⁻¹, suggesting that it may be difficult to obtain signals from all groups using ¹H-glucose (+ 90% D₂O) when labeling a 150 kDa protein. This conclusion is consistent with our suggested upper limit of 132 kDa that was predicted from a relaxation rate of 22.7 s⁻¹ measured in D₂O (no glycerol) at 300 K (see Table 1).

Growth in 100% ¹H-glucose increased the overall protonation level in the sample but showed a smaller than expected increase in the MQ relaxation rate (see Table 1). The additional protons from ¹H-glucose appear to be selectively incorporated into aromatic sites as opposed to aliphatic sites. Figure 5 shows the 1D-¹H spectra of samples grown in H₂O, 99% D₂O + ¹H-glucose, 99% D₂O with ²H-glucose. Selective incorporation into aromatics is a consequence of the pathways for biosynthesis which utilize precursors derived directly from glucose via the pentose pathway while aliphatic residues would use glucose-derived carbons that have undergone H–D exchange during central metabolism (Rosen et al. 1996). The relatively small increase in MQ-relaxation rate for the ¹H-glucose derived samples may be

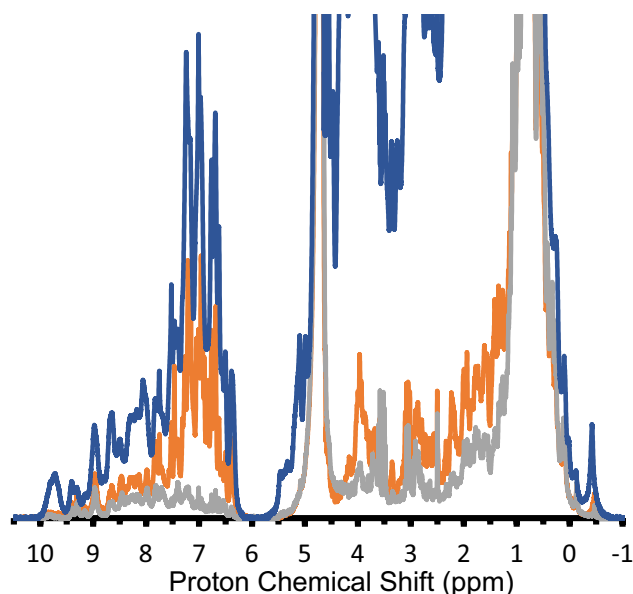


Fig. 5 Effect of Glucose Deuteration on Aromatic and Aliphatic Labeling. Cells were grown in 100% H₂O ¹H-glucose (blue), 99% D₂O with either ¹H glucose (orange) or ²H-glucose (gray)

due to rapid ring flips of Phe and Tyr residues which would reduce dipolar coupling between those aromatics and ILV methyl groups.

Effect of deuterium labeling of ILV precursors

Ketobutyrate and ketovalerate precursors are commercially available with different levels of deuteration (See Scheme 1, supplementary material). The β-hydrogens on ketobutyrate can be protons (HKB) or deuterons (DKB). Ketovalerate can be purchased with both methyls labeled and either a proton (HKV2) or a deuteron at the β-position (DKV2). In addition, ketovalerate with a single ¹³C labeled methyl is available, with all other hydrogens deuterated (KV). We generated protein samples from growth in either 100% H₂O + 100% ¹H-glucose, or 90% D₂O + 90% ²H-Glucose, or 90% D₂O + 100% ¹H-Glucose, varying the Ile and Leu/Val precursor (e.g. HKB + HKV2, HKB + DKV2, HKB + KV, DKB + HKV2, DKB + DKV2, DKB + KV).

A concern with the use of ketoacid precursors for labeling is chemical exchange of the hydrogens on the acidic C_β position on the precursor, e.g. HKB becomes DKB when used with high levels of D₂O in the media. We see no evidence of proton-deuterium exchange for HKB during growth, based on deuterium isotope shifts (see Fig. S4). The ²H isotope shift is – 0.18 ppm (¹³C) from HKB to DKB for samples grown in D₂O. This is consistent with the reported two-bond shift of –0.09 ppm per deuteron (Maltsev et al. 2012), suggesting that the CH₂ group at the C_γ position on Ile remains protonated when HKB is used in D₂O. When DKB is used

in conjunction with $\text{H}_2\text{O}/^1\text{H}$ -glucose we find a small isotope shift -0.05 ppm. This is presumably a three-bond effect due to the incorporation of a proton at C_β during the synthesis of Ile from KB. It is important to note that the source of these protons is from the solvent, not glucose. Growth in $\text{D}_2\text{O} + ^1\text{H}$ -glucose does not produce an isotope shift.

In the case of ketovalerate labeling (HKV2, DKV2, KV) there appears to be partial exchange of the proton on the β position of HKV2 for deuterium when cells are grown in D_2O . Figure 6 shows an overlay of spectra obtained from samples grown with different ketovalerate precursors in D_2O . The resonances obtained from HKV2 labeled samples consist of two overlapping peaks. The lower peak corresponds to isotope shift of $+0.18$ ppm relative to the KV peak, which is consistent with 3 deuterons 3 bonds away, and one deuteron that is two bonds away. The upper HKV2 peak has the same shift as the peak for the DKV2 sample, suggesting that the proton at the C_β position of HKV2 can exchange for a deuteron. The level of exchange is approximately $\sim 50\%$, based on the intensity of each peak. The exchange of H for D in HKV2 also occurs when cells are grown in $90\% \text{D}_2\text{O}$ and $100\% ^1\text{H}$ -glucose, indicating that the source of deuterons is the solvent. The β -deuteron on DKV2

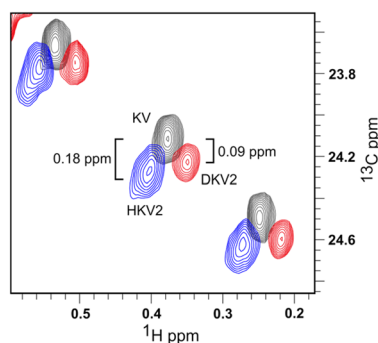


Fig. 6 Deuterium-proton exchange of HKV2. Spectra obtained from cells grown in $90\% \text{D}_2\text{O} + 90\% ^2\text{H}$ glucose in the presence of DKB using different ketovalerate sources. Methyl signals from cells grown in HKV2 show the presence of two overlapping resonances, the lower shifted 0.18 ppm from peaks in the sample grown in KV. The peaks have been displaced laterally to make it easier to see the lineshapes

and KV appear to be resistant to exchange with protons in H_2O since the signals in the DKV2 and KV samples are single peaks and the deuterium isotope shift from HKV2 to KV is 0.18 ppm (not shown). The higher exchange rate of the β -proton versus the β -deuteron is consistent with the higher stability of C-D bonds (Wiberg 1955; Martino et al. 2023).

Predicted and measured relaxation rates indicate that the relaxation of methyls on Leu and Val are not greatly affected by protonation or deuteration at the γ -position of Ile. Likewise, the protonation/deuteration state of the β/γ -position on Leu and Val does not have a large effect on the relaxation of the Ile C_δ methyl. When only one of the methyls of Leu and Val is labeled (KV precursor), there is a modest reduction in the rate of relaxation of Ile C_δ methyls. In the case of DKB labeling in H_2O , the average relaxation rate of Ile C_δ was 52.5 s^{-1} versus 51.1 s^{-1} for DKV2 versus KV, respectively. In the case of DKB labeling in $90\% \text{D}_2\text{O}$, the average relaxation rate of Ile C_δ was 20.4 s^{-1} versus 19.2 s^{-1} for DKV2 versus KV, respectively. Since the change in relaxation rates between DKV2 and KV were small ($\sim 5\%$) we considered each type of labeling (Ile versus Leu/Val) separately when evaluating relaxation rates.

Table 2 gives the measured relaxation rates and estimated upper limit of molecular weight for combinations of growth media (H_2O versus $90\% \text{D}_2\text{O}$) and precursor. Note that the relaxation rates given for HKV2 in D_2O are the average of HKV2 and DKV2 due to partial exchange of the proton with deuterium on HKV2. Figure 7 shows the Ile region of spectra obtained under the four KB labeling schemes presented in Table 2. The quality of the Ile spectrum mirrors the relaxation rate, e.g. $\text{HKB} + \text{H}_2\text{O} < \text{HKB} + \text{D}_2\text{O} \approx \text{DKB} + \text{H}_2\text{O} < \text{DKB} + \text{D}_2\text{O}$. The quality of Leu + Val spectra (Fig. S2) also show a similar trend.

The different relaxation rates that are obtained for different D_2O content, glucose labeling, and ketobutyrate and ketovalerate precursors suggest guidelines for isotopic labeling requirements as a function of MW of the protein (see Fig. 8). For example, the suggested labeling scheme for a protein with a molecular weight of 100 kDa should be $90\% \text{D}_2\text{O}$, ^1H -Glucose, KV and DKB.

Table 2 Effect of deuteration pattern on relaxation rate, upper limit for molecular weight

Growth Media	HKB		DKB		HKV2		DKV2		KV	
	R2 (s^{-1})	MW	R2 (s^{-1})	MW	R2 (s^{-1})	MW	R2 (s^{-1})	MW	R2 (s^{-1})	MW
H_2O	113	27 kDa	49	61 kDa	99	30 kDa	67	45 kDa	49	61 kDa
D_2O (90%)	67	45 kDa	19	158 kDa	58	52 kDa	43	70 kDa	19	158 kDa

All samples were labeled with both a ketobutyrate and a ketovalerate precursor. Deuterated ketobutyrate (DKB) was used for all of the ketovalerate data in this table. The R2 rates refer to Ile peaks for ketobutyrate entries and to Val and Leu peaks for HKV2, DKV2, KV entries. HKB has the β -hydrogens protonated, DKB has the β -hydrogens deuterated. HKV2 has both methyls labeled and the β -hydrogen protonated. DKV2 has both methyl groups labeled and the β -hydrogen deuterated. KV has only one methyl group labeled; all remaining hydrogens are deuterated

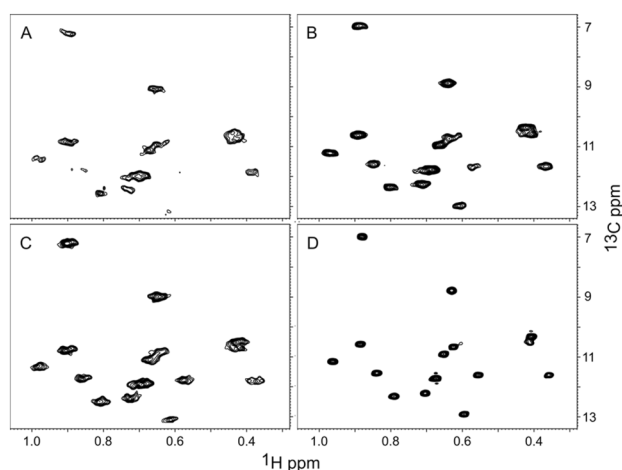


Fig. 7 Effect of solvent and protonation of Ketobutyrate on Quality of Ile Spectra. Cells were grown in H₂O + HKB (A), H₂O + DKB (B), 90% D₂O + HKB (C), 90% D₂O + DKB (D). The quality of the spectra aligns with the MQ relaxation rates: 113 s⁻¹ (A), 40 s⁻¹ (B), 67 s⁻¹ (C), 19 s⁻¹ (D)

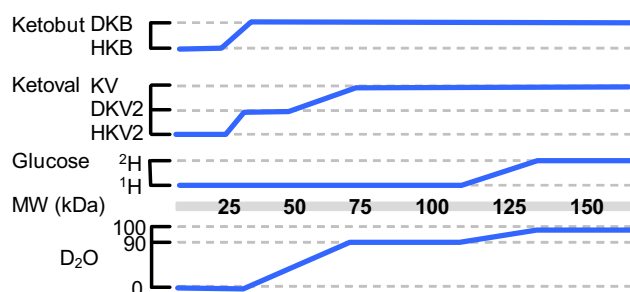
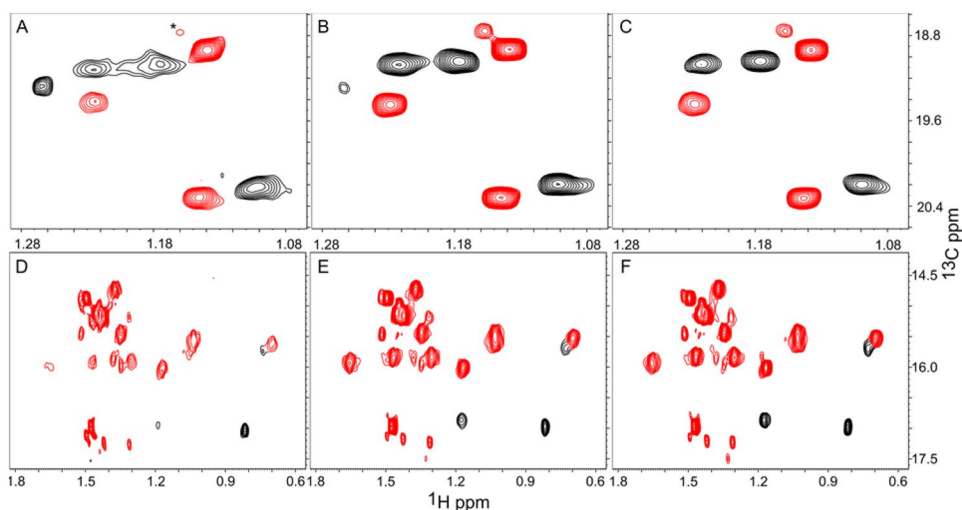


Fig. 8 Guidelines for ILV labeling conditions to obtain HSQC spectra suitable for assignments, relaxation dispersion and other uses that do not depend on very high levels of deuteration. The sloped line segments indicate that the choice of deuterated compound may be affected by the nature of the specific protein. For example, a 50 kDa protein may require growth in 90% D₂O to generate adequate spectra

Fig. 9 Thr/Ala Methyl Spectra. Thr/Ala peaks are colored red. Thr samples are the upper 3 panels. Cells were grown in H₂O media (A, D), 90% D₂O + 100% ¹H-glucose (B, E), 90% (D₂O + ²H-glucose) (C, F). All samples were grown with DKB and KV. The weaker Thr peak, marked with an * in panel A, is adjacent to only Ile, Leu, and Val residues, causing a greater signal loss



Growth on intermediate levels of D₂O has the potential to generate inhomogeneous line broadening due to a distribution of protons and deuterons on a residue that occurs when the ketoacid precursors are converted to amino acids. For example, the β -hydrogens on Leu will be protons in H₂O media and deuterons in D₂O media. In a 50:50 mixed H₂O:D₂O media the β -hydrogens will either be two protons, a proton and a deuteron, or two deuterons, with a distribution of 1:2:1. Since the three-bond isotopic shift is small (~ 0.03 ppm) (Maltsev et al. 2012), only a modest increase in linewidth is expected. This increase in linewidth is outweighed by a decrease in linewidth and increase in signal intensity due to higher levels of deuteration (see Fig. 1). Nevertheless, it may be advisable to express the protein in either 100% H₂O or elevated levels of D₂O (e.g. 90%), to reduce this effect. 100% H₂O should be used when labeling with HKV2. Growth in high levels of D₂O would lead to partial replacement of the β -proton with deuterium, resulting in a decrease in the quality of the spectra (see Fig. 6).

Alanine and threonine labeling

The protein production protocol outlined in Fig. 1 can also be used for Ala + ILV and Thr + ILV labeling. Figure 9 shows spectra of Thr and Ala samples obtained in either 100% H₂O/¹H-Glucose, 90% D₂O + 100% ¹H-glucose, or in 90% D₂O + 90% ²H-glucose. High quality spectra of Thr methyl groups are obtained under all three conditions. In the case of Ala labeling, growth in H₂O results in weak signals for some residues because of their faster MQ relaxation rate. Growth in 90% D₂O with ¹H-glucose produces spectra that are similar to the spectrum obtained in 90% D₂O + ²H-glucose. This similarity is reflected in the similar MQ relaxation rates shown in Fig. 10. As with ILV labeling, the use of ²H-glucose is an unnecessary expense for Thr/Ala labeling for proteins that are smaller than 125 kDa.

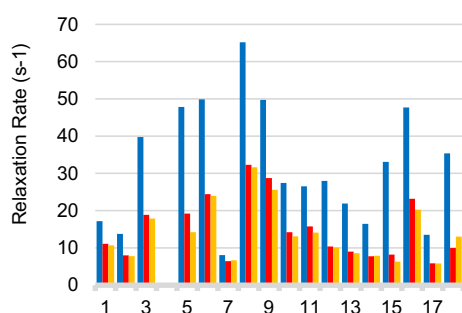


Fig. 10 Thr/Ala Relaxation Rates. Groups 1–3 correspond to the three strongest Thr signals in Fig. 9. Groups 5–18 correspond to Ala methyls. First bar in each group (blue) gives the relaxation rate for H₂O media, the second bar (red) gives the rate for 90% D₂O + 100% ¹H-glucose, the third bar gives the rate for 90% D₂O + 90% ²H-glucose

Calculations of the MQ relaxation rate for Thr and Ala methyl groups show that a subset of methyl groups can show very high MQ relaxation rates because of nearby ILV methyl groups. This effect is seen for a Thr residue in A1-1 (Fig. 9). Consequently, Ala and Thr spectra with more uniform relaxation rates will be obtained if ILV residues are fully deuterated.

In summary, we present a simple method to obtain methyl labeled deuterated protein that does not require complicated media changes and no centrifugation steps, provided the required D₂O levels is less than 97.5%. A single centrifugation step is required for higher levels of deuteration. The labeling protocol requires 8–10 h of growth after overnight adaptation, with longer growth periods required for higher levels of deuteration. The protocol is suitable for ILV, Thr, and Ala labeling. MQ relaxation measurements under different deuterium labeling conditions provide guidelines to assist in determining the optimal labeling strategy for different sized protein. For proteins ~50 kDa spectra that are adequate for resonance assignment purposes can be obtained by growth in H₂O with highly deuterated ketobutyrate and ketovalerate precursors. Increasingly higher quality spectra are obtained by increasing the D₂O content of the media. Deuterons that are provided by the solvent (D₂O) have a much larger effect on reducing the relaxation rate of the methyls than deuterons from ²H-glucose, suggesting that there is little benefit in using ²H-glucose for proteins in the 50–125 kDa range.

Experimental procedures

Cell growth

The following process was used to grow up to 100 ml of culture. The recipe for defined media (PG) (Studier

2005) is given in the supplemental material. H₂O-PG media was made with H₂O and 5 g protonated glucose/L. D₂O-PG media was made with 99% D₂O and 5 gm 99% deuterated glucose/L. Ampicillin was used for selection at 75 mg/L. Cell density was measured using scattering at 600 nm (OD₆₀₀). pET based expression plasmid was transformed into *E. coli* C3013 (New England Biolabs) on Luria broth (LB) plates. After overnight (O/N) growth five to 10 colonies were used to inoculate 10 ml of 50:50 LB:PG-D₂O (50% D₂O) and this culture was grown O/N at 37 °C. The next day 5 ml of the culture was replaced with PG-D₂O to give 75% D₂O and the culture grown for 4 h at 37 °C. For the purposes of varying the D₂O content, the cells were centrifuged and resuspended in PG media with various concentrations of D₂O (0%, 25%, 50%, 75%, 90%, 99%) as well as protonated or deuterated glucose at a cell density of ~0.4. The cells were grown at 37 °C until the optical density reached 1.0. At this point precursors for ILV methyl labeling (e.g. 50 mg/L ketobutyrate, 100 mg/L ketovalerate) were added (Goto et al. 1999). The ILV methyl-precursors in this study varied in their deuteration level (see Scheme S1 for labeling pattern) to assess the effect of deuteration of the precursors on the quality of the spectra. Unless otherwise noted, the most highly deuterated forms of the keto acids were used. Labeling of Ala + ILV methyls followed the method described by Ayala et al. (2009), except that precursors for ILV labeling were also added as described in Godoy-Ruiz et al. (2010). Note that it is important to adjust the pH of the d-succinic acid to 7.0 before addition to the media. Labeling of ILV + Thr methyls followed the method described by Velyvis et al. (2012). In the case of Ala and Thr labeling, we tested growth in 100% H₂O, growth in 90% D₂O with protonated glucose, and in 90% D₂O with 90% deuterated glucose. All cell growth was performed in baffles flasks with a culture volume that was 10–20% of the flask size.

After addition of the precursors the cells were incubated for 30–45 min, isopropyl thiogalactoside (IPTG) was added to a final concentration of 1 mM and the cells were incubated at 30 °C for 16 h, harvested, and the cell pellet was frozen at -20 °C. The final cell density was ~5.0. Occasionally, we experienced slower growth in the 90% D₂O media such that it was not possible to induce the cells within 4–5 h after the adaptation period. In this case we simply grew the cells O/N at 30 °C, shifted the temperature back to 37 °C, and induced them at OD₆₀₀ = 1.

We also explored the removal of the centrifugation step to streamline cell growth. In this case the cells were simply diluted into the media after the 4-h growth period in 75% D₂O. The final concentration of LB during labeling is 2.5% and it is possible to generate D₂O levels as high as 97.5% using this approach. Higher levels of D₂O require centrifugation and resuspension at the required D₂O level.

Enzyme expression and purification

A synthetic glutathione transferase A1-1 gene was constructed using codon frequencies as described in Boël et al. (2016) using a modified pET vector that contained the complete T7 promoter, as described in Shilling et al. (2020). Previous expression using human codons and tac-based expression vector (Stenberg et al. 1992) gave 55 mg/L. The use of optimized codons and use of the complete T7 promoter increased the yield to 400 mg/L. For purification the cell pellet was resuspended in 8 ml 20 mM Tris, 1 mM EDTA (pH 7.8). Lysozyme was added to a concentration of 1 mg/ml. The cells were frozen in liquid nitrogen, thawed in a RT water bath, and 2 ml of 0.5% triton X-100 was added (final triton 0.1%). Cells were incubated for 15 min on ice. MgSO_4 was added to 2 mM and 500 units of benzonase was added to digest DNA to reduce viscosity (15 min). The lysate was centrifuged at 30,000 rpm for 20 min (Ti70 rotor). The cleared lysate was loaded on a CM sepharose column, washed, and the protein eluted using a gradient from 0 to 1 M NaCl. Protein yields were approximately 40 mg/100 ml of culture.

The purified protein was concentrated and exchanged into 10 mM Phosphate, 50 mM NaCl, 1 mM EDTA, 0.02% Azide, 100% D_2O , pH = 7.0, except for the ^{15}N -labeled sample which was exchanged into 10% D_2O . Typical protein concentrations were 500–900 μM .

NMR spectroscopy

Spectra were acquired on a Bruker 500 MHz Neo using a 2-channel inverse cryoprobe (Prodigy) using a standard HMQC sequence with echo-antiecho selection for quadrature detection in carbon (Fig. S3A). The total evolution time in carbon was 25.5 ms (80 complex points). Data was processed and visualized using TopSpin NMR (Bruker). Deuteration levels were obtained by integrating a region of the proton spectra (-0.3 to -0.7 ppm) that does not contain signals from ILV, Thr, or Ala methyl groups and normalizing the integral by the concentration of the protein in the sample.

Measurement of MQ relaxation rate

The HMQC pulse sequence was modified to include an additional relaxation delay during the coherence selection and t_1 evolution period of the sequence (see Fig. S3B). Interleaved spectra were collected with different relaxation delay times. Individual spectra were generated using NMRPipe (Delaglio

et al. 1995) and the relaxation rate constants were obtained from exponential fitting using routines in NMRPipe.

Estimation of rotational correlation time (τ_c) and upper limit for molecular weight

The rotational correlation time of A1-1 at 300 K was estimated by measuring the rates of decay of the ^{15}N α and β spin states in ^{15}N labeled deuterated protein using [^{15}N , ^1H]-TRACT (Lee et al. 2006) (See Fig. S5 for decay curves). The difference in the relaxation rates is consistent with a τ_c of at least 20.5 ns in H_2O or at least 25 ns after adjustment for the increased viscosity of D_2O (Millero et al. 1971). Measurement of the ^{15}N - T_2 (500 MHz) using standard CPMG pulse sequences gave a median relaxation time of 28.8 ms, which is consistent with the average correlation time of ~27 ns predicted by hydrodynamic calculations (median predicted T_2 is 28.6 ms by hydroNMR (García de la Torre et al. 2000)). For purposes of estimating molecular weights, we assumed that the correlation time of A1-1 in D_2O was 30 ns. Higher molecular weights were mimicked by adding sucrose (20%) and the increase in molecular weight was estimated from the increased viscosity (Telis et al. 2007).

Calculation of relaxation rates

Relaxation rates of methyl protons were estimated from determining the proton density surrounding the relevant methyl groups. The proton density was adjusted to reflect the growth conditions (D_2O versus H_2O), the labeling pattern (e.g. γ -protons or γ -deuterons on Ile residues), and the D_2O content of the NMR buffer. In the case of the latter, all exchangeable protons were replaced by deuterons.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10858-024-00441-y>.

Author contributions K.G. and A.L. were involved in developing the methods, G.R. collected the NMR data and wrote the manuscript.

Funding Open Access funding provided by Carnegie Mellon University. This work was supported by funds from the NIH (5R21AI167017), NSF (2112872), and Carnegie Mellon University undergraduate research program. The NMR spectrometer was purchased with funding from the NSF (CHE-9808188, CHE-1039870 and CHE-1726525). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data availability All data is described in the manuscript.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest with the contents of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Amerio C, Asunción Durá M, Noirclerc-Savoye M, Perollier A, Gallet B, Plevin MJ, Vernet T, Franzetti B, Boissbouvier J (2011) A systematic mutagenesis-driven strategy for site-resolved NMR studies of supramolecular assemblies. *J Biomol NMR* 50(3):229–236
- Ayala I, Sounier R, Usé N, Gans P, Boissbouvier J (2009) An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. *J Biomol NMR* 43(2):111–119
- Boël G, Letso R, Neely H, Price WN, Wong KH, Su M, Luff J, Valecha M, Everett JK, Acton TB et al (2016) Codon influence on protein expression in *E. coli* correlates with mRNA levels. *Nature* 529(7586):358–363
- Cai M, Huang Y, Yang R, Craigie R, Clore GM (2016) A simple and robust protocol for high-yield expression of perdeuterated proteins in *Escherichia coli* grown in shaker flasks. *J Biomol NMR* 66(2):85–91
- Cai M, Huang Y, Lloyd J, Craigie R, Clore GM (2021) A simple and cost-effective protocol for high-yield expression of deuterated and selectively isoleucine/leucine/valine methyl protonated proteins in *Escherichia coli* grown in shaker flasks. *J Biomol NMR* 75(2–3):83–87
- Clay MC, Saleh T, Kamatham S, Rossi P, Kalodimos CG (2022) Progress toward automated methyl assignments for methyl-TROSY applications. *Structure* 30(1):69–79.e62
- Crublet E, Kerfah R, Mas G, Noirclerc-Savoye M, Lantze V, Vernet T, Boissbouvier J (2014) A cost-effective protocol for the parallel production of libraries of $^{13}\text{CH}_3$ -specifically labeled mutants for NMR studies of high molecular weight proteins. *Methods Mol Biol* 1091:229–244
- de la Torre JG, Huertas ML, Carrasco B (2000) HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J Magn Reson* 147(1):138–146
- Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6(3):277–293
- Di Martino RMC, Maxwell BD, Piralí T (2023) Deuterium in drug discovery: progress, opportunities and challenges. *Nat Rev Drug Discovery* 22(7):562–584
- Dubey A, Stoyanov N, Viennet T, Chhabra S, Elter S, Borggräfe J, Viegas A, Nowak RP, Burdzhiev N, Petrov O et al (2021) Local deuteration enables NMR observation of methyl groups in proteins from eukaryotic and cell-free expression systems. *Angew Chem Int Ed Engl* 60(25):13783–13787
- Godoy-Ruiz R, Guo C, Tugarinov V (2010) Alanine methyl groups as NMR probes of molecular structure and dynamics in high-molecular-weight proteins. *J Am Chem Soc* 132(51):18340–18350
- Goto NK, Gardner KH, Mueller GA, Willis RC, Kay LE (1999) A robust and cost-effective method for the production of Val, Leu, Ile (δ^1) methyl-protonated ^{15}N -, ^{13}C -, ^2H -labeled proteins. *J Biomol NMR* 13(4):369–374
- John M, Schmitz C, Park AY, Dixon NE, Huber T, Otting G (2007) Sequence-specific and stereospecific assignment of methyl groups using paramagnetic lanthanides. *J Am Chem Soc* 129(44):13749–13757
- Lee D, Hilty C, Wider G, Wüthrich K (2006) Effective rotational correlation times of proteins from NMR relaxation interference. *J Magn Reson* 178(1):72–76
- Lescanne M, Skinner SP, Blok A, Timmer M, Cerofolini L, Fragai M, Luchinat C, Ubbink M (2017) Methyl group assignment using pseudocontact shifts with PARAssign. *J Biomol NMR* 69(4):183–195
- Li J, Byrd RA (2022) A simple protocol for the production of highly deuterated proteins for biophysical studies. *J Biol Chem* 298(8):102253
- Listwan P, Pédelacq JD, Lockard M, Bell C, Terwilliger TC, Waldo GS (2010) The optimization of in vitro high-throughput chemical lysis of *Escherichia coli*. application to ACP domain of the polyketide synthase ppsC from *Mycobacterium tuberculosis*. *J Struct Funct Genomics* 11(1):41–49
- Maltsev AS, Ying J, Bax A (2012) Deuterium isotope shifts for backbone ^1H , ^{15}N and ^{13}C nuclei in intrinsically disordered protein α -synuclein. *J Biomol NMR* 54(2):181–191
- Millero FJ, Dexter R, Hoff E (1971) Density and viscosity of deuterium oxide solutions from 5–70. deg. *J Chem Eng Data* 16(1):85–87
- Nerli S, De Paula VS, McShan AC, Sgourakis NG (2021) Backbone-independent NMR resonance assignments of methyl probes in large proteins. *Nat Commun* 12(1):691
- Ollerenshaw JE, Tugarinov V, Kay LE (2003) Methyl TROSY: explanation and experimental verification. *Magn Reson Chem* 41(10):843–852
- Pervushin K, Riek R, Wider G, Wüthrich K (1997) Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci USA* 94(23):12366–12371
- Pritišanac I, Alderson TR, Güntert P (2020) Automated assignment of methyl NMR spectra from large proteins. *Prog Nucl Magn Reson Spectrosc* 118–119:54–73
- Rosen MK, Gardner KH, Willis RC, Parriss WE, Pawson T, Kay LE (1996) Selective methyl group protonation of perdeuterated proteins. *J Mol Biol* 263(5):627–636
- Rossi P, Swapna GV, Huang YJ, Aramini JM, Anklin C, Conover K, Hamilton K, Xiao R, Acton TB, Ertekin A et al (2010) A microscale protein NMR sample screening pipeline. *J Biomol NMR* 46(1):11–22
- Sattler M, Fesik SW (1996) Use of deuterium labeling in NMR: overcoming a sizeable problem. *Structure* 4(11):1245–1249
- Schanda P, Kupce E, Brutscher B (2005) SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J Biomol NMR* 33(4):199–211
- Shilling PJ, Mirzadeh K, Cumming AJ, Widesheim M, Köck Z, Daley DO (2020) Improved designs for pET expression plasmids increase protein production yield in *Escherichia coli*. *Commun Biol* 3(1):214
- Stenberg G, Björnstedt R, Mannervik B (1992) Heterologous expression of recombinant human glutathione transferase A1–1 from a hepatoma cell line. *Protein Expr Purif* 3(1):80–84
- Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 41(1):207–234

- Telis VRN, Telis-Romero J, Mazzotti HB, Gabas AL (2007) Viscosity of aqueous carbohydrate solutions at different temperatures and concentrations. *Int J Food Prop* 10(1):185–195
- Trivedi R, Nagarajaram HA (2020) Substitution scoring matrices for proteins—an overview. *Protein Sci* 29(11):2150–2163
- Tugarinov V, Kay LE (2003) Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *J Am Chem Soc* 125(45):13868–13878
- Tugarinov V, Kay LE (2004) An isotope labeling strategy for methyl TROSY spectroscopy. *J Biomol NMR* 28(2):165–172
- Velyvis A, Ruschak AM, Kay LE (2012) An economical method for production of (2)H, (13)CH₃-threonine for solution NMR studies of large protein complexes: application to the 670 kDa proteasome. *PLoS ONE* 7(9):e43725
- Wiberg KB (1955) The deuterium isotope effect. *Chem Rev* 55(4):713–743

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.