



NMR Study of larger proteins using isotope labeling

Sung Jean Park*

College of Pharmacy, Gachon University, 534-2 Yeonsu-dong, Yeonsu-gu, Incheon, Republic of Korea

Received Oct 10, 2014; Revised, Dec 5, 2014; Accepted, Dec 11, 2014

Abstract Larger proteins (above molecular weight 50 kDa) usually show slow motional tumbling in solution, which facilitates the decay of NMR signal, resulting in poor signal-to-noise. In the past twenty years, researchers have tried to overcome this problem with higher molecular weight by improvement of hardware (higher magnetic field and cryoprobe), optimization of pulse sequences for larger molecules, and development of isotope-labeling techniques. Actually, GroEL/ES complex (\approx 900 kDa) was successfully studied using combination of above techniques¹. Among the techniques used in large molecular studies, the impact of isotope-labeling for large molecules study is summarized and discussed here.

Keywords NMR, Large molecular weight, Isotope labeling, Methyl TROSY

Large proteins studied by NMR

First breakthrough of larger protein studies by NMR was reported in early 2000. Fiaux *et al* published the GroEL/ES complex in 2002¹. Two-dimensional [¹⁵N,¹H]-correlation spectra was obtained by universal > 97% ²H labeling to overcome the size limit. This work successfully established that TROSY/CRINEPT-based ¹⁵N-¹H correlation experiments in combination with perdeuteration are able to provide collective information of proteins

with molecular weight of up to 900 kDa. In the same year, Rudiger *et al* reported structural information of the p53 core domain-bound Hsp90 (\approx 200 kDa) using TROSY/CRINEPT techniques². The same molecular system was monitored by Park *et al* using methyl-specific labeling method, revealing the mode of binding of the p53 core domain to Hsp90³. The bigger molecular system, L7/L12 ribosome complex corresponding to about 2.3 MDa was also monitored in heteronuclear NMR study⁴. Before this study, this large system can be only detected in cryoelectron microscopy. Even though Christodoulou *et al* detected only several narrow lines in NMR spectra, the information was sufficient for explaining dynamic mobility of the L7/L12 proteins on the ribosome. The highly conserved, 300 kDa cylindrical protease ClpP composed of 14 subunits was also studied mainly using methyl resonances⁵. It showed that the handle region is highly dynamic and the ring structure makes equilibrium between two distinct conformational states. Later, the same group published the 670 kDa 20S proteasome core particle using methyl resonances of Ile, Leu, and Val with highly deuterated background⁶. These studies provided strategies to overcome the size limitation that prevents the quantitative analysis using NMR.

Deuteration and Perdeuteration

To extend the use of NMR to larger protein systems,

* Address correspondence to: **Sung Jean Park**, College of Pharmacy, Gachon University, 534-2 Yeonsu-dong, Yeonsu-gu, Incheon, Republic of Korea, Tel: 82-32-820-6113; E-mail: psjnmr@gachon.ac.kr

one approach is the utilization of deuterium labeling. The usefulness of deuteration in protein study is originated from the reduction of relaxation rate by bulk protons. The larger the sizes of proteins, the bigger the density of protons in proteins, and which result in line-broadening of each resonance. A magnetization (usually of protons) in protein experiences multiple relaxation pathways by surrounding protons that are coupled through space (Fig. 1). This is so called spin diffusion. In a larger protein system, this spin diffusion is more significant so that the spectra usually show a low signal-to-noise ratio. Replacement of proton by deuterium eliminates many of relaxation pathways, which enhances the life-time of a magnetization and improves spectrum quality.

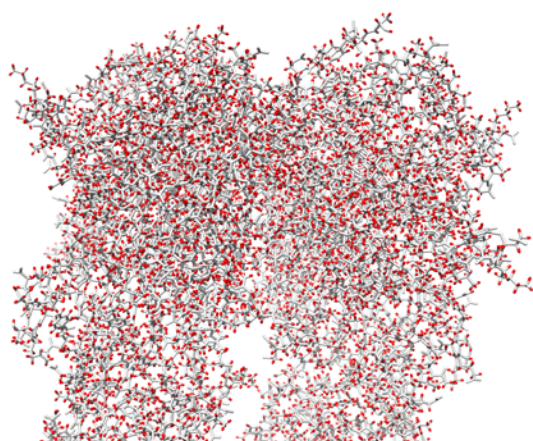


Figure 1. Network of protons (colored red) in the N-terminal domain of Hsp90

Another advantage of deuteration is related to the improved life-time of magnetization from hetero nuclei such as ^{13}C and ^{15}N that are directly attached to protons. The bonded protons highly affect the relaxation of ^{13}C and ^{15}N by dipolar interaction. For example, the C_α relaxation times increase from 16.5ms (100% ^1H) to 130ms (70% ^2H) in a 37kDa Trp repressor–DNA complex. The increased relaxation time is due to the smaller gyromagnetic ratio (γ) of deuterium compared to the proton. Usually, the relaxation rates of hetero nuclei are proportional to the square of the gyromagnetic ratios

$((\gamma_{\text{D}}/\gamma_{\text{H}})^2 \approx 0.02)$, which leads to reduced dipolar relaxation rate of hetero nuclei (about 16 times less). Indeed, the linewidth of $^{13}\text{C}_\alpha$ attached ^2H is not broadened up to 50 kDa, while that of ^1H shows a significant line broadening according to the increase of molecular weight⁷. Thus, deuteration of target protein is necessary for studying large molecules.

The required deuteration level depends on the system. Especially, random fractional deuteration (below uniform 80% deuteration) has been widely used to study proteins. Through fractional deuteration, the transverse relaxation time of amide protons is expected to highly increase. However, in a larger protein system the high deuteration level ($> 95\%$) is a prerequisite for obtaining analyzable spectra. Perdeuterated protein sample has several benefits. First, as described above, strong reduction of external relaxation contribution can be achieved. Second, it is possible to remove the internal dipolar contributions of $^{13}\text{CH}_n$. Third, we can expect the increase in T_2 and T_1 values of $^{13}\text{CH}_n$ and $^1\text{H}_N$ as well as sharper lines by removal of $J(\text{H}, \text{H})$.

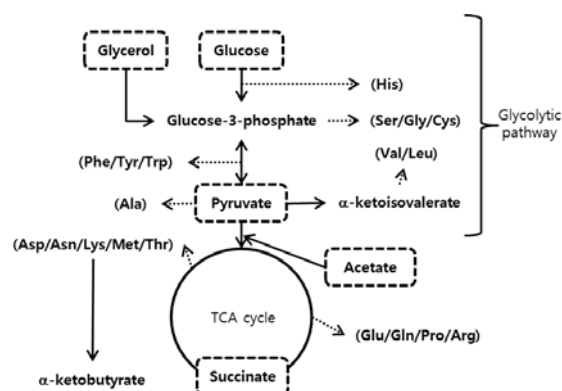


Figure 2. Pathway of amino acid biosynthesis in *E. coli*. The usual carbon sources for deuteration are represented by dotted line.

To produce deuterium-enriched protein in *E. coli*, several chemicals can be used. *E. coli* can synthesize all its amino acids using single nitrogen and carbon sources. Thus, appropriate selection of nitrogen or carbon sources in D_2O growth media enables incorporation of deuterium into amino acids.

Practically, to use only D₂O growth media without deuterated carbon sources provides low-level incorporation of deuterium into amino acids. In general, deuterated carbon sources such as ²H-glucose, ²H-pyruvate, ²H-acetate, ²H-succinate, and ²H-glycerol can be chosen for deuteriation as depicted in Fig. 2. However, one should consider that the carbon sources affect the production yield of target protein when *E. coli* is grown in single carbon source. The glucose is the most widely-used chemical and the production yield should be best among carbon sources, while the cost of ²H-glucose is relatively high. Compared to the growth in minimal media containing normal glucose, the growth rate and the maximum population of *E. coli* usually decrease more than two fold⁸. This may be caused by short of ¹H sources that maintain protein structures via hydrogen bonding.

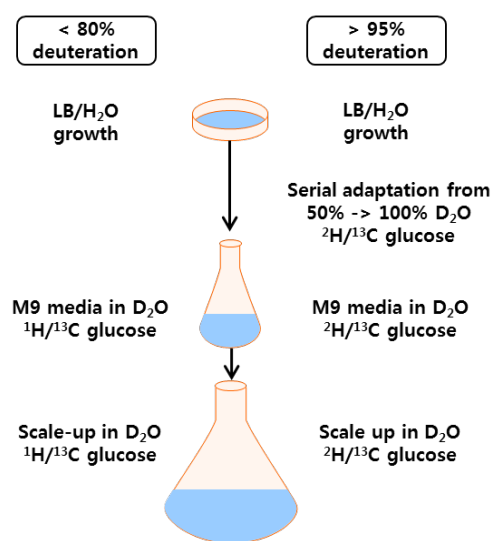


Figure 3. Serial culture schemes of deuteration.

²H-pyruvate and ²H-acetate can be alternative of ²H-glucose while the production yield is lower than that of ²H-glucose due to their toxicity. Actually, several proteins cannot be obtained in ²H-acetate media. Benefit for using these two chemicals is they are cheaper than glucose. Growth on ²H-glycerol will result in isotope enrichment at alternating carbons in most amino acids, including isolated aromatic

carbons in Phe, Tyr, and Trp. For the purpose of perdeuteration of protein, ²H-glucose, ²H-glycerol, and ²H-succinate can be generally applied. Other materials are mainly used for ¹³C labeling or methyl protonation rather than perdeuteration. During perdeuteration, *E. coli* should be serially adapted to the D₂O media from 50% to 100% D₂O since abrupt change in growth media can cause bacterial death (Fig. 3).

Methyl-selective isotope labeling

Fractional deuteration is generally effective when correlation time, τ_c is below 18 ns (accordingly the size of molecules is below 35 kDa). Above 18 ns, the limitation of enhanced life time of HN resonances by deuteration still exists⁹, which means that monitoring HN resonances in larger system is not promising and other labeling strategies including perdeuteration are required for NMR measurement. Monitoring methyl protons can be successfully applied to the larger system. Methyl protons (-CH₃) can offer approximately 3-fold intrinsic sensitivity since three protons are coupled to one carbon nucleus compared to the single amide proton-nitrogen pair. In addition, methyl groups in proteins have high mobility with slow relaxation time as well as their protons do not exchange with solvent protons. The first application of methyl resonance to study protein was done by London et al¹⁰. They monitored the bacterial dihydrofolate reductase (DHFR) by incorporating the methyl-¹³C methionine. Later, Rosen et al showed selectively methyl-protonated samples will be useful in many areas of structural analysis of larger molecules and molecular complexes by NMR¹¹. The report on selective protonation strategy for (¹H-delta methyl)-leucine and (¹H-gamma methyl)-valine in perdeuterated protein greatly accelerated the use of methyl resonances in structural studies in larger complex systems¹². Currently, most of the larger protein systems with molecular weight of above 100 kDa are studied using methyl proton resonances in combination with perdeuteration (Fig. 4).

To prepare a methyl protonated and perdeuterated sample, two chemicals are mainly used. $[3,3\text{-}^2\text{H}]\text{-}^{13}\text{C}$ α -ketobutyrate can be used to incorporate the protons into δ_1 carbon position in perdeuterated condition. $[3\text{-}^2\text{H}]\text{-}^{13}\text{C}$ α -ketoisovalerate can be used to incorporate ^1H - δ methyl-Leucine and -Valine in perdeuterated proteins. These two chemicals are actually metabolites produced in amino acid biosynthesis in *E. coli* as shown in Fig. 2. Thus, ^2H -pyruvate can be used for the same purpose while metabolic scrambling may happen to some extent.

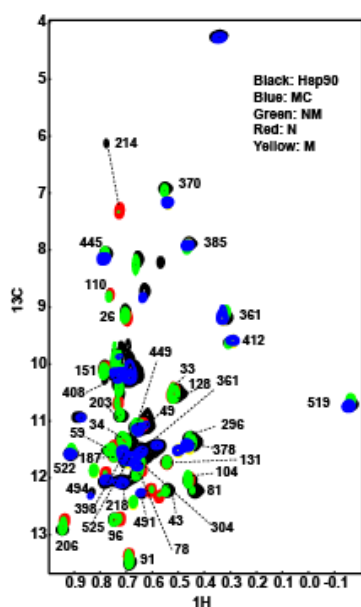


Figure 4. Example of Ile (δ_1) methyl-protonated and perdeuterated NMR spectra of Hsp90¹³. The molecular size of Hsp90 dimer corresponds to approximately 180 kDa. All Ile residues of Hsp90 can be detected in the methyl TROSY spectra. The spectrum was measured in DRX 800 MHz bruker machine.

A simple addition of α -ketoisovalerate and α -ketobutyrate to D_2O media containing ^2H -glucose 1hr prior to induction can easily produce Val, Leu, Ile (δ_1) methyl-protonated ^{15}N -, ^{13}C -, ^2H -labeled proteins. The concentration of α -ketoisovalerate and α -ketobutyrate should not exceed 100 mg/L and 50

mg/L, respectively. This is because $[3,3\text{-}^2\text{H}]\text{-}^{13}\text{C}$ α -ketobutyrate incorporation into Ile can decrease at higher α -ketoisovalerate concentration. As shown in Fig. 2, high level of α -ketoisovalerate increases cellular concentration of Val which activates the enzyme, threonine deaminase. This enzyme degrades Thr into α -ketobutyrate with which *E. coli* can incorporate the non-protonated methyl group into Ile since Thr is mainly originated from full-deuterated glucose.

In summary, the development of labeling techniques together with advance of hardware and software has facilitated the studies on the larger molecular system. Currently, the deuteration and methyl-protonation have become basic techniques for researches and expanded availability of NMR study. In addition to the two basic labeling strategies introduced here, many other specific labeling strategies such as stereo-array isotope labeling and *in vitro* cell-free labeling are currently used and being developed. All these efforts should contribute evolution of NMR techniques, leading NMR as a valuable tool for studying large biological system.

Acknowledgements

This work was supported by the Gachon University Gil Medical Center (Grant number : 2013-36).

References

1. Fiaux J, Bertelsen EB, Horwich AL, Wüthrich K. *Nature* **418**, 207. (2002).
2. Rudiger S, Freund SM, Veprintsev DB, Fersht AR. *Proc. Natl. Acad. Sci. U S A*. **99**, 11085-90. (2002).
3. Park SJ, Borin BN, Martinez-Yamout MA, Dyson HJ. *Nat. Struct. Mol. Biol.* **18**, 537. (2011).
4. Christodoulou J, Larsson G, Fucini P, Connell SR, Pertinhez TA, Hanson CL, Redfield C, Nierhaus KH, Robinson CV, Schleucher J, Dobson CM. *Proc. Natl. Acad. Sci. U S A* **101**, 10949. (2004).
5. Sprangers R, Gribun A, Hwang PM, Houry WA, Kay LE. *Proc. Natl. Acad. Sci. U S A* **102**, 16678. (2005).
6. Sprangers R, Kay LE. *Nature* **445**, 618. (2007).
7. L.E. Kay, *J. Am. Chem. Soc.* **116**, 11655. (1994).
8. Lian L-Y, Middleton DA. *Progress in Nuclear Magnetic Resonance Spectroscopy* **39**, 171. (2011).
9. Nietlispach D *et al.*, *J. Am. Chem. Soc.* **118**, 407. (1996).
10. London RE, Wageman WE, Blakley RL. *FEBS Lett.* **160**, 56. (1983).
11. Rosen MK, Gardner KH, Willis RC, Parris WE, Pawson T, Kay LE. *J. Mol. Biol.* **263**, 627. (1996).
12. Goto NK, Gardner KH, Mueller GA, Willis RC, Kay LE. *J. Biomol. NMR* **13**, 369. (1999).
13. Park SJ, Kostic M, Dyson HJ. *J. Mol. Biol.* **411**, 158. (2011).