Journal of the Korean Magnetic Resonance Society **2011**, *15*, 115-127

DOI : 10.6564/JKMRS.2011.15.2.115

Cost-effective isotope labeling technique developed for $^{15}N/^{13}C$-labeled proteins

Heeyoun Kim, Eunmi Hong and Weontae Lee*

Department of Biochemistry, College of Life Science and Biotechnology, Yonsei University, Seoul 120-749, Korea
(Received November 09, 2011 ; revised November 30, 2011 ; accepted December 13, 2011)

Abstract: A newly developed cost-effective approach to prepare $^{15}N/^{13}C$-labeled protein for NMR studies is presented. This method has been successfully applied to isotopically labeling of PTK6 SH2 domain and MTH 1880 protein. The production method generates cell density using a growing media containing $^{15}NH_4Cl$, $^{12}C_6$-D-glucose. Following a doubling time period for unlabeled metabolite exhaustion and then addition $^{13}C_6$-D-glucose into a M9 growing media, the cells are induced. Our results demonstrate that in order to get full incorporation of ^{13}C, the isotopes are not totally required during the initial growth phase before induction. The addition of small amounts of $^{13}C_6$-D-glucose to the induction phase is sufficient to obtain more than 95% incorporation of isotopes into the protein. Our optimized protocol is two-thirds less costly than the classical method using ^{13}C isotope during the entire growth phase.

Key words: isotope labeling, protein expression, NMR, cost-effective

INTRODUCTION

Isotopic labeled samples are generally required for optimizing biomolecular NMR techniques, performing heteronuclear spin relaxation studies [1], biological screening applications such as 'SAR by NMR' [2], and high throughput structural genomics studies [3].

* To whom correspondence should be addressed. E-mail : wlee@spin.yonsei.ac.kr

The size of macromolecular structures that can be solved by NMR has been dramatically increased over the last decade largely due to the advent of three- and four-dimensional heteronuclear spectroscopy [4-6] and takes advantage of transverse relaxation optimized spectroscopy (TROSY). The high-resolution protein structures being obtained today from triple-resonance NMR techniques rely on the generation of milligram quantities of protein labeled with ^{13}C, ^{15}N, and ^{2}H isotopes [7]. Probably the most serious problem for determining protein structures by NMR is to prepare well-purifying protein samples. A necessary stage of these studies about macromolecular protein structure is the purification by expressing the protein of interest in a bacterial host, using a growing media containing $^{15}NH_4Cl$, $^{13}C_6$-D-glucose as a isotopically labeled source. This is commonly achieved by expressing the protein of interest in a bacterial host, typically *E.coli*, using a growth medium containing ^{15}N and ^{13}C as the sole nitrogen and carbon sources, respectively.

A variety of different strategies for enhancing *E.coli* growth and protein expression have been employed. Minimal media supplements such as trace metal mixtures, vitamin cocktails as well as commercially available algal and microbial hydrolysates have shown enhancements in growth and expression [8-11]. Despite these advances, it is desirable to achieve higher expression yields at lower isotope costs using commonly employed shake flask cultures.

In this paper, we present a simple procedure for the production of isotopically labeled proteins that permit high yields of labeled protein while significantly reducing the isotope cost. We demonstrate that by growing the bacterial cells initially in medium at natural isotopic abundance and

providing isotopically labeled nutrients only after induction, the yield of labeled protein per unit nutrient can be increased more than about two-fold for ^{13}C-labeling. We developed and tested this protocol using as example the SH2 domain of protein tyrosine kinase 6 and MTH1880.

EXPERIMENTAL

Materials

Chemicals of the highest grade available were obtained form Sigma. Isotope enriched (=98% ^{15}N) ^{15}N-NH_4Cl and (=98% ^{13}C) $^{13}C_6$-D-glucose was from Cambridge Isotope Laboratories Inc.

Protein expression and purification

PTK6-SH2 was over expressed in *E. coli* strain BL21 (DE3) transformed with the pGEX 4T-3 / SH2 plasmid construct and MTH1880 was over expressed in *E. coli* strain BL21 (DE3) pLysS transformed with the pET13d / MTH1880 plasmid construct.

A standard M9 minimum media [12] was used as the basic growth medium that was either supplemented with $^{14}NH_4Cl$ or $^{15}NH_4Cl$ and $^{12}C_6$-D-glucose or $^{13}C_6$-D-glucose. For ideal growth, thiamine 0.1% and ampicillin with final concentration of 50 ug/ml were added in 500 mL culture. To prepare ^{15}N-labeled samples of the proteins, cells were grown in growing media supplemented with 10 g/L D-glucose, 1 g/L $^{15}NH_4CL$. For ^{13}C/^{15}N-labeled proteins, ^{13}C-D-labeled glucose was used as the carbon source.

Shaking flask cultures were inoculated from a 2ml culture of cells that were grown from a seed culture overnight in LB media with shaking at 37 °C. The cells were induced with 1 mM isopropyl □-D-thiogalactopy-ranoside (IPTG) when OD_{600} was reached at 0.5-0.6.

SH2 domain of protein tyrosine kinase 6 was purified by GST affinity chromatography followed by fast-protein liquid chromatography (FPLC) using Superdex 75 HR 10/30 column as described by Hong et al. [13]. MTH 1880 was purified by Ni-NTA column (Novagen), eluted with 200mM imidazole. To remove HisTag, fractions containing protein were incubated with bovine thrombin (BioRad) in the ratio of 10 Unit/mg at 25 °C for 6 hours. After thrombin digestion, the protein solution was passed through gel filtration column (BioRad) using Acta System. In this procedure, loading buffer was used with NMR buffer (25 mM sodium phosphate, 300 mM NaCl, 0.002% NaN_3). The purified proteins were concentrated with Centricon-3 concentrators (Millipore), placing in a 5 mm symmetrical micro cell (Shigemi).

Protein concentrations were determined from a Bradford assay using bovine serum albumin (BSA) as the standard. The concentrations of the purified proteins were accurately measured by the absorbance at 280 nm on a BECKMAN DU 530 UV/VIS spectrophotometer. Purified SH2 domain was approximately 1mM and purified MTH 1880 was approximately 2 mM in 500 mL culture.

Mass Spectrometry

The molecular weights of the purified proteins were analyzed using matrix-assisted laser-

desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) on a Voyager DE (PE Biosystems, Foster City, CA) instruments.

NMR Spectroscopy

All NMR experiments were preformed on either a Bruker DRX-500 MHz or a Varian Unity INOVA 500 MHz spectrometer equipped with a triple resonance probes with shielded x, y, z-gradients. Nmr datas of SH2 domain were obtained at 25 °C, MTH 1880 were conducted at 37 °C. The 3D triple resonance experiments, HNCA [14-16], HN(CO)CA [14], HNCACB [16-17], CBCA(CO)NH [18], HNCO [14, 17, 19], HCACO [20], HCCH-TOCSY [21] and ^{13}C-edited-NOESY [22] experiments were conducted on [U-^{13}C, U-^{15}N] SH2 domain and MTH1880.

RESULTS AND DISCUSSION

Gene cloning

The cDNA segment encoding the PTK6 SH2 domain (residues 75-174) and MTH1880 were amplified using a primer′ pair (5′-CGGGATCCGAACCGTGGTTCTTTGGC-3′ and 5′-HGGAATTCACTCGTGCTTCCGGCAGG-3′, lowercase letters were introduced for restriction enzyme sites and termination codon) by PCR and cloned into the *Bam*HI-*Eco*RI sites of pGEX 4T-3 (Amersham Pharmacia Biotech), an *E. coli* expression vector. The expressed recombinant protein contained Glutathione-S-transferase (GST) tag encoded from the vector sequence in its N-terminus.

This vector was used to transform the *E. coli* strain BL21(DE3)pLysS for fusion protein expression.

Protein expression

Expression levels of SH2 domain and MTH1880 were visualized by SDS-PAGE gel (15%). Judging from the Coomassie blue staining of the total, soluble and insoluble proteins on SDS-PAGE gels, all fusion proteins were found almost exclusively in soluble form.

^{13}C-isotope labeling of the proteins

The first step in achieving cost-effective production of labeled protein without compromising yield of bacterial cell mass is to establish the minimal requirement for the metabolite destined for labeling. For that reason, initial experiments were carried out to establish the relationship between generation of cell mass and glucose consumed from our growth medium. Several sets of conditions for bacterial growth were trailed using various glucose quantities. Using the M9 minimum media at temperature of 37 °C, we achieved an OD_{600} of ~0.8 for 1 g/1L of glucose and ~1.4 for 1 g/1L of glucose (Fig. 1). These data provide the basis for the design of optimal growth media for ^{13}C-isotope labeling. Also, we use these results to identify that express level of fusion proteins when carbon source was limited.

Prior to induction, we use these results to verify that optical density of the culture at 600 nm (OD_{600}) reached 0.4-0.6. Then a necessary carbon source of 0.5-1.0 g is added per 500 mL to obtain

the desired cell number. And about 2 g-carbon source necessary to make sufficient cell mass that protein expression accomplish by 1mM IPTG induction. It is the fact that glucose must be used as the carbon source from the culture medium before induction. Because $^{13}C_6$-D-glucose cost are expensive, we wanted to acquire if it was possible to obtain high yields of protein from the culture media that is limited carbon source. In addition, it was also interested how given carbon source in the growing media came from the cell mass that were accumulated during the wanted protein expression.

We performed a trial that $^{12}C_6$-D-glucose was used as the carbon source and the M9-media was exchanged with $^{13}C_6$-D-glucose at the induction phase. Figure 2 illustrates a typical growth for ^{15}N, ^{13}C double labeling of SH2 domain. Initially (time point a), the medium contained both 1 g/1L of $^{12}C_6$-D-glucose, 1 g/1L $^{15}NH_4Cl$ (Fig. 2A) and 2 g/1L of $^{12}C_6$-D-glucose, 1 g/1L $^{15}NH_4Cl$ (Fig. 2B).

Cells were pelleted (time point b) by centrifugation at 6500 g for 20min when the value of OD_{600} reached 0.55-0.6 and the cells were exchanged into a PBS buffer (pH 7.4) slowly. Harvest and media transfer process of cells repeated more than once. Finally, cells were harvested and resuspended in isotopically labeled growth media excluding all carbon sources. The OD_{600} value was slightly decreased probably due to the loss of cells in washing them, and then cells were cultured at growth medium in the absence of glucose during double time of *E. coli*. A consequent rise in the OD_{600} value at time point c marked recovery of growth and exhaustion of an essential nutrient, in this case $^{12}C_6$-D-glucose. The elimination of nutrient at this point allows for any remaining small amounts of unlabeled nutrient to be consumed before protein production is initiated.

At this point (time point c), protein expression was induced by the addition of IPTG of 1 mM and $^{13}C_6$-D-glucose. After induction, the cellular metabolism slow down and OD_{600} value eventually increases as expected. Cells were harvested after induction period (time point d), at which time all glucose was consumed. The harvesting point, as well as the final quantity of labeled glucose, needs to be determined empirically since it depends on the density of the culture at the time of induction, the length of induction for optimal yield of protein, and the response of the host cell to the expressed protein. For the example described above and illustrated in Fig. 2B, approximately 40 mg of label fusion protein was produced from 0.5 L of culture.

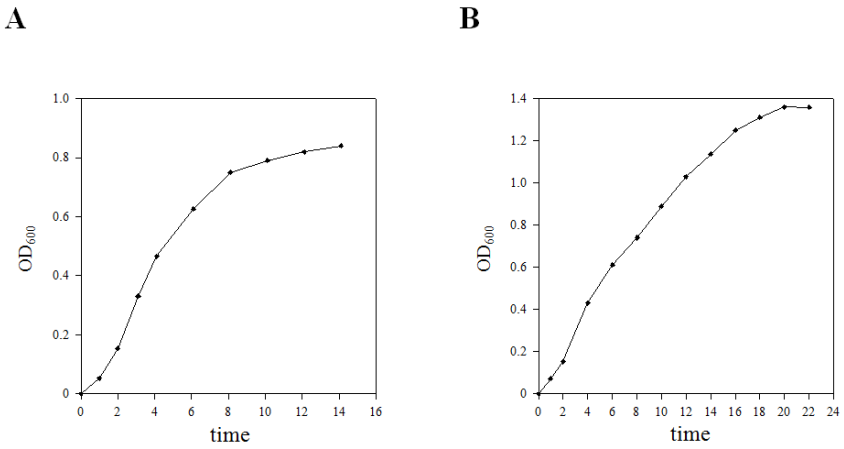

Figure 1. Growth curves for $^{13}C_6$-D-glucose condition as a function of time for $^{15}N/^{13}C$ labeling of the SH2 domain. Cells were induced with IPTG when OD_{600} value reached 0.4-0.5. Cells grew M9 media containing 1 g/1l $^{13}C_6$-D-glucose (A) and 2 g/1l $^{13}C_6$-D-glucose (B).

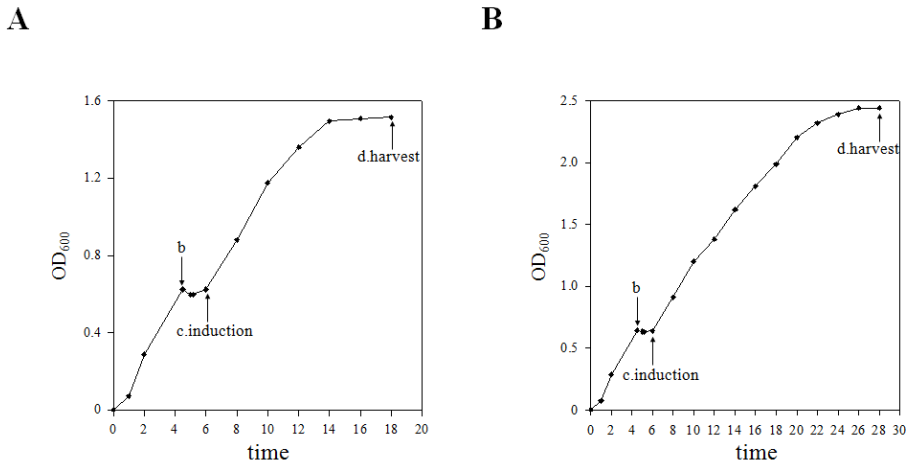

Figure 2. Induction profile for each labeling condition as a function of time for $^{15}N/^{13}C$ labeling of the SH2 domain. Growth curve using M9 media transfer to 1g/1l $^{13}C_6$-D-glucose (A) and 2g/1l $^{13}C_6$-D-glucose (B). The time points are as follows: a. start of run; b. removal of nutrients and media exchange c. addition of labeled nutrient ($^{13}C_6$-D-glucose) and induction; d. cell harvest.

Fig. 3 shows the intensity of the protein expression level using different $^{13}C_6$-D-glucose concentrations for SDS-PAGE band. The protein yields are 2.5 times higher than that of conventional method by growing in shake flasks without media transfer with same amount of isotope.

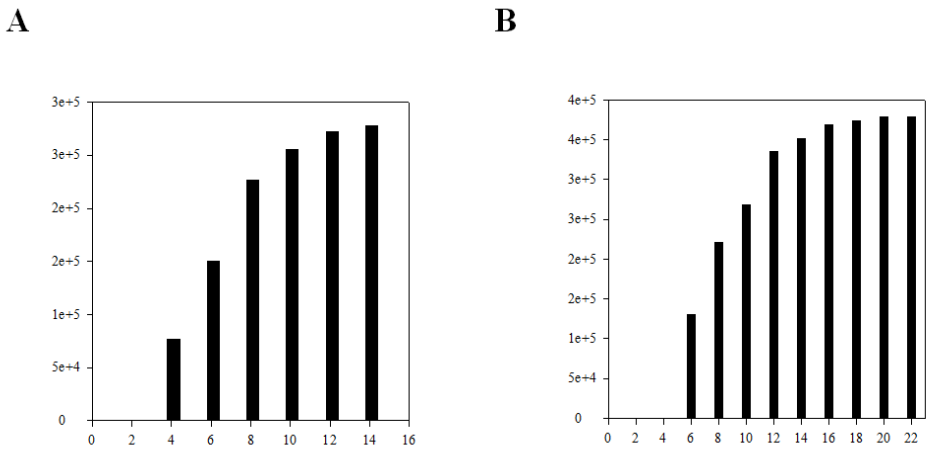

Figure 3. The intensity of the protein expression level using different $^{13}C_6$-D-glucose concentrations as a function of time for SDS-PAGE gel(15%). The SDS-PAGE gel is stained with Coomassie blue. Intensity measurements were made on TotalLab v1.10 Beta 1. In order to establish and optimize the carbon metabolite for this protocol, two sets of condition for cell growth were trailed; 1g $^{13}C_6$-D-glucose per 1L (A) and 2g $^{13}C_6$-D-glucose per 1L (B), respectively.

Characterization of isotope labeled proteins

The level of isotope enrichment was determined using the MALDI-TOF and NMR spectroscopy. Mass spectra of purified ^{13}C, ^{15}N-labeled proteins were used to confirm the molecular masses for SH2 domain and MTH 1880 (data not shown).

Distinctly different recombinant proteins have been expressed using this production method, in all cases improved protein yields were observed at reduced isotope costs. In general, enrichment

levels of ~95% are sufficient for all of the heteronuclear NMR experiments currently employed for assignment and structure determination of proteins. Fig. 4A shows strip plot taken from the HNCA spectrum displaying the $C_{\square}$ sequential connectivity of residues Arg^{11} through Qln^{19} in the $^{13}C/^{15}N$ labeled SH2 domain. HNCACB spectra show the sequential connectivity of the backbone of residue Asn^{38} through Ser^{44} in the $^{13}C/^{15}N$ labeled MTH 1880; the excellent sensitivity observed validates the labeling strategy (Fig. 4B).

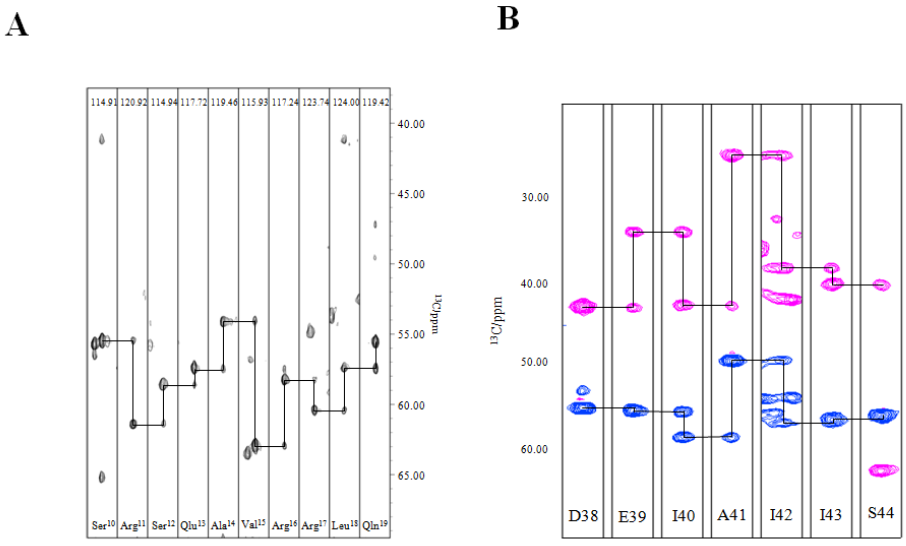

Figure 4. Strip plots from the HNCA, HNCACB spectra record SH2 domain and MTH1880. (A) Strip plots from ^{15}N-plans of the HNCA spectrum of the $^{15}N/^{13}C$-labeled SH2 domain illustrating through-bond sequential assignments for residues Ser^{10}-Qln^{19} (Helix-A region). (B) Strips plots from ^{15}N-plans of the HNCACB spectrum of $^{15}N/^{13}C$-labeled MTH1880, giving the $C_{\square}$ and $C_{\square}$ sequential connectivity between Asn^{38} and Ser^{44}.

CONCLUSIONS

Modifications of the above-described procedure can be easily implemented for different types of labeled nutrients. For example, the labeling protocol presented in this paper also allows for efficient and cost-effective deuterium labeling of large proteins.

This method needed a growth media exchanging to shaking incubator growths before induction phase that require additional set-up time and specialized equipment that is economically prohibitive for expressions of recombinant proteins. Our optimized protocol reduces two-thirds of labeling cost compared to that of the conventional method using ^{13}C-isotope during the entire growth phase.

REFERENCES

1. A.G. Palmer, C. Bracken, *In NMR Supramolecular Chemistry* (pons, M., Ed.). Kluwer Academic Publishers, Dordrecht, pp.171, (1999).
2. W. Shu, Ji, H. Liu, M. Lu, *J. Mol. Biol.* **299**. 1101, (2000).
3. G.T. Montelione, D. Zheng, Y.J. Huang, Gunsalus, K.C. and Szyperski, T. *Nat. Struct. Biol.* **7**, 982, (2000).
4. A. Bax, S. Grzesiek, Acc. Chem. Res. **26**, 131, (1993).
5. G. M. Clore, A. M. Gronenborn, Science **252**, 1390, (1991a).
6. G. M. Clore, A. M. Gronenborn, Prog. Nucl. Magn. Reson. Spectrosc. **23**, 43, (1991b)
7. G. M. Clore, A. M. Gronenborn, *Methods Enzymol.* **239**, 349, (1994).
8. D. Reilly, W.J. Fairbrother, *J. Biomol. NMR* **4**, 459, (1994).
9. M. Jansson, Y.C. Li, S. Anderson, G. Montelione, B. Nilsson, *J. Biomol. NMR* **7**, 131, (1996).
10. M. Kainosho, *Nat. Struct. Biol.* **4**, 858, (1997).

11. M. Cai, H. Ying, K. Sakaguchi, G.M. Clore, A.M. Gronenborn, R. Craigie, *J. Biomol. NMR* **11**, 97, (1998).

12. J. Sambrook, E.F. Fritsch, T. Maniatis, *Molecular Cloning : A Laboratory Manual*. **3**. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. (1989).

13. E. Hong, J Shin, E. Bang, MH. Kim, ST, Lee, W. Lee *J. Biomol. NMR* **19**, 291, (2001)

14. S. Grzesiek, A. Bax, *J. Magn. Reson*. **96**, 432, (1992)

15. L.E. Kay, G.Y. Xu, T. Yamazaki, *J. Magn. Reson*. **A109**, 129, (1994)

16. J. Stonehouse, R.T. Clowes, G.L. Shaw, J. Keeler, E.D. Laue, *J. Biomol. NMR* **5**, 226, (1995)

17. D.R. Muhandiram, L.E. Kay, *J. Magn. Reson*. **B103**, 203, (1994)

18. S. Grzesiek, A. Bax, *J. Magn. Reson*. **99**, 201, (1992)

19. M. Ikura, L.E. Kay, A. Bax, *Biochemisty* **29**, 4659, (1990)

20. A. Grzesiek, and A. Bax, *J. Magn. Reson*. **B102**, 103, (1993)

21. L.E. Kay, G.Y. Xu, A.U. Singer, D.R. Muhandiram, J.D. Forman-Kay *J. Magn. Reson*. **B101**, 333, (1993)

22. A.L. Davis, J. Keeler, E.D. Laue, D. *J.* Moskau, *Magn. Reson*. **98**, 207, (1992)

FOOTNOTES

Abbreviations: SH2, Src homology 2; MTH 1880, Methanobacterium Thermoautotropy H. 1880; NMR, nuclear magnetic resonance; MALDI-TOF MS, matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry