Journal of Biomolecular NMR, **20:** 71–75, 2001.
KLUWER/ESCOM

A method for efficient isotopic labeling of recombinant proteins

Jonathan Marley, Min Lu and Clay Bracken*
Department of Biochemistry, Weill Medical College of Cornell University, 1300 York Avenue, New York, NY 10021, U.S.A.

Received 12 January 2001; Accepted 7 March 2001

Key words: ^{13}C-, ^{15}N-enrichment, deuteration, *Escherichia coli*, heteronuclear NMR, isotope labeling, protein expression

Abstract

A rapid and efficient approach for preparing isotopically labeled recombinant proteins is presented. The method is demonstrated for ^{13}C labeling of the C-terminal domain of angiopoietin-2, ^{15}N labeling of ubiquitin and for ^{2}H/^{13}C/^{15}N labeling of the *Escherichia coli* outer-membrane lipoprotein Lpp-56. The production method generates cell mass using unlabeled rich media followed by exchange into a small volume of labeled media at high cell density. Following a short period for growth recovery and unlabeled metabolite clearance, the cells are induced. The expression yields obtained provide a fourfold to eightfold reduction in isotope costs using simple shake flask growths.

A majority of biomolecular NMR techniques require isotopic labeling (^{2}H, ^{13}C, and/or ^{15}N) of recombinant proteins. Isotopic labeled samples are generally required for optimizing NMR solution conditions (Bagby et al., 1997; Lepre and Moore, 1998), NMR assignments and structure determination (Sattler et al., 1999), performing heteronuclear spin relaxation studies (Palmer and Bracken, 1999), biological screening applications such as 'SAR by NMR' (Shuker et al., 1996), and high throughput structural genomics studies (Montelione et al., 2000). Many of these investigations require large quantities of isotopically labeled proteins, the production of which is often a costly and time consuming aspect of NMR studies.

Currently, most isotopically labeled recombinant proteins are expressed in the bacterial host *E. coli*. Methods of generating heteronuclear labeled samples in *E. coli* commonly use standard or modified versions of M9 minimal media (Sambrook et al., 1989) employing ^{13}C glucose for carbon labeling, ^{15}N ammonium sulfate or ^{15}N ammonium chloride for nitrogen labeling, and deuterium oxide for deuteration. 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 (Reilly and Fairbrother, 1994; Jansson et al., 1996; Kainosho, 1997; Cai et al., 1998). Despite these advances, it is desirable to achieve higher expression yields at lower isotope costs using commonly employed shake flask growths.

In this communication, we present a simple procedure for the production of isotopically labeled proteins that permits high yields of labeled protein while significantly reducing the cost in both isotopes and production time. The method employs two new features. Cell mass is predominately grown on unlabeled rich media allowing rapid growth to high cell densities. Following growth in unlabeled medium, cells are exchanged into an isotopically defined minimal media at higher cell densities optimized for maximal protein expression. The isotopic incorporation rates obtained were 95% for ^{13}C, 90% for ^{15}N and ∼56% overall ^{2}H incorporation for protein expressed in 75% deuterium oxide. Expression yields in all cases were increased fourfold to eightfold per unit isotope.

*To whom correspondence should be addressed. E-mail: wcb2001@med.cornell.edu

Table 1. Recipe for 1 L of supplemented minimal growth medium and 1 L of 5 × M9 salts (autoclaved)

Minimal growth medium	
5 × M9 Salts	200.0 ml
D-glucose Stock (20 g/100 ml) (0.2_m filter sterilized)[a]	20.0 ml
Basal Vitamins Eagle Media[b]	10.0 ml
1 M $MgSO_4$ (autoclaved)	2.0 ml
1 M $CaCl_2$ (autoclaved)	0.1 ml
5 × M9 Salts (autoclaved)	
KH_2PO_4	15.0 g
$Na_2HPO_4 \bullet 7H_2O$	64.0 g
NaCl	2.5 g
NH_4Cl[a] (pH adjusted to 7.2 with NaOH)	5.0 g

[a]For isotopic labeling $^{15}NH_4Cl$ and ^{13}C D-glucose were substituted for unlabeled NH_4Cl and D-glucose respectively. For incorporation of deuterium, 75% $^{2}H_2O$ was used in place of deionized water for preparing the solution.
[b]Basal Vitamins Eagle Media was purchased from Life Sciences Technology. The above recipe is a modification of M9 minimal media (Sambrook et al., 1989).

The expression protocol was developed incorporating the following ideas. First, the growth rate for *E. coli* in minimal media is generally slower. Longer growth rates can result in reduced expression due to cytotoxic effects associated with the plasmid and plasmid gene products exerting selective pressure (Makrides, 1996; Baneyx, 1999). In addition, the rate of protein expression increases exponentially in relation to increasing *E. coli* growth rates. A change in *E. coli* doubling times from 100 min to 24 min results in a 10-fold increase in the number of ribosomes per cell (Bremer and Dennis, 1987). Therefore, rapid *E. coli* growth rates can have favorable effects on protein yields. Second, Cai and co-workers have demonstrated, using carefully controlled fermenter growths in minimal media, that isotopically labeled cell mass is not a requirement for high level $^{13}C/^{15}N$ incorporation (Cai et al., 1998). Therefore, isotope consumption can be reduced by generating the majority of cell mass using unlabeled media. Third, exchanging cells into fresh growth media immediately prior to induction can increase protein expression levels by removing byproducts inhibitory to growth and expression. Fourth, the cell density for optimal induction may differ for freshly exchanged media and conditioned media.

We tested the method on three proteins, an outer membrane lipoprotein found in gram negative bacteria designated Lpp-56 (Shu et al., 2000), ubiquitin (Ecker et al., 1987) and the C-terminal domain of angiopoietin-2 designated ANG2. Lpp-56 was ex-

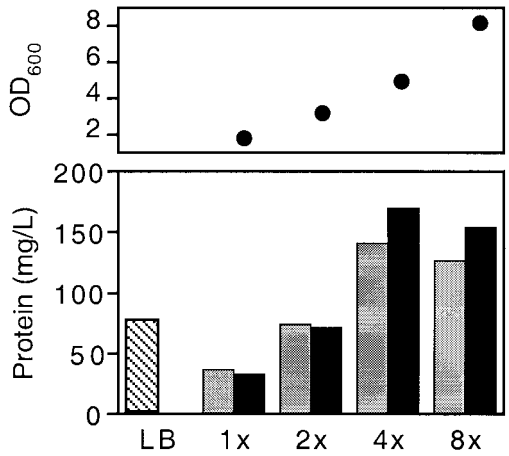

Figure 1. Effect of concentration on protein expression level. The top graph displays the final OD_{600} at harvest as a function of cell concentration for the ^{12}C minimal media growth. The first column of the bar graph (LB) shows the control expression of ANG2 grown on 1 L of LB media (striped bar), in a 2 L baffled shake flask, to $OD_{600} \approx 0.7$ followed by a 4 h induction. Minimal media samples were grown on 250 ml of either ^{13}C or ^{12}C labeled minimal media (Table 1) in 1 L baffled shake flasks, denoted by gray and black bars, respectively. The concentrations of glucose in the ^{13}C and ^{12}C growths were 4 g/L and 8 g/L, respectively. The columns 1×, 2×, 4×, and 8× indicate the concentration factor that corresponds respectively to 250 ml, 500 ml, 1000 ml and 2000 ml of LB growth harvested at $OD_{600} \approx 0.7$, then exchanged into 250 ml of minimal media for induction. After induction and harvest, the cells were chemically lysed, then flash frozen in liquid nitrogen and stored at −80 °C for later purification using NiNTA affinity chromatography (Qiagen). The purified protein was judged to be >95% pure by gel electrophoresis. Expression yields were quantified by UV absorbance using a calculated extinction coefficient of $\epsilon_{278} = 2.27$ Abs. cm^{-1} mg^{-1} (Pace and Schmid, 1997).

pressed using the pLPP56 plasmid (Shu et al., 2000) encoded in the pAED4 vector (Doering, 1992). The ubiquitin construct pGLUB was a generous gift from Dr. Tracy Handel (Lazar et al., 1997). The ANG2 sequence was encoded in the pET27 vector (Novagen), as a fusion to a cleavable N-terminal His-tag for affinity purification. All proteins were expressed in the *E. coli* BL21(DE3)/pLysS strain using the T7 expression system (Studier et al., 1990). The general protocol for isotope labeling is as follows. Cells were grown in 1 L of LB at 37 °C shaken at 240 rpm. Upon reaching optical cell densities at 600 nm (OD_{600}) ~0.7, the cells were pelleted by a 30 min centrifugation at 5000 × g. The cells were then washed and pelleted using an M9 salt solution, excluding all nitrogen and carbon sources. The cell pellet was resuspended in isotopically labeled minimal media (Table 1), then incubated to allow the recovery of growth and clearance of unlabeled metabolites. Protein expression was induced after 1 h by addition of isopropylthio-β-D-galactoside to a concentration of 0.8 mM. After a 4 h incubation period the cells were harvested. For isotope labeling the expression medium substituted one or more of the following isotopically labeled chemicals: $^{15}NH_4Cl$ (99%), ^{1}H-, ^{13}C-glucose (98%), and deuterium oxide (75% ^{2}H labeled) (Cambridge Isotope Laboratories).

To determine the optimal cell concentration for induction and expression in freshly exchanged minimal media, cells grown on unlabeled LB-broth were harvested at $OD_{600}\approx0.7$ during mid-log phase growth. The cell pellets were resuspended in 250 ml of either ^{13}C labeled minimal media at cell concentrations onefold (1×), twofold (2×), fourfold (4×), and eightfold (8×) greater relative to the originating LB cell growths (e.g. 1× results in an $OD_{600}\approx0.7$ and 8× results in $OD_{600}\approx5.6$ suspended in 250 ml of minimal media). An identical series of minimal media growths using unlabeled ^{12}C-glucose was tested at a higher concentration of glucose (8 g/L) to determine if glucose concentration limits expression at high cell densities. A 1 L control growth using LB-broth was used to determine the protein yield achieved for growth and induction on rich media. In this case the same media was used for both the growth and induction at $OD_{600}\approx0.7$. The protein yields from the rich media, ^{13}C minimal media, and ^{12}C minimal media growths are shown in Figure 1. The standard rich media growth (LB) generates ~75 ± 6 mg/L LB (n = 4 growths) of pure protein, whereas the 1× concentrated minimal media growth displays roughly half the expression level observed in rich media. The highest yields were attained by concentrating the LB growths 4× in minimal media prior to induction, resulting in a purified protein yield of 140 mg/L ^{13}C minimal media and 165 mg/L in the ^{12}C minimal media growth. The modest expression gains observed between ^{13}C and the ^{12}C growths upon a twofold increase in glucose concentrations indicates that glucose starvation is not a significant limiting factor.

The molecular masses of ANG2 proteins in the ^{13}C-labeling experiments were determined using MALDI-TOF. For the four ^{13}C-labeling experiments, the average molecular weight was 28985 ± 25 Da, the calculated masses for unlabeled and 100% ^{13}C-labeled protein are 27824 and 29047 Da respectively, indicating 95 ± 2% ^{13}C incorporation. The extent of ^{15}N label incorporation were tested by expressing the protein ubiquitin using the optimized labeling procedure. Out of a 250 mL minimal media expression 25.5 mg of HPLC purified protein was obtained with ~90% ^{15}N label incorporation. This is a fourfold improvement over the highest expression yields (~25 mg/L minimal media) that we have obtained for ubiquitin expressed using standard minimal media growth protocols.

We employed the method to generate a $^{2}H/^{13}C/^{15}N$ labeled sample of the Lpp-56 protein. Lpp-56 forms a highly anisotropic 18 kDa α-helical coiled-coil trimer composed of three 6 kDa monomeric subunits. Prior attempts at heteronuclear $^{2}H/^{13}C/^{15}N$ labeling of the Lpp-56 protein resulted in complete loss of protein expression following the prolonged growth in deuterated media using established $^{2}H/^{13}C/^{15}N$ labeling protocols (Gardner and Kay, 1998). For expression in 75% deuterated media, cell pellets were transferred from a 1 L LB growth (100% $^{1}H_2O$) upon reaching $OD_{600}\approx0.7$ to 250 ml 75% $^{2}H_2O$ $^{13}C/^{15}N$ minimal media. After transfer into deuterated media a 1 h growth lag was observed followed by a slow increase in OD_{600}. The cells were induced after 1 h, then harvested after a 9 h induction period. The protein was purified by HPLC and lyophilized, producing 18 mg of pure protein, corresponding to an expression yield of 72 mg/L minimal media. This yield compares favorably with the 80–100 mg/L LB yields we typically obtain from the rich-broth expressions of Lpp-56. The incorporation rates for ^{15}N and ^{13}C were estimated by NMR to be ~90% and ~85%, respectively. The extent of $^{15}N/^{13}C$ incorporation is slightly lower than in the individual isotopically labeled growths. However, the high expression yield overcompensates for the sensitivity loss since higher concentrations and/or larger sample volumes can be employed. Increases in the iso-

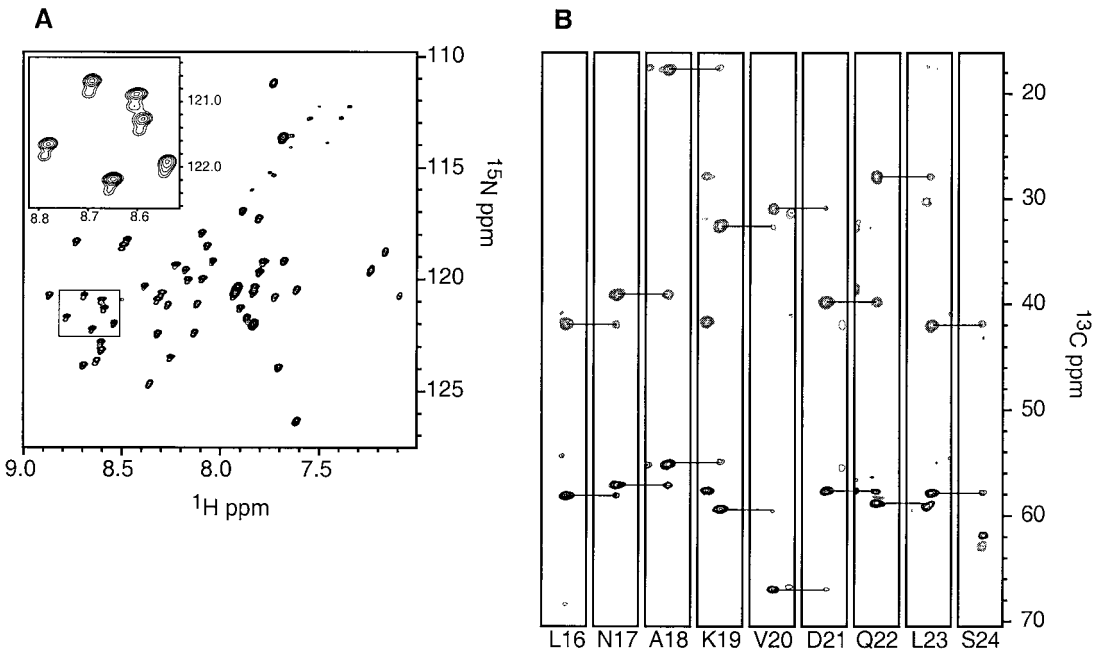

Figure 2. (A) Heteronuclear sensitivity enhanced ^{1}H-^{15}N TROSY spectra (Rance et al., 1999) and (B) strip plots from the sensitivity enhanced TROSY-HNCACB spectra displaying correlations for residues 16 through 24 of ~1.6 mM (monomer concentration) ^{2}H/^{13}C/^{15}N-labeled Lpp-56 in 93% H_2O/7% D_2O, 50 mM NaPi, 0.1% NaN_3, pH 5.8 at 35 °C acquired on a Varian Inova 600 NMR spectrometer. The ^{15}N-TROSY spectrum of the Lpp-56 molecule displays well-resolved resonances. The expansion of the peaks, shown in the inset box in Figure 2A, displays an isotope induced chemical shift effect on ^{15}N. The stronger upfield shifted cross peak is attached to $^{2}H^{\alpha}$ and the weaker downfield shifted peak is due to the $^{1}H^{\alpha}$ isotope species. The excellent signal-to-noise observed in both spectra demonstrates the success of the labeling procedure.

topic incorporation rates may be obtained by including a longer period for clearance of unlabeled nutrients before induction, to account for reduction of *E. coli* metabolic rates in deuterated media. The experimentally determined mass was 6653.1 Da for the triply labeled sample compared to the unlabeled average mass of 6153.7 Da. This indicates an overall deuterium incorporation level of ~56%. The sensitivity enhanced ^{15}N-TROSY spectrum (Figure 2) displays peaks with a deuterium isotope shift on the N^{H} due to the adjacent $^{2}H^{\alpha}$ (Figure 2, inset). The backbone resonances were assigned using a suite of TROSY pulse sequences (Loria et al., 1999; Salzmann et al., 1999) acquired on the ^{2}H/^{13}C/^{15}N labeled Lpp-56 sample. Figure 2B shows slices taken from the sensitivity enhanced TROSY-HNCACB experiment displaying the sequential connectivity of the polypeptide backbone of residues 16 through 24 in the coiled-coil of ^{2}H/^{13}C/^{15}N labeled Lpp-56; the excellent sensitivity observed validates the labeling strategy.

A number of distinctly different recombinant proteins have been expressed using the production method, in all cases improved protein yields were observed at reduced isotope costs. We find that the extent of isotope incorporation is up to 95% for ^{13}C and ~90% for ^{15}N, allowing a majority of NMR experiments to be successfully run. Using a standard suite of triple resonance pulse sequences, complete backbone assignments have now been obtained on two distinct protein samples produced using this method. The method offers an alternative to fermenter growths that require additional setup time and specialized equipment which is economically prohibitive for parallel expressions of recombinant proteins. The time preparing isotopically labeled samples using this method is slightly greater than growth on unlabeled rich media, offering a time savings in comparison to standard minimal media growth procedures which require significantly longer periods of cell growth to achieve cell densities suitable for expression. Additionally, the smaller media volumes required make the production method ideal for selective side-chain or amino acid labeling procedures (Goto and Kay, 2000) that employ expensive isotope labeled amino acids or precursors.

Acknowledgements

This research has been supported by a Grant-In-Aid from the American Heart Association and Stanley Stahl fellow to C.B. and by NIH grant AI42382 to M.L. We thank Julie Kneller for help in protein expression; David Eliezer, J. Patrick Loria, and Tam Nguyen for helpful discussions; and Arthur G. Palmer III for critical reading of the manuscript.

References

Bagby, S., Tong, K.I., Liu, D., Alattia, J.R. and Ikura, M. (1997) *J. Biomol. NMR*, **10**, 279–282.

Baneyx, F. (1999) *Curr. Opin. Biotechnol.*, **10**, 411–441.

Bremer, H. and Dennis, P.P. (1987) In *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology* (Neidhart, F.C., Ed.), American Society of Microbiology, Washington, DC, pp. 1527–1542.

Cai, M., Ying, H., Sakaguchi, K., Clore, G.M., Gronenborn, A.M. and Craigie, R. (1998) *J. Biomol. NMR*, **11**, 97–102.

Doering, D. (1992) *Functional and Structural Studies of a Small f-actin Binding Protein*, Massachusetts Institute of Technology, Cambridge, MA.

Ecker, D.J., Butt, T.R., Marsh, J., Sternberg, E.J., Margolis, N., Monia, B.P., Jonnalagadda, S., Khan, M.I., Wever, P.L., Mueller, L. and Crooke, S.T. (1987) *J. Biol. Chem.*, **262**, 14213–14221.

Gardner, K.H. and Kay, L.E. (1998) *Annu. Rev. Biomol. Struct.*, **27**, 357–406.

Goto, N.K. and Kay, L.E. (2000) *Curr. Opin. Struct. Biol.*, **10**, 585–592.

Jansson, M., Li, Y.C., Anderson, S., Montelione, G. and Nilsson, B. (1996) *J. Biomol. NMR*, **7**, 131–141.

Kainosho, M. (1997) *Nat. Struct. Biol.*, **4**, 854–857.

Lazar, G.A., Desjarlais, J.R. and Handel, T.M. (1997) *Protein Sci.*, **6**, 1167–1178.

Lepre, C.A. and Moore, J.M. (1998) *J. Biomol. NMR*, **12**, 493–499.

Loria, J.P., Rance, M. and Palmer, A.G. (1999) *J. Magn. Reson.*, **141**, 180–184.

Makrides, S.C. (1996) *Microbiol. Rev.*, **60**, 512–538.

Montelione, G.T., Zheng, D., Huang, Y.J., Gunsalus, K.C. and Szyperski, T. (2000) *Nat. Struct. Biol.*, **7**, 982–985.

Pace, C.N. and Schmid, F.Z. (1997) In *Protein Structure: A Practical Approach* (Creighton, T.E., Ed.), IRL Press at Oxford Press, New York, NY.

Palmer, A.G. and Bracken, C. (1999) In *NMR in Supramolecular Chemistry* (Pons, M., Ed.), Kluwer Academic Publishers, Dordrecht, pp. 171–190.

Rance, M., Loria, J.P. and Palmer, A.G. (1999) *J. Magn. Reson.*, **136**, 92–101.

Reilly, D. and Fairbrother, W.J. (1994) *J. Biomol. NMR*, **4**, 459–462.

Salzmann, M., Wider, G., Pervushin, K., Senn, H. and Wüthrich, K. (1999) *J. Am. Chem. Soc.*, **121**, 844–848.

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

Sattler, M., Schleucher, J. and Griesinger, C. (1999) *Prog. NMR Spectrosc.*, **34**, 93–158.

Shu, W., Liu, J., Ji, H. and Lu, M. (2000) *J. Mol. Biol.*, **299**, 1101–1112.

Shuker, S.B., Hajduk, P.J., Meadows, R.P. and Fesik, S.W. (1996) *Science*, **274**, 1531–1534.

Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorf, J.W. (1990) *Methods Enzymol.*, **185**, 60–89.