Journal of Biomolecular NMR, 5 (1995) 93–96
ESCOM

J-Bio NMR 240

A novel method for selective isotope labeling of bacterially expressed proteins

Karen M. Lee[a], Elliot J. Androphy[b] and James D. Baleja[a,*]

[a]*Department of Biochemistry, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111, U.S.A.*
[b]*Department of Dermatology, New England Medical Center and Tufts University School of Medicine, 750 Washington Street, Boston, MA 02111, U.S.A.*

Received 11 July 1994
Accepted 26 October 1994

Keywords: Protein expression; T7 RNA polymerase; Rifampicin; Papillomavirus; E2

Summary

A novel method for isotope labeling in selected amino acids is presented for use with the T7 RNA polymerase system. The protocol is illustrated with the DNA-binding domain from the E2 protein of bovine papillomavirus, BPV-1. On addition of rifampicin, protein expression occurs exclusively from the gene controlled by the T7 promoter. Since the bacteria are now dedicated to the production of E2 protein, labeling with specific amino acids is efficiently performed. For example, 10 mg/l of ^{15}N-labeled phenylalanine is shown to be sufficient for incorporation of the label, without scrambling, and without the use of an auxotrophic strain.

Recent advances in NMR have made it possible to determine the solution conformations of medium-sized proteins up to several hundred amino acids long. The new methods involve implementation of triple-resonance experiments that require the protein to be labeled with the NMR-observable isotopes ^{15}N and ^{13}C. Since isotope-labeled nutrients are expensive, optimization of protein production is needed. At the same time, spectra can be difficult to interpret, and proteins that are selectively labeled in a specific amino acid are advantageous for resonance assignments.

The T7 RNA polymerase system is one of the most widely used for bacterially expressed proteins (Tabor and Richardson, 1985; Studier and Moffatt, 1986; Studier et al., 1990). In this system the gene encoding the protein of interest is placed under the control of the T7 promoter, which is recognized as a translation start site by the RNA polymerase from phage T7, but not by the RNA polymerase of host *E. coli*. T7 RNA polymerase is supplied either by T7 phage infection, or, more conveniently, by induction of the T7 RNA polymerase gene under an IPTG-inducible promoter. Since the T7 RNA polymerase elongates chains about five times faster than does *E. coli* RNA polymerase, the proteins from genes preceded by the T7 promoter are often (but not always) expressed at high levels.

Rifampicin is an antibiotic that acts on the RNA polymerase of *E. coli*, but not the RNA polymerase of phage T7 (Wehrli et al., 1968; Studier et al., 1990). After high levels of T7 RNA polymerase are induced, rifampicin can be added to inhibit transcription of genes by the host enzyme and to allow only transcription of genes directed by the T7 RNA polymerase promoter (Tabor and Richardson, 1985; Dubendorff and Studier, 1991; Duverger et al., 1991; Kashiwagi et al., 1991; Nevin and Pratt, 1991; Kampfenkel and Braun, 1993; Notarnicola and Richardson, 1993). Radioactive isotopes added to the cell culture after rifampicin addition are observed to incorporate only into proteins whose genes are being transcribed by T7 RNA polymerase (Tabor and Richardson, 1985).

High-level expression systems other than the T7 RNA polymerase system exist where the cellular machinery is directing expression of mainly one protein. On addition of an NMR-active isotope-labeled amino acid at induction, rapid production of the protein has been observed in these systems to circumvent isotopic scrambling and

*To whom correspondence should be addressed.

Abbreviations: IPTG, isopropylthiogalactoside; SR media, synthetic rich media; OD_{600}, optical density or absorbance at 600 nm; HSQC, heteronuclear single-quantum coherence; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

allow selective labeling by amino acid residue type, even without the use of amino acid auxotrophs (Hibler et al., 1989; Muchmore et al., 1989). Without the addition of rifampicin, our protein, the DNA-binding domain (residues 310–410) of the E2 protein from bovine papillomavirus, BPV-1, is expressed in the T7 RNA polymerase system at relatively low levels (see below). Therefore, we wanted to see if the addition of rifampicin would increase protein production and if the level of protein expression was sufficiently high to evaluate the efficiency of incorporating specifically labeled amino acids into the protein.

The gene encoding the E2 protein was placed behind a T7 RNA polymerase promoter in the pET series of plasmid vectors, which also confer ampicillin resistance (Studier et al., 1990; Prakash et al., 1992). As a result of the cloning procedure, two amino acids (Met-Ala) are added to the N-terminus of the E2 protein. The pET plasmid was used to transform BL21(DE3) cells, which carry the gene for T7 RNA polymerase under an IPTG-inducible promoter. In addition, the cells harbor a pLysS plasmid that produces low levels of T7 lysozyme and confers chloramphenicol resistance. T7 lysozyme inhibits T7 RNA polymerase in uninduced cells, allowing the gene for a mildly toxic protein to be stably carried under the T7 promoter. Upon addition of IPTG, the higher level of T7 RNA polymerase overwhelms the inhibitory effect of the lysozyme (Studier et al., 1990).

Colonies were grown overnight in LB broth containing 25 mg/l ampicillin and 25 mg/l chloramphenicol at 29 °C (Sambrook et al., 1989). In order to determine the effect of rifampicin on the level of protein expression, 50 ml cultures were inoculated at a ratio of 1:50 and grown at 37 °C. The media used was synthetic rich (SR) broth, containing amino acid chemicals (Muchmore et al., 1989). Flasks were induced with 0.4 mM IPTG when the cultures reached an OD_{600} of 0.5, and 0–50 mg/l rifampicin was added 1.5 h after induction. 0.1 ml of the cell cultures were harvested 4.5 h post-induction. Typical cell OD_{600} readings ranged from 1.91 for the flask without rifampicin to 1.36 for the flask with 50 mg/l rifampicin. 40 µl of sample buffer (1.5 M urea, 20% glycerol, 2% SDS, 1 mM β-mercaptoethanol, 62.5 mM Tris.HCl, pH 6.8) was added and the samples boiled for 2 min. To each well, 10 µl was applied in a 15% polyacrylamide gel developed using a Laemmli system (Sambrook et al., 1989). Proteins were visualized using Coomassie Blue staining.

The addition of rifampicin clearly results in higher production of E2 protein (Fig. 1). The first lane is a control and contains uninduced cells with no IPTG or rifampicin added. The second lane illustrates the 'standard' induction conditions, using 0.4 mM IPTG and no rifampicin. Although the amount of E2 protein is low (densitometry shows about 2.8% of total cell protein), this is sufficient to produce a 1 mM sample (in 400 µl) from a 1 l culture. However, on addition of rifampicin, E2 protein production increases, and at 25–50 mg/l of rifampicin, E2 is now the dominant protein in the cell (about 21% from densitometric tracings). This allows for production of a 2 mM sample from a 500 ml culture (about four times higher). Note that, although the percentage of E2 protein in the rifampicin culture is about seven times higher (21% vs. 2.8%), the total amount of protein is lower (about 60%). We have also examined the effect of rifampicin in LB broth and in minimal media, where the increase in E2 protein production is similar, and we have checked E2 production with the higher concentration of rifampicin (200 mg/l) that is typically used (Tabor and Richardson, 1985; Dubendorff and Studier, 1991; Duverger et al., 1991; Kashiwagi et al., 1991; Nevin and Pratt, 1991; Kampfenkel and Braun, 1993; Notarnicola and Richardson, 1993). Although E2 protein production is still about 20% of total protein, the total amount of protein produced was found to be about 50% lower, and therefore E2 protein production was suboptimal.

For other proteins that are very efficiently expressed (for example to levels similar to the last lane of Fig. 1), the addition of rifampicin appears to have marginal effect (Studier and Moffatt, 1986), suggesting that the T7 RNA polymerase is already effectively competing with host *E. coli* RNA polymerase. In our case, during induction of T7 RNA polymerase by IPTG alone, it appears that the E2 message is either not being transcribed or translated efficiently. By inhibition of the host *E. coli* RNA polymerase, the nucleotides, and subsequently the tRNA and amino acids, become available exclusively for T7 RNA polymerase (Ellard, 1991). Although addition of rifampi-

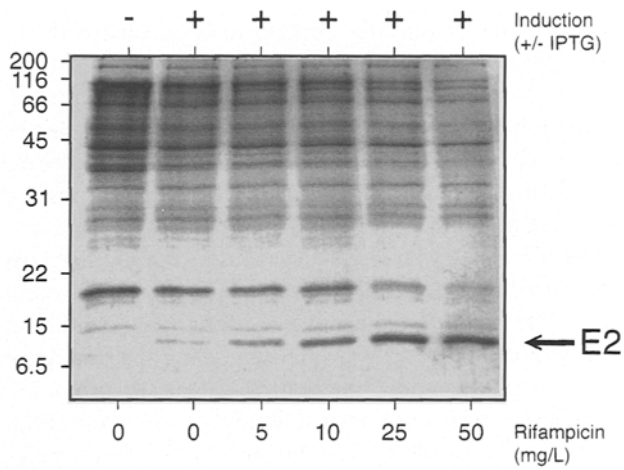

Fig. 1. Polyacrylamide gel, illustrating that addition of rifampicin increases expression of E2 protein. Equal volumes of *E. coli* culture were centrifuged and the cell pellets were boiled in SDS. The proteins were separated by SDS-PAGE on a 15% gel and visualized by Coomassie Blue staining. In cultures marked '+' for IPTG, the final concentration of IPTG was 0.4 mM. The concentration of rifampicin added after induction with IPTG is shown at the bottom. Molecular weight standards are shown at the left, with the molecular mass given in kDa. The position of the DNA-binding domain of E2 protein is indicated by an arrow.

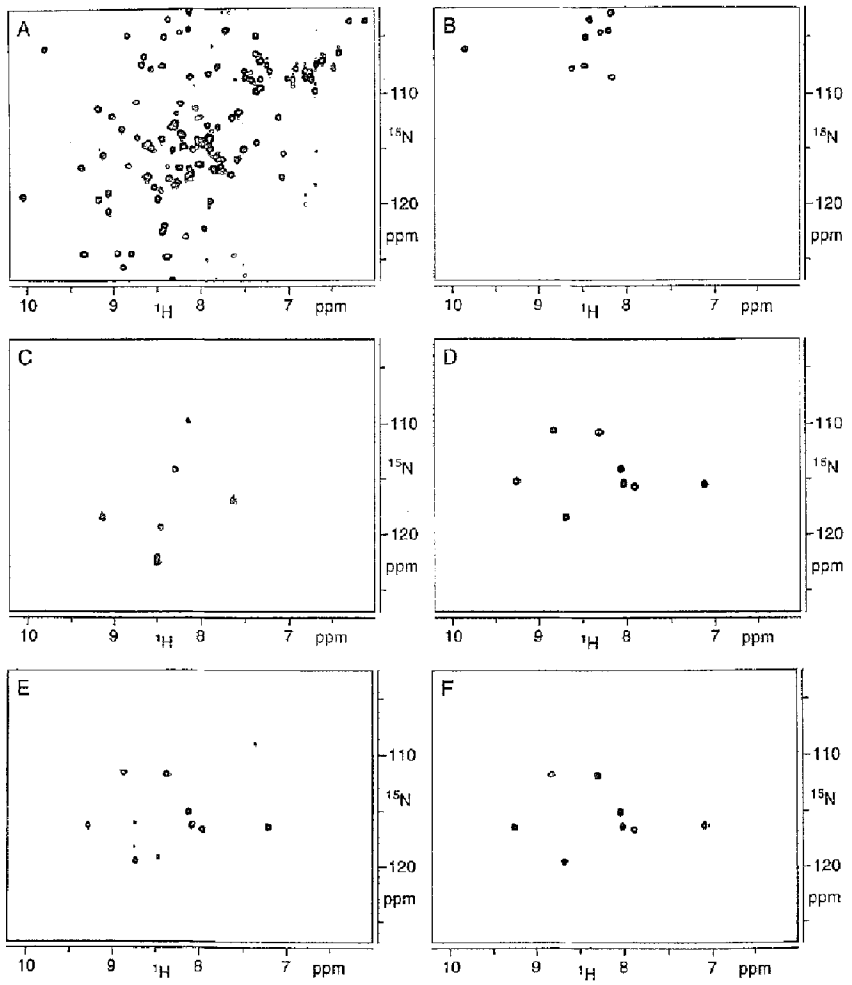

Fig. 2. ^{15}N-^{1}H HSQC correlation spectra of E2 protein. Selectively ^{15}N-enriched E2 proteins were produced by adding the ^{15}N-labeled amino acid at the time of induction. For each of the 128 increments, 32 scans were recorded, resulting in a total collection time of 1.5 h. (A) Uniform ^{15}N-labeled; (B) ^{15}N-glycine; (C) ^{15}N-leucine; (D) ^{15}N-phenylalanine (50 mg); (E) ^{15}N-phenylalanine, added both before induction (75 mg) and after induction (50 mg); (F) ^{15}N-phenylalanine (10 mg). In spectra C and E, peaks near 9.1 and 7.6 ppm are slightly distorted because of the post-acquisition removal of streaks in the ^{1}H dimension.

cin at 1.5 h post-induction suppresses further production of T7 RNA polymerase, the enzyme appears to be stable and to continue to transcribe from its promoter. Almost all protein produced over the next 3 h is E2 protein, while the other proteins of *E. coli* undergo degradation. This results in both higher amounts of E2 protein and lower background levels (Tabor and Richardson, 1985).

Given that almost all new protein production was E2, it seemed plausible to attempt specific ^{15}N labeling without the use of amino acid auxotrophs, and yet still achieve high levels of ^{15}N-isotope incorporation. Other than the addition of rifampicin, the timing of addition of isotope-labeled material and the protein purification procedures were standard. Briefly, protein ^{15}N-labeled in specific amino acids was produced by using 500 ml of SR media, omitting the amino acid of interest. When the OD_{600} reached 0.5, cells were induced with 0.4 mM IPTG and the ^{15}N-labeled amino acid was added (glycine, 250 mg/l; leucine, 50 mg/l; phenylalanine, 50 mg/l). Rifampicin was added to 50 mg/l at 1.5 h, and cells were harvested by centrifugation 3 h later. E2 protein was purified from cell lysate by Sepharose fast flow ion exchange and by G50 gel filtration (Pharmacia, Piscataway, NJ) chromatography (Prakash et al., 1992). ^{15}N-^{1}H correlation NMR spectra were recorded on 0.4 to 1.3 mM samples at 35 °C, pH 5.5–6.1, using a heteronuclear single-quantum coherence (HSQC) pulse sequence on a Bruker AMX500 spectrometer with a spin-lock for water suppression (Messerle et al., 1989).

Spectra are presented in Fig. 2 for protein uniformly ^{15}N-labeled in all amino acids and specifically ^{15}N-labeled in glycine, leucine and phenylalanine, respectively. E2 contains nine glycines, seven leucines and eight phenylalanines, which is equal to the number of ^{15}N-^{1}H cross peaks observed in the HSQC spectra (Figs. 2B–D). Despite an expectation that the label on glycine might end up on serine (Neidhardt, 1987; Muchmore et al., 1989), no scrambling of its label was observed (Fig. 2B). The extent of ^{15}N labeling is about 75%, with some ^{14}N-glycine incorporated from the endogenous pool (estimated to be about

80 mg/l; see below). In this case, an amino acid auxotroph would be useful. Also, no scrambling was observed with either ^{15}N-leucine (Fig. 2C) or ^{15}N-phenylalanine (Fig. 2D). Although our proteins are purified to greater than 97%, a high level of protein purity is probably unnecessary for 3D and ^{15}N-selective 2D experiments, since the contaminating background proteins are expected to be mostly unlabeled (Tabor and Richardson, 1985).

In one preparation of ^{15}N-phenylalanine-labeled E2, the amino acid was added twice during growth; it was included initially in the media, and again at 1 h after induction (total of 125 mg/l ^{15}N-phenylalanine). This method of labeling shows other amino acids that are labeled at low levels (about 10%; Fig. 2E). The low-intensity peaks have not been identified, but likely include glutamate, because of its central role in amino acid metabolism (Neidhart, 1987). The degree of scrambling observed is thus dependent on the moment the label is introduced during growth. If added much before T7 RNA polymerase induction, it is scrambled by amino acid interconversion. However, if added late after T7 RNA polymerase induction, any E2 protein produced up to that point would incorporate only ^{14}N-labeled amino acids. Moreover, the addition of rifampicin at a precise time is important for optimal protein production and label incorporation. We allow induction of T7 RNA polymerase for 1.5 h, so that sufficient levels of enzyme can build up before inhibiting host RNA polymerase with rifampicin (thus preventing further T7 RNA polymerase production).

The lowest amount of ^{15}N-phenylalanine that could be added and still yielded an informative HSQC spectrum was 10 mg/l (Fig. 2F). On the other hand, a protein produced with only 2 mg/l ^{15}N-phenylalanine gave an HSQC with low signal-to-noise and with some scrambling of the label (data not shown). To determine the size of the endogenous ^{14}N-phenylalanine pool from catabolism of the other ^{14}N amino acids, cross peaks to amide protons were examined in TOCSY spectra without decoupling of ^{15}N for the samples that were produced with 10, 50 and 125 mg/l ^{15}N-phenylalanine in the media. Amide protons attached to ^{15}N were identified by the $^{1}J_{^{15}N,HN}$ splitting of about 90 Hz, whereas those attached to ^{14}N appeared as a singlet in the middle of the doublet. All samples showed that about 30 mg/l ^{14}N-phenylalanine was produced by catabolism of other amino acids prior to induction of E2 protein expression, thus accounting for the low signal-to-noise ratio in the sample produced using only 2 mg/l ^{15}N-phenylalanine (only about 6% of the phenylalanine would be ^{15}N-labeled).

For the preparation of a 1 mM solution of protein, selectively ^{15}N-labeled in phenylalanine without the use of rifampicin, standard procedures would suggest the use of 1 l of media containing 200 mg of ^{15}N-phenylalanine (Hibler et al., 1989) (although we have not determined the ^{15}N-label requirement in this case). Alternatively, a phenylalanine auxotroph could have been generated. Using this, we would have needed about 50 mg/l of ^{15}N-phenylalanine (Muchmore et al., 1989). By way of contrast, the procedure using rifampicin demonstrates that the same sample can be produced by using 250 ml of media containing only 2.5 mg of ^{15}N-phenylalanine, which is a much more efficient use of an isotopically labeled amino acid.

In summary, rifampicin in the T7 RNA polymerase expression system dramatically improves production of E2 DNA-binding protein. This method was shown to allow the incorporation of ^{15}N-glycine, ^{15}N-leucine and ^{15}N-phenylalanine without scrambling. Amino acid auxotrophs were not required for these amino acids and therefore it appears that other amino acids also could be specifically labeled without using auxotrophs. Samples that were selectively ^{15}N-labeled by residue type have been very helpful in the sequential resonance assignment of this protein (unpublished experiments).

Acknowledgements

We thank Jim Sudmeier for maintenance of the NMR spectrometer facility. This work was supported in part by the Charlton Fund, Tufts University.

References

Dubendorff, J.W. and Studier, F.W. (1991) *J. Mol. Biol.*, **219**, 61–68.

Duverger, N., Murry-Brelier, A., Latta, M., Reboul, S., Castro, G., Mayaux, J.-F., Fruchart, J.-C., Taylor, J.M., Steinmetz, A. and Denèfle, P. (1991) *Eur. J. Biochem.*, **201**, 373–383.

Ellard, G.A. (1991) *J. Antimicrob. Chemother.*, **28**, 347–355.

Hibler, D.W., Harpold, L., Dell'Acqua, M., Pourmotabbed, T., Gerlt, J.A., Wilbe, J.A. and Bolton, P.H. (1989) *Methods Enzymol.*, **177**, 74–86.

Kampfenkel, K. and Braun, V. (1993) *J. Biol. Chem.*, **268**, 6050–6057.

Kashiwagi, K., Suzuki, T., Suzuki, F., Furuchi, T., Kobayashi, H. and Igarashi, K. (1991) *J. Biol. Chem.*, **266**, 20922–20927.

Messerle, B.A., Wider, G., Otting, G., Weber, C. and Wüthrich, K. (1989) *J. Magn. Reson.*, **85**, 608–613.

Muchmore, D.C., McIntosh, L.P., Russell, C.B., Anderson, D.E. and Dahlquist, F.W. (1989) *Methods Enzymol.*, **177**, 44–73.

Neidhardt, F.C. (Ed.) (1987) *Escherichia coli and Salmonella Typhimurium*. Vol. 1, American Society for Microbiology, Washington, DC.

Nevin, D.E. and Pratt, J.M. (1991) *FEBS Lett.*, **291**, 259–263.

Notarnicola, S.M. and Richardson, C.C. (1993) *J. Biol. Chem.*, **268**, 27198–27207.

Prakash, S.S., Grossman, S.R., Pepinsky, R.B., Laimins, L.A. and Androphy, E.J. (1992) *Genes Dev.*, **6**, 105–116.

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

Studier, F.W. and Moffatt, B.A. (1986) *J. Mol. Biol.*, **189**, 113–130.

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

Tabor, S. and Richardson, C.C. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1074–1078.

Wehrli, W., Knüsel, F., Schmid, K. and Staehelin, M. (1968) *Proc. Natl. Acad. Sci. USA*, **61**, 667–673.