J Biomol NMR (2008) 42:59–67
DOI 10.1007/s10858-008-9264-0

ARTICLE

A simple method for amino acid selective isotope labeling of recombinant proteins in *E. coli*

Kit I. Tong · Masayuki Yamamoto · Toshiyuki Tanaka

Received: 26 February 2008 / Accepted: 6 August 2008 / Published online: 2 September 2008

Abstract A simple and user-friendly method of labeling protein selectively with amino acids in vivo is introduced. This technique does not require the use of transaminase-deficient or auxotrophic strains. By manipulating the product feedback inhibitory loops of the *E. coli* amino acid metabolic pathways and, if necessary, by using enzyme inhibitors, proteins were labeled efficiently in vivo even with amino acid types that are central to the metabolic pathways, such as glutamine. The sequential backbone resonance assignment of the Neh2 domain of Nrf2 transcriptional factor, an intrinsically disordered protein with high spectral degeneracy, was achieved using this labeling method.

Keywords Amino acid selective isotope labeling · Intrinsically disordered protein · Neh2 · ^{15}N-glutamate labeling · ^{15}N-glutamine labeling

K. I. Tong · M. Yamamoto
Center for Tsukuba Advanced Research Alliance, Graduate School of Comprehensive Human Sciences, JST-ERATO Environmental Response Project, University of Tsukuba, Tsukuba 305-8577, Japan

Present Address:
K. I. Tong · M. Yamamoto (✉)
Department of Medical Biochemistry, Tohoku University Graduate School of Medicine, Sendai 980-8575, Japan
e-mail: masi@mail.tains.tohoku.ac.jp

T. Tanaka (✉)
Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba 305-8572, Japan
e-mail: ttanaka@tara.tsukuba.ac.jp

Introduction

Intrinsically unstructured proteins (Dyson and Wright 2005) lack intrinsic globular structures under physiological conditions, but often acquire inducible rigid tertiary folds upon binding with ligands or other proteins. When carrying out NMR structural analysis on these proteins, even partially folded, we often encounter extensive overlap of resonances and weak sequential nuclear Overhauser effects (NOEs). Together with the fact that carbon chemical shifts of some amino acid types, such as glutamine, glutamate, arginine, and histidine, are not typically unique, clusters of these amino acid types may lead to further ambiguity during structural analysis. These spectral complexities have made sequence-specific backbone resonance assignments of these proteins challenging.

Development of a series of heteronuclear three-dimensional (3D) NMR experiments taking advantage of slower spin relaxation rates of unfolded polypeptides in combination with constant-time evolution periods has helped to achieve resonance assignments of unfolded or partially unfolded proteins (Liu et al. 2000). Apart from these advances in NMR methodologies, another common approach is in vivo selective incorporation of ^{15}N- and/or ^{13}C-labeled amino acids into the proteins under study (Senn et al. 1987; Muchmore et al. 1989; Lee et al. 1995). Recently, there are many reports on preparing samples for NMR analysis by cell-free (CF) protein synthesis, which have the benefits of reducing reaction volumes, quantities of expensive or unusual labeled amino acids, and isotopic scrambling (Kigawa et al. 1995; Morita et al. 2004; Kainosho et al. 2006; Staunton et al. 2006). The CF system is not a cell-based system and, in principle, can incorporate a variety of reagents, such as protease inhibitors, chaperones, ligands, or detergents that may facilitate protein

synthesis in vitro, protein folding, and post-translational protein stability. Therefore, the CF system is especially useful for producing cytotoxic proteins and integral membrane proteins (Katzen et al. 2005). Automation of the CF synthesis has also allowed the implementation of high-throughput screening in NMR-based structural proteomics projects (Vinarov and Markley 2005). In addition, taking advantage of the CF synthesis and the dual amino acid-selective $^{13}C/^{15}N$ labeling technique (Kainosho and Tsuji 1982), partial sequence-specific backbone assignment has been achieved by a combinatorial selective labeling method (Parker et al. 2004).

The major disadvantage of the CF synthesis is low yields of protein since not all proteins are successfully synthesized in vitro to a satisfactory amount required for NMR data acquisition within a reasonable cost (Tyler et al. 2005; Staunton et al. 2006). For example, a wheat-germ based CF system preferable for eukaryotic protein production was reported to yield proteins varying from 0.1 to 2.3 mg/ml reaction (Endo and Sawasaki 2006). The cost of sufficient protein production for NMR analyses would be lower when one uses in-house cell extracts and reagents, but it demands both effort and expertise to maintain a good CF system in quality and quantity (Staunton et al. 2006). Low signal-to-noise ratio due to poor NMR sample production could also be compensated by new NMR technology such as ultra-high sensitivity cryoprobes, depending on their accessibility to individual laboratories. Therefore, an effective in vivo expression method for specific labeling remains a viable alternative for protein NMR.

Escherichia coli is a commonly used expression system in preparing protein samples for NMR studies. Stable isotope labeling has been well established for easy manipulation of culturing conditions in the bacterial system. Many methods have been presented in the past for selective amino acid labeling in vivo (Senn et al. 1987; Muchmore et al. 1989; Lee et al. 1995; Waugh 1996; Fiaux et al. 2004). Since there are metabolic pathways to shuffle one amino acid type to another in *E. coli*, the use of a transaminase-deficient or an auxotrophic mutant strain has been widely used. However, some of these auxotrophic mutant strains have low growing rate or yield, which leads to sub-optimal protein production. On the other hand, some heterologous proteins are not efficiently expressed in *E. coli* due to rare codon usage. Although *E. coli* strains that carry extra copies of rare codon tRNA genes for high-level expression are now attainable from the market, there is no commercially available strain harboring both transaminase deficiency and extra rare codon tRNA genes. Alternatively, one might mutate all of the rare codons within their coding regions to enhance the production of heterologous proteins in *E. coli*. However, the rare codons for arginine (AGA, AGG, and CGA), leucine (CUA), isoleucine (AUA), proline (CCC), and glycine (GGA and GGG), which were reported to create translational problems for cloned heterogeneous genes (Kane 1995), have high total occurrence in genes from eukaryotes (for example, approximately 9% for the typical human, mouse, or rat genes) (Wada et al. 1992). Therefore, it would be significantly laborious to mutate all rare codons in expressing constructs even when the actual percentage of rare codons in any particular gene is below the reported value.

Here we describe an in vivo expression method for amino acid selective labeling of proteins by manipulating the product feedback inhibitory loops of amino acid metabolic pathways without employing transaminase-deficient or auxotrophic mutant strains. For amino acids such as glutamine and glutamate that are easy to metabolize into other amino acid types, inhibitors of the enzymes responsible for these amino acid type inter-conversions were also used to improve labeling specificity. This method has been successfully applied, in combination with data obtained from uniform heteronuclear isotopic labeling, to achieve sequence-specific backbone resonance assignment of a non-globular protein, the regulatory Neh2 (Nrf2-ECH Homology 2) domain of Nrf2 (NF-E2 related factor 2) (Itoh et al. 1999; Tong et al. 2006a, b), which is a transcription factor for regulating cellular electrophilic/oxidative stress response.

Materials and methods

Expression of the recombinant protein

The plasmid containing mouse Neh2 cDNA (Met-1 to Gly-98; pET15b-Neh2; Tong et al. 2006a) was transformed into *E. coli* BL21-CodonPlus(DE3)-RIL cells (Stratagene) and cells were grown at 37°C in modified M9 minimal medium containing 1 g/l of ammonium chloride, 5 g/l of D-glucose, and 100 μg/ml of ampicillin. For uniformly ^{15}N-labeled protein, *E. coli* was grown in modified M9 minimal medium prepared with ^{15}N-ammonium chloride (Isotec Inc.) as the sole nitrogen source. Over-expression of the recombinant protein was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at OD_{600} of 0.7. Cells were harvested 4 h after IPTG induction. For glutamine-selective labeling, *E. coli* was grown in modified M9 minimal medium supplemented with 0.1 g/l of each non-labeled amino acid except for glutamine and glutamate. At OD_{600} of 0.7, 1 g/l of each non-labeled amino acid except for glutamine and glutamate, 2 g/l of non-labeled glutamate, 0.1 g/l of $^{15}N_1$-glutamine, 75 mg/l of 6-diazo-5-oxo-L-norleucine, 180 mg/l of L-methionine sulfoximine, and 180 mg/l of L-methionine sulfone (Sigma) were added to the culture. Over-expression was induced 15 min after the addition of the amino acids and inhibitors by adding a final

concentration of 0.5 mM IPTG. The culture was then harvested 90 min after IPTG induction. For ^{15}N-glutamate labeled Neh2, 0.1 g/l of each non-labeled amino acid except for glutamine and glutamate were added to modified M9 minimal medium. At 15 min prior to IPTG induction, 4 g of Asp, Gln, Tyr, and Ile; 3 g of Leu, Val, Phe, Ala, and Asn; 1 g of Met, Ser, Lys, Trp, Cys, His, Thr, Pro, Gly, and Arg; 1 g of L-methionine sulfoximine and L-methionine sulfone; 0.25 g of disodium succinate, disodium maleate, and aminooxyacetate; and 0.1 g of ^{15}N-glutamate were added to each liter of culture. Cells were harvested 90 min post IPTG induction. For labeling with $^{15}N_4$-arginine, $^{15}N_1$-leucine or $^{15}N_2$-lysine, cells were grown in modified M9 minimal medium supplemented with 0.1 g/l of all non-labeled amino acids except for arginine, leucine or lysine, respectively. At 15 min before induction, 0.1 g/l of $^{15}N_4$-arginine, $^{15}N_1$-leucine or $^{15}N_2$-lysine and 1 g/l of the rest of non-labeled amino acids were added to the culture medium. Cells were induced as mentioned above. For $^{15}N_2/^{13}C_5$-glutamine and α-^{15}N-lysine labeling, culture was grown in modified M9 minimal medium supplemented with 0.1 g/l of non-labeled amino acids except for glutamine, glutamate, and lysine. At 15 min before induction, 1 g/l of non-labeled amino acids except for glutamine, glutamate, and lysine, 2 g/l of non-labeled glutamate, 0.1 g/l of $^{15}N_2/^{13}C_5$-glutamine, 0.1 g/l of α-^{15}N-lysine, 75 mg/l of 6-diazo-5-oxo-L-norleucine, 180 mg/l of L-methionine sulfoximine, and 180 mg/l of L-methionine sulfone were added to the medium. Cells were harvested 90 min post IPTG induction. All isotopically labeled amino acids were purchased from Cambridge Isotope Laboratory.

Protein purification

Cells were resuspended in lysis buffer [20 mM Tris–HCl, pH 7.5, 300 mM NaCl, 0.5% (v/v) Triton X-100, 5 mM $MgCl_2$, 1 mg/ml of lysozyme, 20 μg/ml of DNaseI, 1 mM imidazole, 2 mM benzamidine-HCl, and 1 mM phenylmethylsulfonyl fluoride]. Cell lysis was achieved by sonicating the cell suspension using a Branson sonifier 450 on ice. Extracts were clarified by centrifugation at 27,000*g* for 30 min at 4°C.

His-tagged proteins in the soluble extracts were purified with Pro-Bond resin (Invitrogen) according to manufacturer's instructions. After dialyzing against 20 mM Tris–HCl, pH 7.5, 20% glycerol, and 1 mM dithiothreitol, the His-tag was removed by thrombin (Calbiochem). The protein was then further purified by Q2 (BIO-RAD) and Superdex S75 PG16/60 (Pharmacia) columns. For NMR data acquisition, the protein sample was concentrated to 1.0–1.2 mM and exchanged into 20 mM sodium phosphate, pH 8.0, 100 mM NaCl, 10 mM $[^2H]_{10}$-dithiothreitol, 1 mM 4-(2-aminoethyl)-benzenesulfonyl fluoride, and 90% H_2O/10% ^2H_2O (Isotec Inc.).

NMR experiments and data analysis

NMR experiments were run at 25°C on a Bruker AVANCE DRX 800 or a Varian UNITY INOVA 500 spectrometer, using 5-mm triple-resonance probes. Two-dimensional ^{1}H-^{15}N HSQC (Kay et al. 1992) and 3D HNCO (Muhandiram and Kay 1994) experiments were performed as previously described (Tong et al. 2006a). All data were processed using NMRPipe, NMRDraw (Delaglio et al. 1995) and analyzed using Pipp (Garrett et al. 1991).

MALDI-TOF mass spectrometry analysis

Approximately 50 μg of $^{15}N_1$-glutamine labeled Neh2 samples prepared without suppressing amino acid interconversion, with inhibitors, or with both inhibitors and excess non-labeled amino acids were separated on 15% SDS-PAGE. After staining with Coomassie blue G250, gel bands were excised with clean blades and chopped into small pieces. The gel plug was then destained with 50 mM ammonium bicarbonate/50% methanol, reduced with 10 mM dithiothreitol, alkylated with 55 mM iodoacetamide, and equilibrated using 100 mM ammonium bicarbonate. The in-gel enzymatic digestion was performed overnight (~16 h) at 37°C with endoproteinase Asp-N (Roche Applied Science) or trypsin (Promega). The digested peptides were extracted with 5% formic acid (FA)/50% acetonitrile (ACN), and the volume of extract was reduced to approximately 20 μl by speed vacuum. The extract was then purified using ZipTip μ-C18 (Millipore) and eluted in 5 μl of 0.1% FA/50% ACN. The eluted peptide sample was mixed 1:1 with the matrix solution (10 mg/ml of α-cyano-4-hydroxycinnamic acid in 0.1% FA/50% ACN), and then spotted on a 384-well stainless steel plate for matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS). Reflectron MS analyses were performed on a 4800 MALDI TOF/TOF™ Analyzer (Applied Biosystems). All ion spectra were recorded in the positive mode with an accelerating voltage of 20.0 kV. The spectrometer was externally calibrated using Cal Mix 1 and 2 standard mixtures (Applied Biosystems). The spectra were analyzed and compared using DataExplorer Software version 4.9 (Applied Biosystems). The simulation of the isotopic distribution pattern of the fragment, ^{21}DILWRQ^{26} or ^{35}EVFDFSQR^{42} was carried out as described elsewhere (Yabuki et al. 1998), but using an in-house program.

Results and discussion

Globally or partially disordered proteins usually give extreme spectral degeneracy, which complicates the

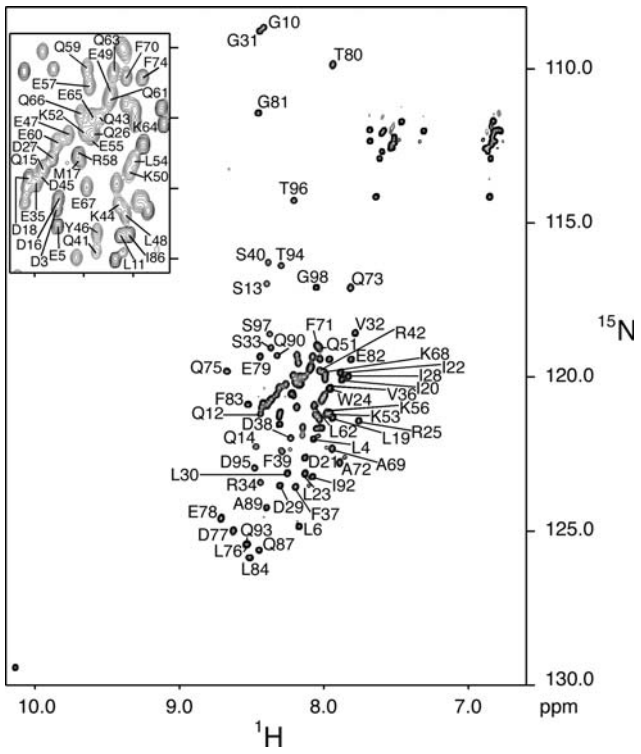

Fig. 1 ^{1}H-^{15}N HSQC spectrum of uniformly ^{15}N-labeled Neh2 recorded at 800 MHz and 25°C. Sequential assignments of the backbone resonances of Neh2 are as indicated by a one-letter amino acid code and a residue number

assignment process. Although the Neh2 domain of Nrf2 transcriptional factor comprises only 98 residues, the protein shows overlapped, narrowly dispersed amide resonance signals in the HSQC spectrum (Fig. 1). In addition, there are many amino acids for which carbon chemical shifts are not typically unique: 16 glutamines, 12 glutamates, 4 arginines, 7 lysines, 10 aspartates, and 11 leucines. The high content of these residues in the sequence of Neh2 (over 60%) made the correlations of NH resonances with their corresponding $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts rather ambiguous. Furthermore, there are multiple pairs of dipeptides made from the above non-unique amino acid types, for example three EK and three QK, in the Neh2 sequence. In order to overcome these ambiguities, the use of specific labeling was considered.

Neh2 contains rare codons for arginines (AGA, AGG, and two CGAs), isoleucine (ATA), leucine (CTA), and glycines (four GGAs). They account for approximately 10% of the nucleic acid sequence that encodes Neh2 and caused the low yield of the protein in *E. coli* without extra rare codon tRNA gene supplementation (Table 1). Since transaminase-deficient or auxotrophic strains bearing extra rare codon tRNA genes are not commercially available to date, the previously described in vivo expression method was not feasible. On the other hand, the protein yield of the Neh2 domain using an *E. coli*-based in vitro protein synthesis kit (Rapid Translation System, RTS, Roche) was also very low (approximately 20–30 μg/ml reaction). In addition, labeling glutamine or glutamate specifically by the Roche RTS system is not practical due to high concentration of L-glutamate salt in the reagents (Manufacturer's note; Staunton et al. 2006). Although glutathione *S*-transferase has been prepared in high quantity by using home made cell extracts from BL21(DE3) strain carrying a plasmid pMINOR that encodes minor tRNA genes (Chumpolkulwong et al. 2006), cell extracts prepared from another strain supplemented with rare tRNA genes, BL21-CodonPlus(DE3)-RIL (Stratagene), did not improve the yield for some other proteins (Tong and Ikura, unpublished observations). Hence, successful protein expression in the described CF system may not necessarily depend on codon

Table 1 Protein yield per liter in different bacterial hosts and culture media

	BL21(DE3)pLysS[a] (mg)	BL21-CodonPlus(DE3)-RIL[b] (mg)
Uniform ^{15}N[c]	1.0	4.0
^{15}N-glutamine with inhibitors and excess amino acids[d]	ND[f]	3.5
^{15}N-glutamate with inhibitors and excess amino acids[e]	ND[f]	3.0

[a] Novagen

[b] Stratagene

[c] Modified M9 minimal medium containing 1 g/l of ^{15}N-ammonium chloride as the sole nitrogen source, 5 g/l of D-glucose, and 100 μg/ml of ampicillin. Cells were grown at 37°C and harvested 4 h after IPTG induction

[d] Modified M9 minimal medium supplemented with 0.1 g/l of each non-labeled amino acid except Gln and Glu. Fifteen minutes prior to IPTG induction, 1 g of each non-labeled amino acid except Gln and Glu, 2 g of non-labeled Glu, 0.1 g of [$^{15}N_1$]-Gln, 75 mg of 6-diazo-5-oxo-L-norleucine, 180 mg of L-methionine sulfoximine, and 180 mg of L-methionine sulfone were added to each liter of culture. Cells were grown at 37°C and harvested 1.5 h after IPTG induction

[e] Modified M9 minimal medium supplemented with 0.1 g/l of each non-labeled amino acid except Glu and Gln. Fifteen minutes prior to IPTG induction, 1 g of each non-labeled amino acid except Glu, 0.1 g of [^{15}N]-Glu, 88 mg of 6-diazo-5-oxo-L-norleucine, 235 mg of L-methionine sulfoximine, and 235 mg of L-methionine sulfone were added to each liter of culture. Cells were grown at 37°C and harvested 1.5 h after IPTG induction

[f] Not determined

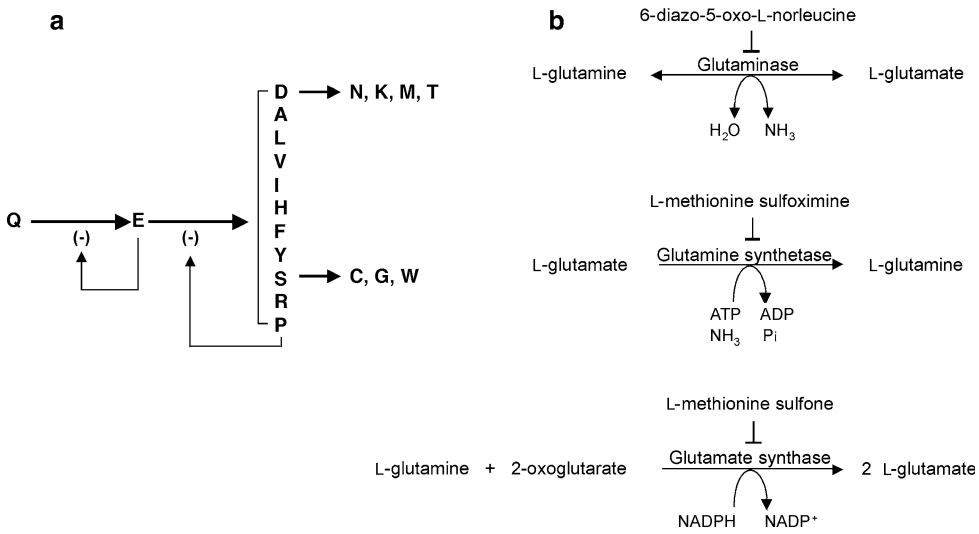

Fig. 2 Schematic diagrams of (**a**) a simplified and partial view of the end-product feedback inhibition for amino acid biosynthetic pathway, and (**b**) the enzymes and corresponding inhibitors involved in the metabolic conversion between glutamine and glutamate

usage alone. Consequently, we decided to use the *E. coli* strain most suitable for in vivo expression of Neh2.

One of the difficult amino acid types to be labeled in vivo is glutamine. Glutamine is easily transformed into glutamate. Once glutamate is being converted, there are pathways to shuffle its amino group to aspartate, alanine, leucine, valine, isoleucine, histidine, phenylalanine, tyrosine, serine, arginine, and proline (Fig. 2a). Therefore, labeling glutamine specifically in vivo becomes difficult. In an earlier work, side-chain amide groups of glutamine (and asparagine at the same time) were successfully labeled in *E. coli* by manipulating the ammonia assimilation in the metabolic pathways among asparagine, aspartate, glutamine, and glutamate (Tate et al. 1992). Another group, on the other hand, demonstrated selective labeling with 1-^{13}C glutamine, but they observed its scrambling to glutamate, proline, and arginine, and the absolute ^{13}C-labeling ratio was less than 20% (Takeuchi et al. 2007). When considering the metabolic pathway of glutamine, enzymes such as glutaminase, glutamine synthetase, and glutamate synthase catalyze the metabolic inter-conversion between glutamine and glutamate in *E. coli* (Fig. 2b). Inhibitors to these enzymes are known and can be used for amino acid specific labeling. In addition, there are many product feedback inhibitory mechanisms for regulating the amino group shuffling or the metabolic conversion from glutamate to other amino acid types (Fig. 2a; Muchmore et al. 1989; Keseler et al. 2005). Indeed, glutamate has been shown to inhibit the catalytic activity of glutamate synthase of *E. coli* (Miller and Stadtman 1972).

For completing backbone assignment, we had prepared NMR samples of the Neh2 domain with and without inhibitors for glutamate/glutamine inter-conversion or excess amount of non-labeled amino acid supplementation. Without any suppression, high numbers of undesired cross-peaks including glutamates, aspartates, leucines, isoleucines, valines, and phenylalanines, were seen when expressing Neh2 in modified M9 minimal medium supplemented with 0.1 g/l of each non-labeled amino acid and ^{15}N-glutamine (Fig. 3a; Table 2). Although inhibitors have efficiently controlled the conversion from glutamate to glutamine in the wheat-germ CF protein synthesis system (Morita et al. 2004), addition of inhibitors of bacterial glutaminase, glutamine synthetase, and glutamate synthase such as 6-diazo-5-oxo-L-norleucine (Prusiner and Stadtman 1976), L-methionine sulfoximine (Weisbrod and Meister 1973), and L-methionine sulfone (Rendina and Orme-Johnson 1978), respectively, was not efficient in enhancing specificity when labeling protein in bacteria (Fig. 3b). One possible reason is that enzymes for amino acid metabolism could be actively expressed and replenished in live cells, whereas these enzymes, if any in the CF system, can be easily exhausted, since this system is a semi-purified preparation that contains the necessary translation machinery for protein production and cannot respond to stimuli for transcriptional regulation unlike live cells. In addition, the efficiency of inhibitors in cells would be different depending on their inhibition mechanisms (irreversible inhibitions by 6-diazo-5-oxo-L-norleucine and L-methionine sulfoximine or a competitive inhibition by L-methionine sulfone). In our tested conditions, the relative intensity of the non-targeted ^{15}N-glutamate in the sample produced with inhibitors for glutamate/glutamine inter-conversion had only approximately 20% reduction when compared with that obtained without inhibitors (Table 2). In addition, aspartate, leucine, isoleucine, valine, and phenylalanine were also similarly cross-labeled.

However, there was drastic reduction of cross labeling when ten-fold of non-labeled amino acids were added to the medium, which raises their intra-cellular concentrations (Britten and McClure 1962); (Table 2). Although there were still leaky conversions to glutamates, aspartates,

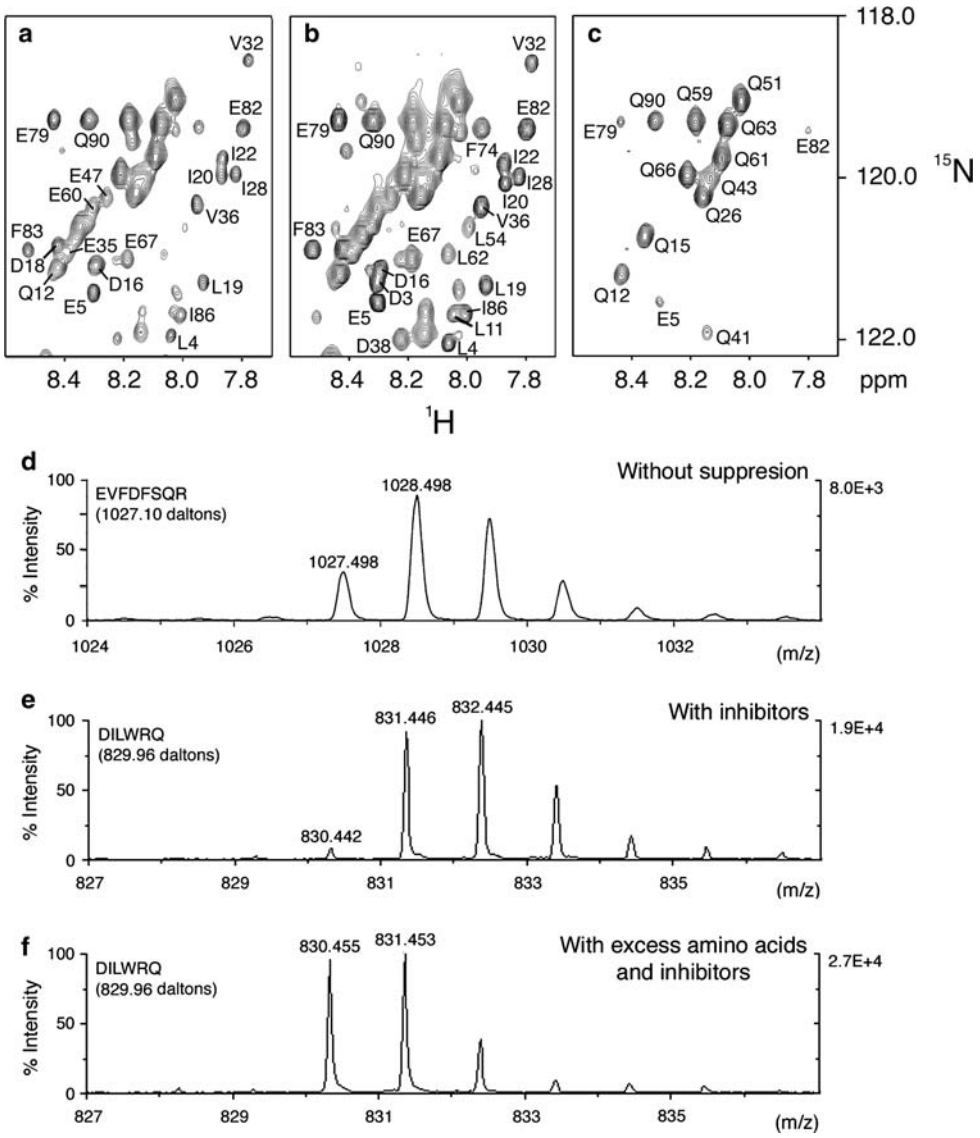

Fig. 3 (**a–c**) The overlapping region of the ^{1}H-^{15}N HSQC spectra of ^{15}N-glutamine labeled Neh2 expressed (**a**) without suppression, (**b**) with inhibitors for glutamate/glutamine inter-conversion, or (**c**) with both excess amino acids and inhibitors as described in Table 2. These spectra were recorded and plotted identically using the same protein concentration. Difference in the peak intensities of glutamine among the spectra reflects the level of ^{15}N enrichment of each sample as described in the text. Some of the isolated peaks are labeled. Eleven out of the total of sixteen glutamines of Neh2, which exist in the overlapping region, are indicated in (**c**). (**d–f**) Portion of MALDI-TOF mass spectrometry spectra showing the ^{15}N-labeled and non-labeled populations of the peptide digested, by either trypsin or endoproteinase Asp-N, from the Neh2 domain expressed without suppression (**d**), with inhibitors (**e**), or with both excess amino acids and inhibitors (**f**) as described in (**a–c**) and Table 2. Theoretical molecular mass of the peptide without ^{15}N incorporation is indicated below the peptide sequence

leucines, and valines, they were less significant when compared with adding inhibitors alone. In addition, by adding a combination of excess non-labeled amino acids and enzyme inhibitors, the residual undesired cross peak signals were corrected to 9% or less relative to ^{15}N-glutamine (Fig. 3c; Table 2). In particular, there was reduction in relative intensity of 77% and 89% for non-targeted ^{15}N-glutamate and ^{15}N-aspartate, respectively. This demonstrated that by manipulating the product feedback loops of *E. coli*'s amino acid metabolic pathways, the conversion from glutamine to the downstream amino acid types was controlled (Fig. 2a). For other less problematic amino acid types such as arginine, lysine, and leucine, clean labeling was obtained by only supplementing 0.1 g/l of ^{15}N-labeled amino acid of interest and 1 g/l of the rest of non-labeled amino acids 15 min prior to induction (data not shown).

One obvious concern of using an excess of non-labeled amino acids in culture would be isotopic dilution. To address the extent of this problem, samples of ^{15}N-glutamine-labeled Neh2 obtained using different conditions were digested by endoproteinase Asp-N or trypsin, and were subjected to MALDI-TOF mass spectrometry. The isotopic distribution of the fragments containing only one glutamine residue, ^{21}DILWRQ^{26} or ^{35}EVFDFSQR^{42}, was analyzed taking into consideration the degree of metabolic conversion in each NMR sample (Table 2), the known average percentage of natural abundance of carbon-13 in reagents (1.11%, West et al. 2006) and the shape of the peaks in each mass spectrometry profile (Fig. 3d–f). The level of ^{15}N enrichment of glutamine in the non-suppressed NMR sample was estimated to be approximately 58%, which increased up to 96% when cells were treated with inhibitors for glutamate/glutamine inter-conversion. However, in both cases, the degree of metabolic conversion to other amino acid types was high, especially for non-targeted labeling of glutamate, aspartate, leucine, isoleucine,

Table 2 Relative resonance intensities of ^{15}N-glutamine labeled Neh2 prepared with or without suppression[a,b]

	Amino acid types						
	Gln	Glu	Asp	Leu	Ile	Val	Phe
Without suppression[c]	1.00	0.39	0.37	0.08	0.09	0.09	0.20
Enzyme inhibitors[c,d]	1.00	0.31	0.31	0.10	0.12	0.09	0.20
Excess amino acids[c,e]	1.00	0.14	0.06	0.05	ND[f]	0.04	ND[f]
Inhibitors and excess amino acids[c,d,e]	1.00	0.09	0.04	0.02	ND[f]	0.02	0.05

[a] Resonance intensities were taken from well-resolved peaks only and normalized by the relative peak intensity of each amino-acid type relative to glutamine of uniformly ^{15}N-labeled sample

[b] Glutamine resonance intensities were set as 1.00

[c] Modified M9 minimal medium supplemented with 0.1 g/l of each non-labeled amino acid except Gln and Glu. Fifteen minutes prior to IPTG induction, 0.1 g of [$^{15}N_1$]-Gln was added to each liter of culture

[d] Seventy-five mg of 6-diazo-5-oxo-L-norleucine, 180 mg of L-methionine sulfoximine, and 180 mg of L-methionine sulfone were added to each liter of culture 15 min prior to IPTG induction

[e] One g of each non-labeled amino acid except Gln and Glu, 2 g of non-labeled Glu were added to each liter of culture 15 min prior to IPTG induction

[f] ND = Not detectable

and phenylalanine (Fig. 3a, b; Table 2). This suggests that inhibitors might have helped to reduce isotopic dilution (from glutamate to glutamine) but failed to increase specificity (from glutamine to glutamate). Conversely, there was a reduction to 36% of the ^{15}N enrichment of glutamine in the sample prepared simultaneously with excess non-labeled amino acids and enzyme inhibitors. The HSQC spectrum showed very low cross labeling (Fig. 3c; Table 2), and the quality of the signals of this HSQC spectrum was sufficiently clean and strong for unambiguous backbone assignment of glutamine signals. Taken together, although isotopic dilution due to the basal glutamate/glutamine inter-conversion could not be avoided, the quality of ^{15}N-glutamine labeling allows unambiguous backbone assignment without restriction to the use of auxotrophic mutants or CF system. It is worth noting that without any difficulty we were able to distinguish three pairs of QK dipeptides within the Neh2 sequence by acquiring a HNCO spectrum on the protein sample that was labeled with both ^{15}N/^{13}C-glutamine and ^{15}N-lysine (data not shown).

Glutamate is another problematic amino acid type when labeling in vivo as described earlier. Therefore, a significant amount of cross labeling was seen even after adapting the similar procedure used for labeling glutamine (Fig. 4a). For example, about 45% relative resonance intensity of ^{15}N-aspartate signal was observed when the medium was supplemented by 1 g/l of non-labeled amino acids, 0.1 g/l of ^{15}N-glutamate and enzyme inhibitors (Table 3, Condition 1). The relative ^{15}N-aspartate signal dropped to less than 22% after doubling some of the non-labeled amino acids including aspartate (Table 3, Condition 2). However, further increase of non-labeled aspartate did not suffice for further decrease of glutamate–aspartate conversion (data not

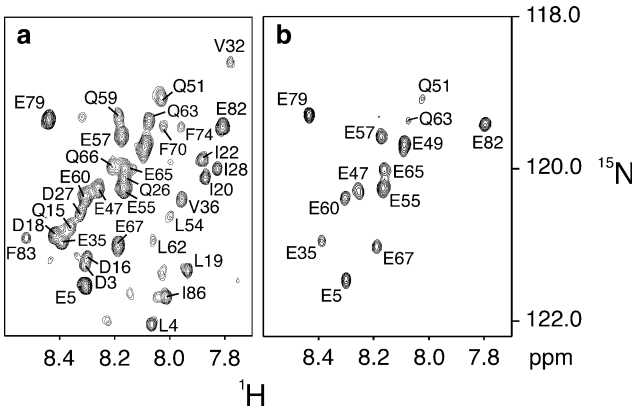

Fig. 4 The overlapping region of the ^{1}H-^{15}N HSQC spectra of ^{15}N-glutamate labeled Neh2 expressed using condition 1 (**a**) or condition 3 (**b**) as described in Table 3. These spectra were recorded and plotted identically using the same protein concentration

Table 3 Relative resonance intensities of ^{15}N-glutamate labeled Neh2 prepared with various concentrations of non-labeled amino acids or inhibitors[a,b]

	Amino acid types						
	Glu	Gln	Asp	Leu	Ile	Val	Phe
Condition 1[c,d]	1.00	0.62	0.45	0.24	0.28	0.16	0.22
Condition 2[c,e]	1.00	0.39	0.22	0.07	0.09	0.07	0.10
Condition 3[c,f]	1.00	0.49	0.07	0.03	ND[g]	ND[g]	0.07

[a] Resonance intensities were taken from well-resolved peaks only and normalized by the relative peak intensity of each amino-acid type relative to glutamate of uniformly ^{15}N-labeled sample

[b] Glutamate resonance intensities were set as 1.00

[c] Modified M9 minimal medium supplemented with 0.1 g/l of each non-labeled amino acid except Glu and Gln. Fifteen minutes prior to IPTG induction, 0.1 g of [^{15}N]-Glu was added to each liter of culture

[d] One g of each non-labeled amino acid except Glu, 88 mg of 6-diazo-5-oxo-L-norleucine, 235 mg of L-methionine sulfoximine, and 235 mg of L-methionine sulfone were added to each liter of culture 15 min prior to IPTG induction

[e] Non-labeled amino acid mixture [2 g of Asp, Gln, Leu, Val, Phe, Ala, Tyr, and Ile; 1 g of Asn, Met, Ser, Lys, Trp, Cys, His, Thr, Pro, Gly, and Arg], 360 mg of L-methionine sulfoximine, and 380 mg of L-methionine sulfone were added to each liter of culture 15 min prior to IPTG induction

[f] Non-labeled amino acid mixture [4 g of Asp, Gln, Tyr, and Ile; 3 g of Leu, Val, Phe, Ala, and Asn; 1 g of Met, Ser, Lys, Trp, Cys, His, Thr, Pro, Gly, and Arg], 1 g of L-methionine sulfoximine, 1 g of L-methionine sulfone, 250 mg of disodium succinate, 250 mg of disodium maleate, and 250 mg of amino-oxyacetate were added to each liter of culture 15 min prior to IPTG induction

[g] ND = Not detectable

shown). The residual 22% was corrected to less than 7% after a combination of excess non-labeled aspartate and inhibitors to aspartate transaminase, such as succinate, maleate, and amino-oxyacetate (Michuda and Martinez-Carrion 1970; John et al. 1978) was applied (Table 3, Condition 3). Scrambling to other amino acid types such as leucine, isoleucine, valine, and phenylalanine was also brought down to approximately 7% or below. On the other hand, we didn't observe clean reduction of relative ^{15}N-glutamine intensity even when the amount of non-labeled glutamine was elevated to 2 g or above (Table 3). However, these data demonstrated that feedback inhibition has an impact in controlling undesired cross-peaks when labeling glutamate specifically in bacteria (Fig. 4a, b; Table 3).

There is another merit of using an excess of non-labeled amino acids. *E. coli* cells grew to higher optical density than usual, which resulted in a better protein yield. The final yield of the specific amino acid labeled Neh2 was similar when compared with that of the uniformly ^{15}N-labeled protein prepared using typical M9 medium. Also, during the process of fine-tuning the bacterial growth conditions, we realized that addition of enzyme inhibitors led to a cutback by about 70% in protein yield (data not shown), but concurrent supplementation of extra non-labeled amino acids has corrected this adversity (Table 1). A reduction of 12.5 to 25% of final protein mass observed in ^{15}N-glutamine and ^{15}N-glutamate labeling may also be accounted for by the shorter induction time of 1.5 h rather than 4 h. The shorter induction time was adopted for expressing selectively amino acid labeled protein in order to prevent metabolic scrambling and the appearance of undesired cross-peaks.

In conclusion, an alternative in vivo expression procedure for selective ^{15}N-glutamine or ^{15}N-glutamate labeling of proteins that may not be efficiently produced in commercial cell-free protein synthesis systems or generated at low yield in available transaminase-deficient mutant bacteria was described. This manipulation only requires addition of an excess of non-labeled amino acids to the medium for inducing product feedback inhibitions and if needed enzyme inhibitors for the cases like glutamate and glutamine. Therefore, it allows diversified choice of any commercially available *E. coli* cell strains that are most suitable for optimizing protein yield, such as those strains bearing extra rare codon tRNA genes or mutations in genes of thioredoxin reductase and glutathione reductase to enhance disulfide bond formation in the cytoplasm and protein folding.

Acknowledgments We would like to thank Dr. Stefan Bagby of the University of Bath for valuable discussion and critical reading of the manuscript. We would also like to thank Dr. Yasutake Katoh for the use of the plasmid, Drs. Yu Wang and Priscilla T. Y. Leung of the Genome Research Centre in the University of Hong Kong for mass spectrometry analysis, and the Colorado State University Macromolecular Resources Facility for the help in the initial mass spectrometry analysis. This work was supported in part by grants-in-aid from JST-ERATO and the Ministry of Education, Culture, Sports, Science, and Technology. T.T. was supported by the 21st Century COE Program and the Protein 3000 project.

References

Britten RJ, McClure FT (1962) The amino acid pool in *Escherichia coli*. Bacteriol Rev 26:292–335

Chumpolkulwong N, Sakamoto K, Hayashi A, Iraha F, Shinya N, Matsuda N, Kiga D, Urushibata A, Shirouzu M, Oki K, Kigawa T, Yokoyama S (2006) Translation of 'rare' codons in a cell-free protein synthesis system from *Escherichia coli*. J Struct Funct Genomics 7:31–36

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 6:277–293

Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 6:197–208

Endo Y, Sawasaki T (2006) Cell-free expression systems for eukaryotic protein production. Curr Opin Biotechnol 17:373–380

Fiaux J, Bertelsen EB, Horwich AL, Wüthrich K (2004) Uniform and residue-specific ^{15}N-labeling of proteins on a highly deuterated background. J Biomol NMR 29:289–297

Garrett DS, Powers R, Gronenborn AM, Clore GM (1991) A common sense approach to peak picking in two-, three-, and four-dimensional spectra using automatic computer analysis of contour diagrams. J Magn Reson 95:214–220

Itoh K, Wakabayashi N, Katoh Y, Ishii T, Igarashi K, Engel JD, Yamamoto M (1999) Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain. Genes Dev 13:76–86

John RA, Charteris A, Fowler LJ (1978) The reaction of amino-oxyacetate with pyridoxal phosphate-dependent enzymes. Biochem J 171:771–779

Kainosho M, Tsuji T (1982) Assignment of the three methionyl carbonyl carbon resonances in *Streptomyces* subtilisin inhibitor by a carbon-13 and nitrogen-15 double-labeling technique. A new strategy for structural studies of proteins in solution. Biochemistry 21:6273–6279

Kainosho M, Torizawa T, Iwashita Y, Terauchi T, Ono AM, Güntert P (2006) Optimal isotope labelling for NMR protein structure determinations. Nature 440:52–57

Kane JF (1995) Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. Curr Opin Biotechnol 6:494–500

Katzen F, Chang G, Kudlicki W (2005) The past, present and future of cell-free protein synthesis. Trends Biotechnol 23:150–156

Kay LE, Keifer P, Saarinen T (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. J Am Chem Soc 114:10663–10665

Keseler IM, Collado-Vides J, Gama-Castro S, Ingraham J, Paley S, Paulsen IT, Peralta-Gil M, Karp PD (2005) EcoCyc: a comprehensive database resource for *Escherichia coli*. Nucleic Acids Res 33:D334–D337

Kigawa T, Muto Y, Yokoyama S (1995) Cell-free synthesis and amino acid-selective stable isotope labeling of proteins for NMR analysis. J Biomol NMR 6:129–134

Lee KM, Androphy EJ, Baleja JD (1995) A novel method for selective isotope labeling of bacterially expressed proteins. J Biomol NMR 5:93–96

Liu A, Riek R, Wider G, von Schroetter C, Zahn R, Wüthrich K (2000) NMR experiments for resonance assignments of ^{13}C, ^{15}N

doubly-labeled flexible polypeptides: application to the human prion protein hPrP(23–230). J Biomol NMR 16:127–138

Michuda CM, Martinez-Carrion M (1970) The isozymes of glutamate-aspartate transaminase. Mechanism of inhibition of dicarboxylic acids. J Biol Chem 245:262–269

Miller RE, Stadtman ER (1972) Glutamate Synthase from *Escherichia coli.* An iron-sulfide flavoprotein. J Biol Chem 247: 7407–7419

Morita EH, Shimizu M, Ogasawara T, Endo Y, Tanaka R, Kohno T (2004) A novel way of amino acid-specific assignment in ^{1}H-^{15}N HSQC spectra with a wheat germ cell-free protein synthesis system. J Biomol NMR 30:37–45

Muchmore DC, McIntosh LP, Russell CB, Anderson DE, Dahlquist FW (1989) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. Methods Enzymol 177:44–73

Muhandiram DR, Kay LE (1994) Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. J Magn Reson 103B:203–216

Parker MJ, Aulton-Jones M, Hounslow AM, Craven CJ (2004) A combinatorial selective labeling method for the assignment of backbone amide NMR resonances. J Am Chem Soc 126: 5020–5021

Prusiner S, Stadtman ER (1976) Regulation of glutaminase B in *Escherichia coli.* II. Modulation of activity by carboxylate and borate ions. J Biol Chem 251:3457–3462

Rendina AR, Orme-Johnson WH (1978) Glutamate synthase: on the kinetic mechanism of the enzyme from *Escherichia coli* W. Biochemistry 17:5388–5393

Senn H, Eugster A, Otting G, Suter F, Wüthrich K (1987) ^{15}N-labeled P22 c2 repressor for nuclear magnetic resonance studies of protein-DNA interactions. Eur Biophys J 14:301–306

Staunton D, Schlinkert R, Zanetti G, Colebrook SA, Campbell ID (2006) Cell-free expression and selective isotope labelling in protein NMR. Magn Reson Chem 44:S2–S9

Takeuchi K, Ng E, Malia TJ, Wagner G (2007) 1-^{13}C amino acid selective labeling in a $^{2}H^{15}N$ background for NMR studies of large proteins. J Biomol NMR 38:89–98

Tate S, Tate NU, Ravera MW, Jaye M, Inagaki F (1992) A novel ^{15}N-labeling method to selectively observe $^{15}NH_2$ resonances of proteins in ^{1}H-detected heteronuclear correlation spectroscopy. FEBS Lett 297:39–42

Tong KI, Katoh Y, Kusunoki H, Itoh K, Tanaka T, Yamamoto M (2006a) Keap1 recruits Neh2 through binding to ETGE and DLG motifs: characterization of the two-site molecular recognition model. Mol Cell Biol 26:2887–2900

Tong KI, Kobayashi A, Katsuoka F, Yamamoto M (2006b) Two-site substrate recognition model for the Keap1-Nrf2 system: a hinge and latch mechanism. Biol Chem 387:1311–1320

Tyler RC, Aceti DJ, Bingman CA, Cornilescu CC, Fox BG, Frederick RO, Jeon WB, Lee MS, Newman CS, Peterson FC, Phillips GN Jr, Shahan MN, Singh S, Song J, Sreenath HK, Tyler EM, Ulrich EL, Vinarov DA, Vojtik FC, Volkman BF, Wrobel RL, Zhao Q, Markley JL (2005) Comparison of cell-based and cell-free protocols for producing target proteins from the *Arabidopsis thaliana* genome for structural studies. Proteins 59:633–643

Vinarov DA, Markley JL (2005) High-throughput automated platform for nuclear magnetic resonance-based structural proteomics. Expert Rev Proteomics 2:49–55

Wada K, Wada Y, Ishibashi F, Gojobori T, Ikemura T (1992) Codon usage tabulated from the GenBank genetic sequence data. Nucleic Acids Res 20:2111–2118

Waugh DS (1996) Genetic tools for selective labeling of proteins with α-^{15}N-amino acids. J Biomol NMR 8:184–192

Weisbrod RE, Meister A (1973) Studies on glutamine synthetase from *Escherichia coli.* Formation of pyrrolidone carboxylate and inhibition by methionine sulfoximine. J Biol Chem 248:3997–4002

West JB, Bowen GJ, Cerling TE, Ehleringer JR (2006) Stable isotopes as one of nature's ecological recorders. Trends Ecol Evol 21:408–414

Yabuki T, Kigawa T, Dohmae N, Takio K, Terada T, Ito Y, Laue ED, Cooper JA, Kainosho M, Yokoyama S (1998) Dual amino acid-selective and site-directed stable-isotope labeling of the human c-Ha-Ras protein by cell-free synthesis. J Biomol NMR 11:295–306