

$1\text{-}^{13}\text{C}$ amino acid selective labeling in a $^2\text{H}^{15}\text{N}$ background for NMR studies of large proteins

Koh Takeuchi · Elise Ng · Thomas J. Malia ·
Gerhard Wagner

Received: 29 November 2006 / Accepted: 20 February 2007 / Published online: 28 March 2007
© Springer Science+Business Media B.V. 2007

Abstract Isotope labeling by residue type (LBRT) has long been an important tool for resonance assignments at the limit where other approaches, such as triple-resonance experiments or NOESY methods do not succeed in yielding complete assignments. While LBRT has become less important for small proteins it can be the method of last resort for completing assignments of the most challenging protein systems. Here we present an approach where LBRT is achieved by adding protonated ^{14}N amino acids that are ^{13}C labeled at the carbonyl position to a medium for uniform deuteration and ^{15}N labeling. This has three important benefits over conventional ^{15}N LBRT in a deuterated background: (1) selective TROSY-HNCO cross peaks can be observed with high sensitivity for amino-acid pairs connected by the labeling, and the amide proton of the residue following the ^{13}C labeled amino acid is very sharp since its alpha position is deuterated, (2) the ^{13}C label at the carbonyl position is less prone to scrambling than the ^{15}N at the α -amino position, and (3) the peaks for the $1\text{-}^{13}\text{C}$ labeled amino acids can be identified easily from the large intensity reduction in the $^1\text{H}\text{-}^{15}\text{N}$ TROSY-HSQC spectrum for some residues that do not significantly scramble nitrogens, such as alanine and tyrosine. This approach is cost effective and has been successfully applied to proteins larger than 40 kDa.

Keywords Amino acid selective labeling · Carbonyl carbon · Large protein · Nuclear magnetic resonance (NMR) · Resonance assignment

Introduction

NMR resonance assignments of proteins as required for structure determination and interaction studies are typically pursued with triple-resonance experiments, which are often sufficient for assigning small proteins below 20 kDa. Frequently, however, isotope labeling by residue type (LBRT) is utilized to resolve ambiguities, to locate segments of linked spin systems in the primary structure, or to assign residues in segments that do not exhibit cross peaks in triple-resonance experiments due to unfavorable relaxation properties or other effects (McIntosh and Dahlquist 1990). Transverse relaxation rates can be dramatically reduced by perdeuteration (LeMaster 1989), which has enabled efficient application of triple resonance experiments to larger proteins and is crucial for the success of the TROSY strategy. As a consequence, systems in the molecular weight range up to 100 kDa and beyond have been successfully studied (Fiaux et al. 2002; Horst et al. 2005). Deuteration methods have reduced the need of LBRT for assigning smaller proteins but knowledge of the amino acid type of unassigned HSQC correlation peaks would be invaluable for assigning very large proteins at the upper size limit. However, bacterial expression of proteins in $^2\text{H}_2\text{O}$ or on deuterated media has made amino acid-specific labeling challenging and expensive. This has only been accomplished in a few cases through the incorporation of protonated or deuterated ^{15}N labeled amino acids into deuterated proteins (Metzler et al. 1996; Kelly et al. 1999; Fiaux et al. 2004). However, incorporation of protonated

Electronic Supplementary Material The online version of this article (doi:10.1007/s10858-007-9152-z) contains supplementary material, which is available to authorized users.

K. Takeuchi · E. Ng · T. J. Malia · G. Wagner (✉)
Department of Biochemistry and Molecular Pharmacology,
Harvard Medical School, 240 Longwood Avenue, Boston, MA
02115, USA
e-mail: gerhard_wagner@hms.harvard.edu

^{15}N labeled amino acids has the disadvantage of line broadening, due to dipole interaction between the amide-proton resonance of the labeled residue and its own α proton. Incorporation of $^2\text{H}^{15}\text{N}$ amino acids is usually quite expensive. Thus, the use of LBRT has not been easily available for assigning spectra of large proteins that crucially depend on the line-narrowing effect of deuteration.

Here we describe a simple approach for efficient LBRT that can be pursued for expression in a deuterated background. By adding a $1\text{-}^{13}\text{C}$ labeled protonated amino acid to [^{15}N , ^{12}C , ^2H] labeled expression media, a sample is produced where: (1) the $^1\text{H}\text{-}^{15}\text{N}$ cross peak of the subsequent residue can be detected from a TROSY HNC O experiment with narrow line shape due to the deuteration of the α -proton, (2) the $^1\text{H}\text{-}^{15}\text{N}$ cross peak of the labeled amino acid type is strongly reduced. The ^{13}C label at the carbonyl position is less prone to scrambling than the ^{15}N at the α -amino position. The labeling patterns were determined on a low molecular weight protein, the 1st SH3 domain of the multi-domain signaling protein Nck. The applicability to large proteins was then tested successfully on the catalytic domain of calcineurin, a 40 kDa paramagnetic protein and VDAC1, a 30 kDa outer mitochondrial membrane protein immersed in a detergent micelle.

Carbonyl labeling strategy

We describe an information-rich and cost-effective strategy for LBRT in a deuterated background relying on the use of carbonyl ^{13}C ($1\text{-}^{13}\text{C}$)-labeled amino acids. In this scheme, an *E. coli* strain harboring a protein expression vector was grown in D_2O (or H_2O) on M9 media supplemented with ^2H glucose, ^{15}N HN_4Cl and $^2\text{H}^{15}\text{N}$ Celtone[®] base media, which is an algal hydrolysate, containing all $^2\text{H}^{15}\text{N}$ amino acids except for asparagine and glutamine. $1\text{-}^{13}\text{C}$ labeled amino acids were then added at 4–9 times the amount present in the Celtone media. With this approach, the residue type targeted with this labeling procedure is expected to be 80–90% positively (^{13}C) labeled for the carbonyls as well as negatively (^{14}N) labeled for the amides of a specific amino acid. Thus, from a simple two-dimensional version of the most sensitive triple resonance experiment TROSY-HN(CO), the amino acid succeeding the labeled amino acid is easily identified. In addition, information for the assignment of the $^1\text{H}\text{-}^{15}\text{N}$ cross peak of the $1\text{-}^{13}\text{C}$ labeled amino acid itself can be obtained from the strong intensity reductions in the ^{15}N TROSY HSQC spectra as compared to those of the uniformly ^2H and ^{15}N labeled protein. This essentially doubles the amount of information than that can be obtained over the common strategy of using ^{15}N labeled amino acids. This approach avoids the problem of fast amide relaxation due to the α proton as observed when

introducing non-deuterated amino acids since the α protons adjacent to the observed resonances are highly deuterated. Furthermore, relaxation of carbonyl carbons is nearly insensitive to proton incorporation. The strategy is also cost efficient when compared with the previously described method that utilizes $^2\text{H}^{15}\text{N}$ labeled amino acids (Supplemental figure) (Fiaux et al. 2004). Except for arginine, histidine, lysine and threonine, the cost is significantly lower for $1\text{-}^{13}\text{C}$ than for $^{15}\text{N}\text{-}^2\text{H}$ labeling. In addition, $1\text{-}^{13}\text{C}$ labeling is much more selective than amide labeling for some amino acids, since the carbonyl moiety cannot be scrambled by transaminases. While ^{15}N labeling is widely used compared to $1\text{-}^{13}\text{C}$ labeling, the strategy is akin to early implementation of $1\text{-}^{13}\text{C}$ -labeling strategies for site specific assignment of protein (Kainosho and Tsuji 1982; Kato et al. 1991).

Recent development of *in vitro* protein expression systems has provided an additional strategy for expressing protein (Kigawa et al. 1999; Sawasaki et al. 2002; Kigawa et al. 2004; Morita et al. 2004; Torizawa et al. 2004; Vinarov et al. 2004; Koglin et al. 2006). Compared to the *in vivo* system used here, an *in vitro* system might provide better selectivity due to the lack of metabolic enzymes that introduce scrambling (Yabuki et al. 1998; Morita et al. 2004; Ozawa et al. 2004; Staunton et al. 2006). However, commercial reactions for cell free systems are still very expensive compared to an *in vivo* expression system using *E. coli* (Tyler et al. 2005; Staunton et al. 2006). Although the cost of *in vitro* expression can be reduced by using in house reagents, this usually requires significant effort to optimize expression systems to produce sufficient amounts of protein for NMR analysis. Therefore, *in vivo* LBRT remains more practical than *in vitro* labeling in cell-free systems, at least for the very near future.

Materials and methods

All chemicals were purchased from Sigma unless otherwise noted. $1\text{-}^{13}\text{C}$ amino acids were purchased from Cambridge Isotope laboratories. Celtone[®] base media were purchased from Spectra Stable Isotopes. An NFAT peptide that contains the calcineurin binding sequence $\text{H}_4\text{N}\text{-GPHPVI-VITGPHEE-COOH}$ was chemically synthesized by the Tufts-New England Medical Center Peptide Synthesis Facility, Boston, MA, USA.

Expression and purification of $1\text{-}^{13}\text{C}$ selectively labeled 1st SH3 domain of Nck (Nck SH3.1) in uniformly ^{15}N -labeled background

$1\text{-}^{13}\text{C}$ selective labeling in a uniformly ^{15}N -labeled background was applied to Nck SH3.1. Nck SH3.1 is a

67 amino acid protein, consisting of 18 amino acid types (no phenylalanine and cysteine). The gene for the Nck SH3.1 was cloned in to the pET30 vector from Novagen as described in previous literature (Park et al. 2006). Nck SH3.1 was expressed in RosettaTM (DE3) *E. coli* cells (Novagen) at 37°C and protein expression was induced for 24 h. Of the 18 ¹⁻¹³C amino acid types observed in NMR spectra of Nck SH3.1, 11 amino acids that are less expensive than ²H¹⁵N amino acids, were tried (i.e. alanine, aspartate, glutamate, glutamine, glycine, isoleucine, leucine, proline, serine, tyrosine, and valine). Tryptophan was not tried because the ¹⁻¹³C labeled form is expensive (Supplemental figure). For each of the ¹⁻¹³C labeled samples, the cells were cultured in ¹⁵N M9 Celtone media containing 8.5 g/L Na₂HPO₄, 3 g/L KH₂PO₄, 0.5 g/L NaCl, 2mM MgCl₂, 0.1 mM CaCl₂, 1.5 g/L of glucose, 1 g/L of ¹⁵NH₄Cl, and 1 g/L ¹⁵N Celtone[®] base powder. Immediately after induction, the ¹⁻¹³C amino acids are added to give a fourfold excess of the amount present in typical Celtone base media as reported by the supplier (212 mg/L, 268mg/L, 286 mg/L, 50 mg/L, 178 mg/L, 87 mg/L, 233 mg/L, 156 mg/L, 123 mg/L, 106 mg/L, or 125 mg/L of ¹⁻¹³C labeled alanine, aspartate, glutamate, glutamine, glycine, isoleucine, leucine, proline, serine, tyrosine, or valine, respectively, were added). Then, at 16 h after induction, the same amount of the respective amino acids was added again to each of the ¹⁻¹³C labeled samples to replenish the stock of labeled amino acid and diminish the effect of scrambling. The cells were incubated for another 8 h for a total of 24 h induction. Uniformly ¹⁵N labeled sample was also expressed in the same way without adding extra amino acids. The protein was purified with Ni-NTA affinity chromatography as previously described (Park et al. 2006). Nck SH3.1 is expressed as a mixture of monomer and dimer. To obtain the monomeric form the protein was heated to 95°C for 5 min in diluted solution. The procedure converts the dimer into a monomer, and the monomer is separated by gel filtration on a Superdex-75 column (Amersham bioscience, Piscataway, NJ, USA) and then concentrated again. The spectra obtained from the procedure were identical with previously reported spectra of Nck SH3.1 (Park et al. 2006).

Expression and purification of ¹⁻¹³C selectively labeled catalytic domain of calcineurin in a uniformly ²H¹⁵N-labeled background

¹⁻¹³C selective labeling in a uniformly ²H¹⁵N-labeled background was applied to the catalytic domain of the paramagnetic protein, calcineurin, which consists of 347 residues and has a molecular weight of 40 kDa. The catalytic domain of human calcineurin A α (CnCat), comprising residues 2–347 with substitutions Y341S, L343A,

and M347D was cloned in to the pGEX vector from Amersham Biosciences as described previously (Aramburu et al. 1999) The catalytic domain of calcineurin was expressed in RosettaTM (DE3) *E. coli* cells (Novagen) at 37°C and protein expression was induced for 24 h. Leucine and valine selective labeling were tried. For each of the two ¹⁻¹³C labeled samples, the cells were cultured in modified H₂O ²H¹⁵N M9 Celtone media containing 8.5 g/L Na₂HPO₄, 3 g/L KH₂PO₄, 0.5 g/L NaCl, 2 mM MgCl₂, 0.1 mM CaCl₂, 10 mg/ml FeCl₃, 10 mg/L ZnSO₄, 1.5 g/L of ²H glucose, 1 g/L of ¹⁵NH₄Cl, and 1 g/L ²H¹⁵N Celtone[®] base powder. Immediately after induction, 200 mg/L and 210 mg/L of ¹⁻¹³C labeled leucine or valine were added, respectively. Then, at 12 h after induction, the same amount of the respective amino acids was added again to each of the ¹⁻¹³C labeled samples. The cells were incubated for another 12 h for a total of 24 h. Following incubation, the cells were harvested and the protein was purified as previously described (Roehrl et al. 2004).

Expression and purification of ¹⁻¹³C alanine selectively labeled human voltage-dependent anion channel 1 (VDAC1) in a uniformly ²H¹⁵N-labeled background

Carbonyl-¹³C alanine selectively labeled VDAC1 was expressed and purified as described in (Malia and Wagner 2007) with minor modifications. Briefly, the gene for the VDAC1 was cloned in to the pET21 vector from Novagen. VDAC1 with a C-terminal His-tag was expressed as inclusion bodies in BL21 (DE3) *E. coli* cells (Novagen) at 37 °C. The cells were cultured in D₂O ²H¹⁵N M9 Celtone media containing 8.5 g/L Na₂HPO₄, 3 g/L KH₂PO₄, 0.5 g/L NaCl, 2 mM MgCl₂, 0.1 mM CaCl₂, 1.5 g/L of ²H glucose, 1 g/L of ¹⁵NH₄Cl, and 1 g/L ²H¹⁵N Celtone[®] base powder. Immediately before induction, 522 mg/L of ¹⁻¹³C labeled alanine was added. The total induction time was 3 hr. Inclusion bodies were isolated and VDAC1 was purified by Ni-NTA affinity chromatography under denaturing conditions. The denatured VDAC1 was then refolded into LDAO micelles. As the final step of purification, refolded VDAC1 was subjected to gel filtration chromatography. The buffer was then exchanged to 25 mM sodium phosphate (pH 5.0), 5 mM DTT, 1 mM EDTA, and 5% D₂O for NMR experiments.

NMR experiments

NMR spectra were recorded on a Bruker Avance 500 spectrometer for Nck SH3.1 and a Bruker Avance 750 instrument for calcineurin and VDAC1. All spectra were recorded at 25°C in buffer containing 10 mM sodium phosphate (pH 6.8), 100 mM NaCl, 3mM dithiothreitol and 7% D₂O. For Nck3.1, a ¹H-¹⁵N HSQC experiment

without ^{13}C decoupling was recorded with a spectral width of 6,510 Hz for proton and 1,318 Hz for nitrogen; 512 and 256 complex data points were recorded for proton and nitrogen, respectively. A 2D HN(CO) was recorded with the same spectral width as the HSQC with 512 and 38 complexed data points for proton and nitrogen, respectively. 4 and 16 scans were applied for each increment of HSQC and HN(CO) experiments, respectively. Protein concentrations were within the range of 0.33 to 1.0 mM. For calcineurin in complex with the NFAT peptide, a TROSY HN(CO) was recorded with a spectral width of 9,766 and 2,431 Hz with 512 and 56 complex data points for proton and nitrogen, respectively. 72 scans were accumulated for each increment. The protein concentrations of the samples were 0.3 mM, and the NFAT peptide was added at 1.2 molar ratio compared to calcineurin (Aramburu et al. 1999). For VDAC1, a TROSY HN(CO) was recorded with a spectral width of 9,766 and 2,355 Hz with 512 and 42 complex data points for proton and nitrogen, respectively. 640 scans were accumulated for each increment. The protein concentrations of the samples were 0.6 mM. All spectra were processed in Xwinnmr (Bruker, Germany) or nmrPipe (Delaglio et al. 1995) and analyzed with either Sparky (Goddard and Kneller 2002) or CARA (Keller 2004).

Results and discussion

To evaluate the aforementioned approach, we first applied the scheme to Nck SH3.1, an 8.2 kDa protein domain that consists of 67 amino acids. Since this is a small test protein we used protonated media. Bacteria were grown in H_2O with ^1H glucose, and ^{15}N Celtone base media rather than D_2O , ^2H glucose and $^2\text{H}^{15}\text{N}$ Celtone. Nck SH3.1 was purified as previously described (Park et al. 2006). Protein expression was induced for 24 h, which is a significantly longer period than typically used for expression of selectively labeled proteins. The long induction period is required for protein expression in a deuterated medium because of the reduced expression rate compared to a protonated medium. For example, deuterated calcineurin can only be expressed at reasonable amounts with a 24 h induction period. Only 1/5–1/10 of the final amount is expressed after 3 h of induction, which is the typical induction time used for the expression of ^{15}N selectively labeled proteins. 2D HN(CO) and ^{13}C coupled ^1H - ^{15}N HSQC spectra were recorded and analyzed using previously obtained assignments (Park et al. 2006). In 2D HN(CO) spectra of the selectively $1\text{-}^{13}\text{C}$ labeled samples, resonances for the residues following the $1\text{-}^{13}\text{C}$ labeled amino acid are observed (Fig. 1a). In the ^1H - ^{15}N HSQC spectra of the selectively $1\text{-}^{13}\text{C}$ labeled samples,

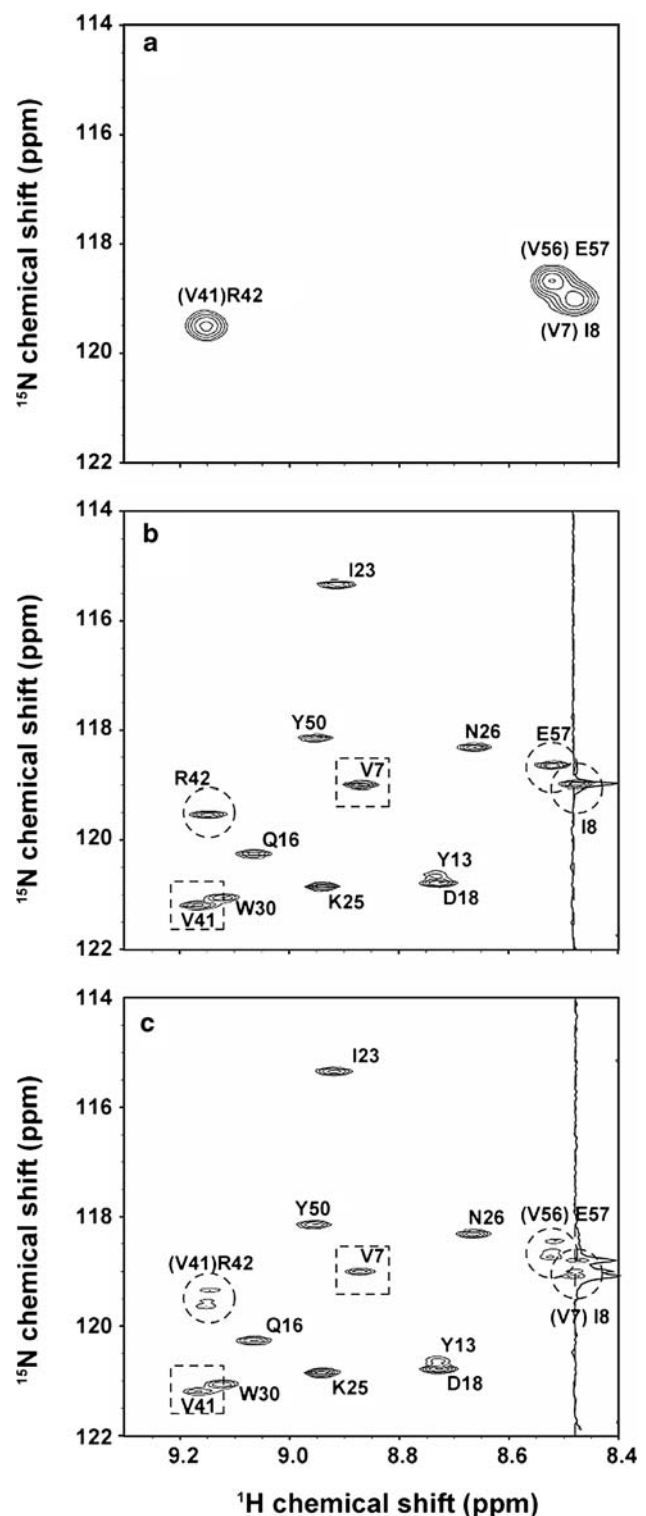


Fig. 1 (a) Section of 2D HN(CO) spectra of $1\text{-}^{13}\text{C}$ valine selectively labeled Nck SH3.1. Resonances in (a) correspond to those succeeding $1\text{-}^{13}\text{C}$ labeled amino acids (indicated in parentheses). **b** and **c** Sections of ^{13}C coupled ^1H - ^{15}N HSQC spectra of (b) uniformly ^{15}N -labeled and (c) $1\text{-}^{13}\text{C}$ valine selectively labeled Nck SH3.1. Peaks with decreased intensity (boxed) or split by $^{15}\text{N}(i)\text{-}^{13}\text{C}(i-1)$ coupling (circled) are marked. For circled resonances in (c), the preceding residues are indicated in parentheses

resonances were perturbed in two different ways as compared to those in the uniformly ^{15}N -labeled sample (Fig. 1b, c). (1) Some resonances exhibited decreased intensity (boxed) resulting from the incorporation of one type of (^{14}N , $1\text{-}^{13}\text{C}$) labeled amino acid. (2) Other resonances were split into three peaks (circled), due to $^{15}\text{N}(i)\text{-}^{13}\text{C}(i-1)$ coupling, which occurs for those resonances that have $1\text{-}^{13}\text{C}$ -labeled amino acids as their predecessor. For these resonances, the center peak represents the population of N-H groups facing ^{12}C -in the preceding residue (originating from the unlabeled amino acid in the Celtone media) and the top and bottom peaks are doublet components due to the presence ^{13}C at the carbonyl position of the preceding residue. The top and bottom peaks show E.COSY type splitting, since both amide proton and nitrogen couple with the same $1\text{-}^{13}\text{C}$ of the preceding residue. The splittings in proton dimension are ~ 4.5 Hz corresponding to the $^2J_{\text{H-C}}$ coupling, while splittings in the nitrogen dimension correspond to the $^1J_{\text{N-C}}$ coupling. The position of the doublet is slightly shifted up-field (~ 2.5 Hz) in the nitrogen dimension due to the ^{13}C isotope shift caused by the preceding carbonyl carbon.

As mentioned above, ^{13}C selective labeling allows for the identification of the residues succeeding the $1\text{-}^{13}\text{C}$ labeled amino acids by comparing HN(CO) spectra of selectively labeled sample with that of uniformly $^{15}\text{N}^{13}\text{C}$ labeled Nck SH3.1 (Fig. 2). From the HN(CO) spectra in Fig. 2, it is obvious that $1\text{-}^{13}\text{C}$ incorporation for isoleucine, leucine, tyrosine, and valine is very selective. For these amino acids, no more than 5% incorporation of ^{13}C labeling for the non-targeted amino acids is observed (Fig. 3). $1\text{-}^{13}\text{C}$ incorporation is also selective for alanine and proline, with no more than 15% incorporation of ^{13}C labeling to non-targeted amino acids (Fig. 3). The absolute ^{13}C -labeling ratio for these residues is calculated from the volume of the center peak versus split peaks in a ^{13}C coupled HSQC spectra (Fig. 1c). Alanine, isoleucine, leucine, proline, tyrosine and valine showed 53 ± 6 , 83 ± 4 , 80 ± 4 , 88 , 80 ± 2 , and $75 \pm 8\%$ incorporation of ^{13}C atoms to carbonyl positions, respectively, which are sufficiently high to provide useful information for difficult systems. Glycine, glutamate, and glutamine are scrambled to some specific amino acids, however. Glycine is metabolized to serine and tryptophan (Figs. 2f, 3e) by serine hydroxymethyltransferase conversion of glycine to serine and subsequent conversion of serine to tryptophan. Glutamate and glutamine were metabolized to arginine, glutamine/glutamate, and proline (Fig. 2d, e, and Fig. 3c, d). Glutamate is the precursor of arginine, glutamine, and proline, and glutamate and glutamine are inter-convertible by transaminases. While the selectivity of labeling is not perfect for these residues, the resultant metabolites for glycine, glutamate and glutamine are limited to a specific

subset of amino acids. Thus, $1\text{-}^{13}\text{C}$ selective labeling is useful for these amino acids. The absolute ^{13}C -labeling ratio for glycine was 68% and $<20\%$ for glutamine and glutamate. On the other hand, $1\text{-}^{13}\text{C}$ labeling of aspartate and serine did not work well, because of extensive scrambling, leading to the observation of resonances in more than half of the observed signals. The absolute ^{13}C -labeling ratio for these residues is also quite low ($<20\%$).

Since ^{14}N incorporation results in reduced resonance intensity in $^1\text{H}\text{-}^{15}\text{N}$ HSQC spectra, the resonances from targeted amino acids can be identified. This is basically the same information that can be obtained from ^{15}N selective labeling. The degree of ^{14}N incorporation for each of these resonances was calculated by comparing the intensity of the resonances in each of the ^{14}N , $1\text{-}^{13}\text{C}$ selective labeling experiments with those of the uniformly ^{15}N -labeled sample (Fig. 4). Comparison between $1\text{-}^{13}\text{C}$ (Fig. 3) and ^{14}N incorporation (Fig. 4) revealed that $1\text{-}^{13}\text{C}$ incorporation is remarkably specific compared to ^{14}N incorporation. The differences are quite obvious for isoleucine, leucine, and valine. While scrambling compromises the effectiveness of amide selective labeling for these residues, it does not occur with the ^{13}C atoms. For selective $1\text{-}^{13}\text{C}$ isoleucine labeling experiment, ^{14}N -labeling was also observed in alanine, leucine, and valine; for the selective $1\text{-}^{13}\text{C}$ leucine labeling experiment, ^{14}N -labeling was also observed in alanine, valine, and isoleucine; when providing ^{14}N -valine, ^{14}N -labeling appeared in alanine at more than 0.7 times the amount of that in valine. Instead, carbonyl ^{13}C labeling remained present solely in the specific amino acid that was added for each experiment. As expected from the *E. coli* metabolic pathway, amide selective labeling of aspartate, glutamate, glutamine, and serine was not successful and no specific incorporation is observed for these residues. This is strikingly different from the selective $1\text{-}^{13}\text{C}$ labeling for glutamate and glutamine, which retained sufficiently high specificity to allow interpretation of the data. These results show that for some residues, such as glutamine, glutamate, isoleucine, leucine, and valine utilization of $1\text{-}^{13}\text{C}$ labeled amino acids presents a clear advantage over the conventional amide labeling strategy alone by eliminating labeling ambiguity for these amino acids. This result also indicates that transaminases or other enzymes that scramble the ^{14}N and ^{15}N nitrogens render conventional ^{15}N labeling difficult to apply for certain amino acids, which is consistent with the literature (McIntosh and Dahlquist 1990). Auxotrophic strains could be used to avoid scrambling; however, growing these strains in D_2O media is often challenging. In addition, it requires maintaining a battery of autotrophic strains. Alternatively, one might want to use an *in vitro* expression system for these residues. ^{14}N -glycine was metabolized to ^{14}N -serine and ^{14}N -tryptophan as was $1\text{-}^{13}\text{C}$ labeling. As seen in Fig. 4, we found that ^{14}N incorporation for alanine was over

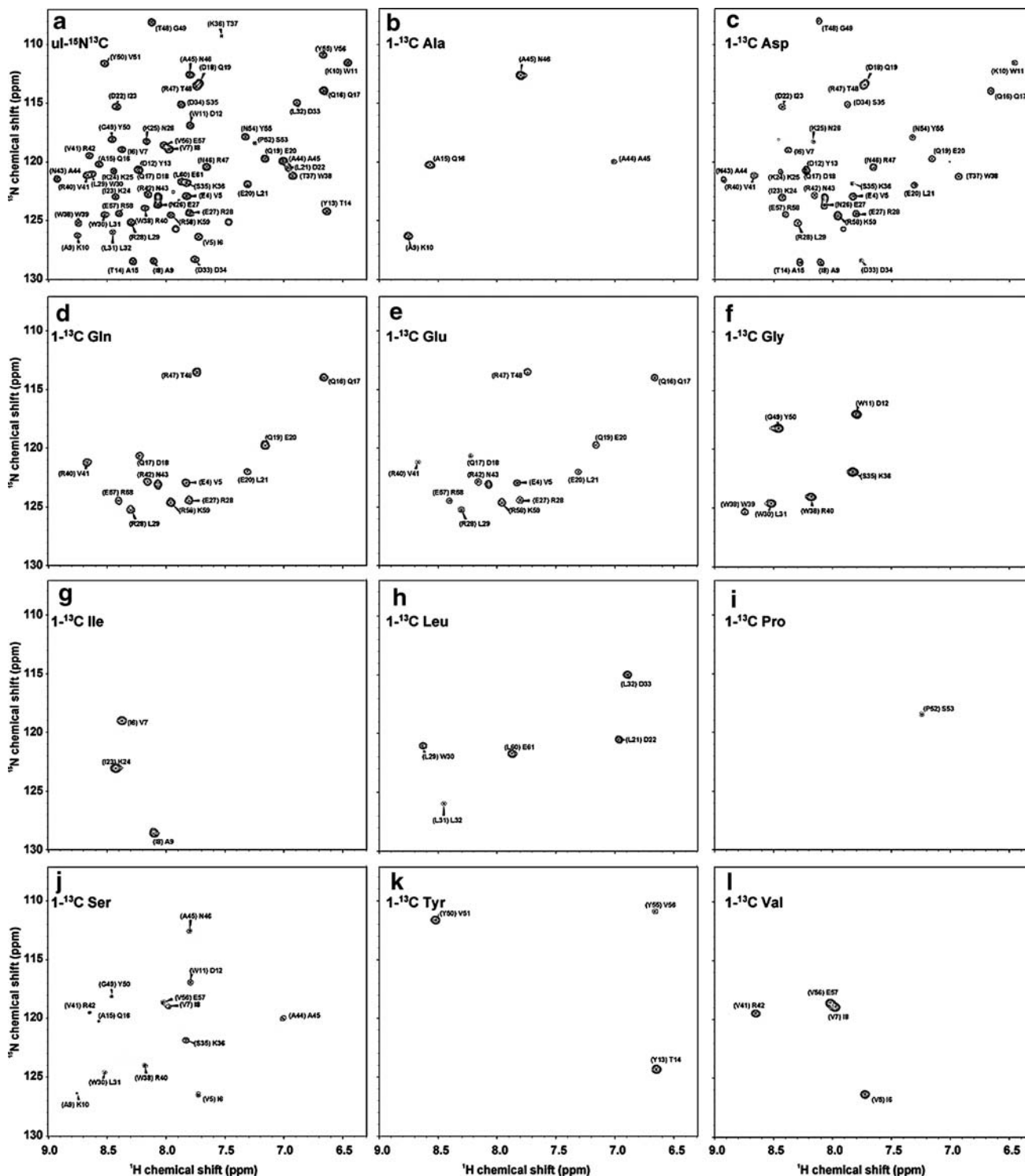


Fig. 2 2D HN(CO) spectra of **a** uniformly $^{15}\text{N}/^{13}\text{C}$ -labeled, **b** $1\text{-}^{13}\text{C}$ alanine, **c** $1\text{-}^{13}\text{C}$ aspartate, **d** $1\text{-}^{13}\text{C}$ glutamine, **e** $1\text{-}^{13}\text{C}$ glutamate, **f** $1\text{-}^{13}\text{C}$ glycine, **g** $1\text{-}^{13}\text{C}$ isoleucine, **h** $1\text{-}^{13}\text{C}$ leucine, **i** $1\text{-}^{13}\text{C}$ proline, **j**

$1\text{-}^{13}\text{C}$ serine, **k** $1\text{-}^{13}\text{C}$ tyrosine, and **l** $1\text{-}^{13}\text{C}$ valine selectively labeled Nck SH3.1. Resonances correspond to those succeeding $1\text{-}^{13}\text{C}$ labeled amino acids (indicated in *parentheses*)

70% and quite specific to alanine. ^{14}N -tyrosine labeling was also selective for the targeted amino acid. For these residues, the labeling strategy provides both the labeled amino acids

and their successors, which doubles the amount of information compared to the conventional amide labeling strategy.

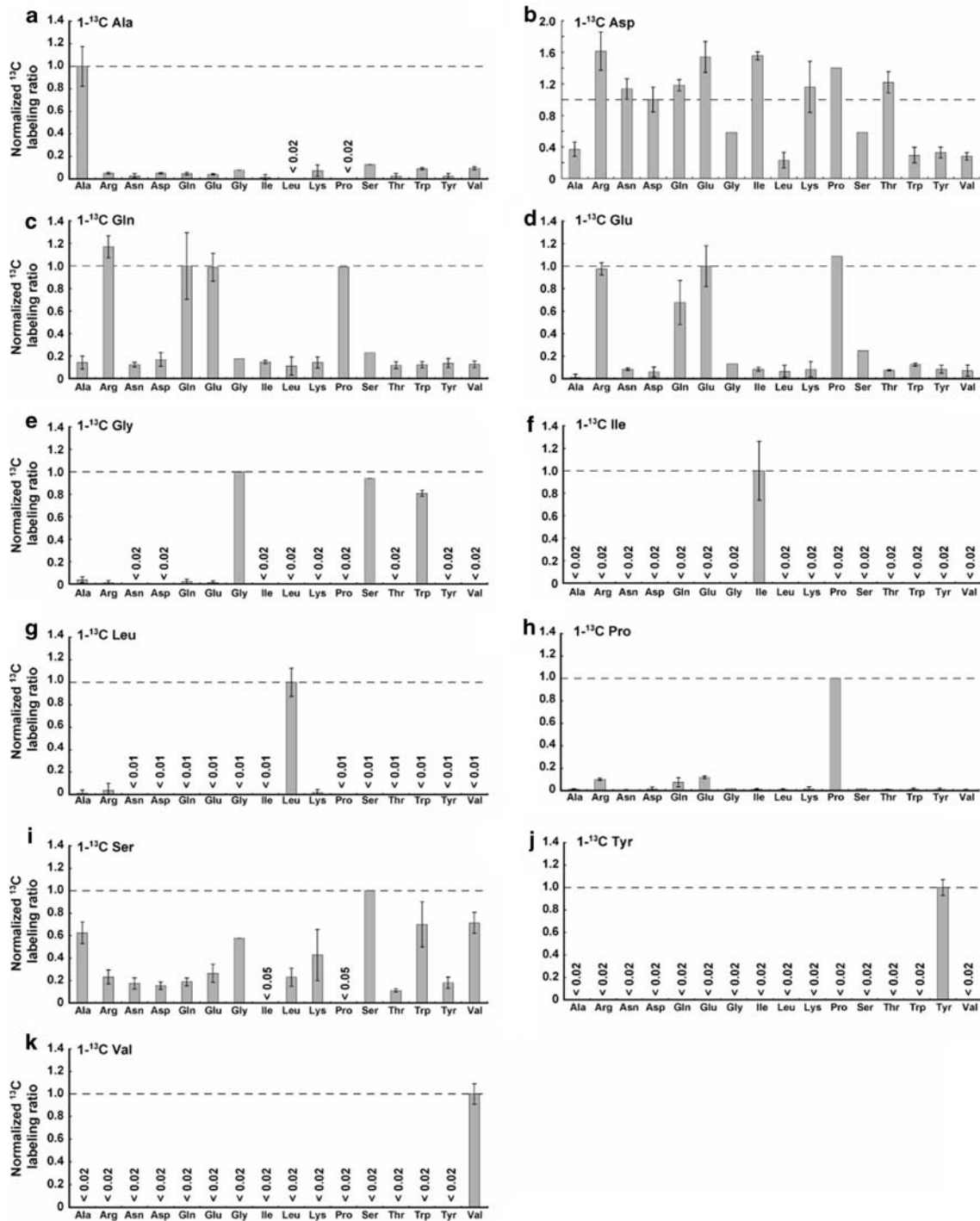


Fig. 3 Summary of ^{13}C incorporation for **a** $1\text{-}^{13}\text{C}$ alanine, **b** $1\text{-}^{13}\text{C}$ aspartate, **c** $1\text{-}^{13}\text{C}$ glutamine, **d** $1\text{-}^{13}\text{C}$ glutamate, **e** $1\text{-}^{13}\text{C}$ glycine, **f** $1\text{-}^{13}\text{C}$ isoleucine, **g** $1\text{-}^{13}\text{C}$ leucine, **h** $1\text{-}^{13}\text{C}$ proline, **i** $1\text{-}^{13}\text{C}$ serine, **j** $1\text{-}^{13}\text{C}$ tyrosine, and **k** $1\text{-}^{13}\text{C}$ valine $1\text{-}^{13}\text{C}$ selectively labeled Nck SH3.1. ^{13}C incorporation ratio relative to the ^{13}C incorporation to the

targeting residues is shown. Standard error bars were calculated using data from each amino acid type for each experiment. ^{13}C -labeling of lysine was used for VDAC1 and it was found that there is no serious scrambling, and the data could be used for assignments

The strategy described here was then applied to the 40 kDa catalytic domain of the phosphatase calcineurin, which contains a paramagnetic Fe^{3+} in its catalytic center. Since it is difficult to back exchange a significant portion of

the amide protons when it is expressed in D_2O , we used H_2O instead of D_2O while the other media components are all deuterated, such as ^2H glucose, and $^2\text{H}^{15}\text{N}$ Celtone base media. This procedure is known to deuterate a protein to a

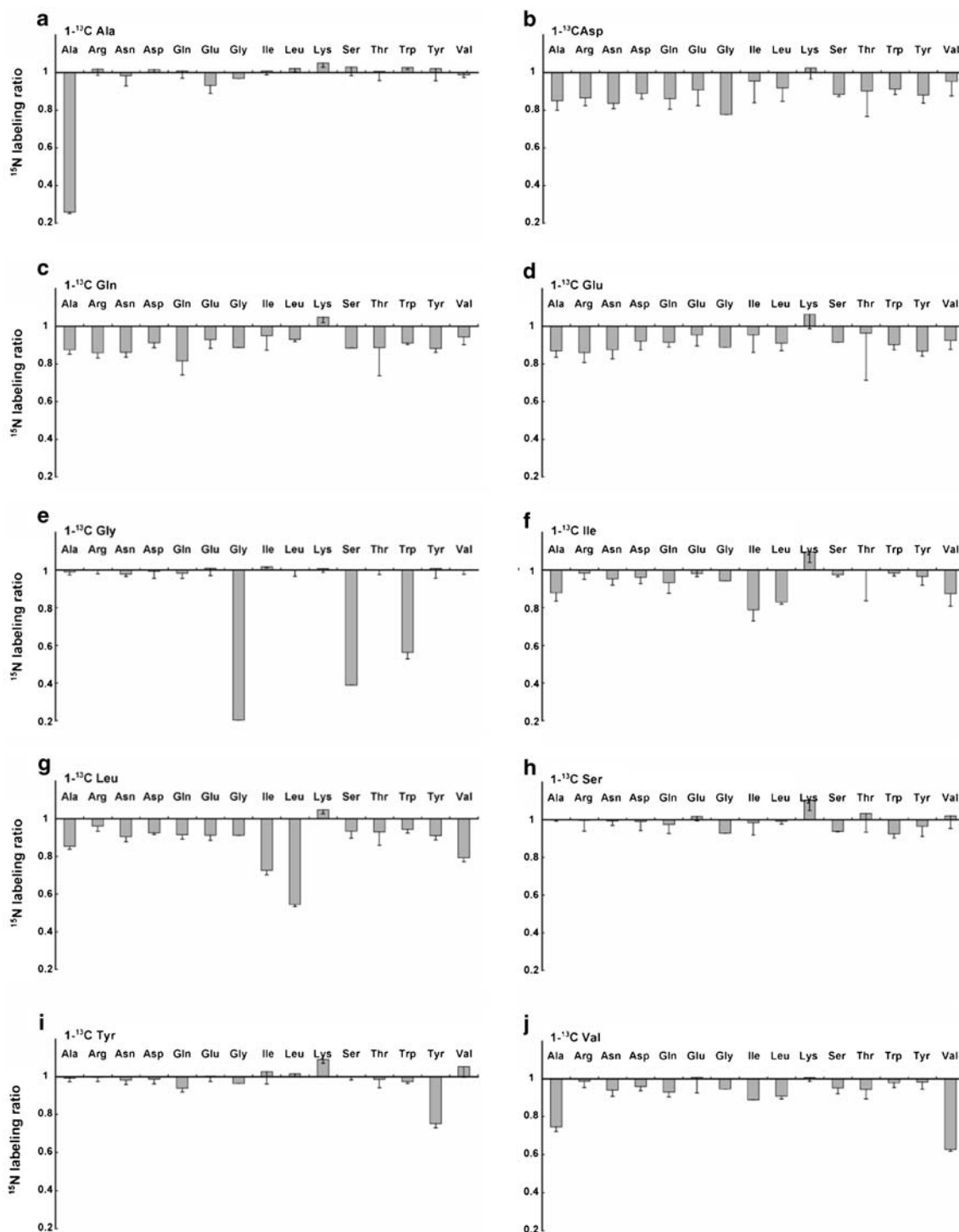


Fig. 4 Summary of ^{15}N labeling ratio for **a** $1\text{-}^{13}\text{C}$ alanine, **b** $1\text{-}^{13}\text{C}$ aspartate, **c** $1\text{-}^{13}\text{C}$ glutamine, **d** $1\text{-}^{13}\text{C}$ glutamate, **e** $1\text{-}^{13}\text{C}$ glycine, **f** $1\text{-}^{13}\text{C}$ isoleucine, **g** $1\text{-}^{13}\text{C}$ leucine, **h** $1\text{-}^{13}\text{C}$ serine, **i** $1\text{-}^{13}\text{C}$ tyrosine, and **j** $1\text{-}^{13}\text{C}$ valine $1\text{-}^{13}\text{C}$ selectively labeled Nck SH3.1. The ^{15}N -labeling

ratio was calculated by comparing the intensity of each resonance in the HSQC spectra of the selectively and the uniformly ^{15}N labeled sample. Standard error bars were calculated using data from each amino acid type for each experiment

sufficient level (Lohr et al. 2003). If limited back exchange is not a problem, or for applications to larger molecular weight systems, one can simply use D_2O instead of H_2O . Labeled calcineurin was purified as previously described

(Roehrl et al. 2004). Figure 5 shows $1\text{-}^{13}\text{C}$ Leu-or Val-labeled $^2\text{H}^{15}\text{N}$ calcineurin resonances in complex with the NFAT peptide, a binding partner of the protein. 25 of 37 and 16 of 25 resonances are clearly observed in the $1\text{-}^{13}\text{C}$

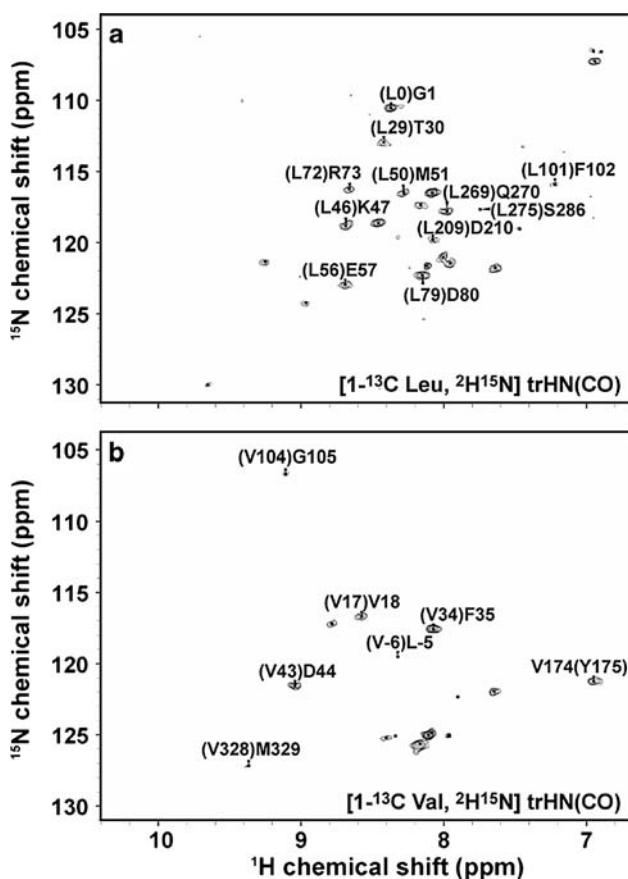


Fig. 5 2D TROSY-HN(CO) spectra of (a) $1\text{-}^{13}\text{C}$ leucine and (b) $1\text{-}^{13}\text{C}$ valine selectively labeled $^2\text{H}^{15}\text{N}$ calcineurin bound to the NFAT peptide. Assignments for the observed resonances are labeled. The assignments for preceding residues are also shown in parentheses

Leu- and Val-labeled spectra, respectively. While we lose about 30% of the resonances, due to paramagnetic broadening, the ratio of observed resonances is similar to the ratio of resonances that is observed in TROSY spectra of the protein. The method was also successfully applied to the 30 kDa mitochondrial membrane protein VDAC1, which has an effective molecular weight > 60 kDa, due to its incorporation into a detergent micelle. The protein exhibits extensive signal broadening due to apparent monomer/multimer chemical exchange (data not shown). Figure 6 shows a TROSY HN(CO) spectrum of a $1\text{-}^{13}\text{C}$ Ala, $U\text{-}^2\text{H}$, ^{15}N -VDAC1 sample. Of the 21 alanine successors, 16 could be assigned with the aid of this experiment. These results prove the ability of this labeling strategy to reduce ambiguity in the mainchain assignment of large molecular weight proteins.

In conclusion, $1\text{-}^{13}\text{C}$ selective labeling in a $^2\text{H}^{15}\text{N}$ background was successfully applied to systems larger than 40 kDa. As shown here, this method is also more informative, cost effective and specific than other *in vivo* selective labeling strategies published thus far,

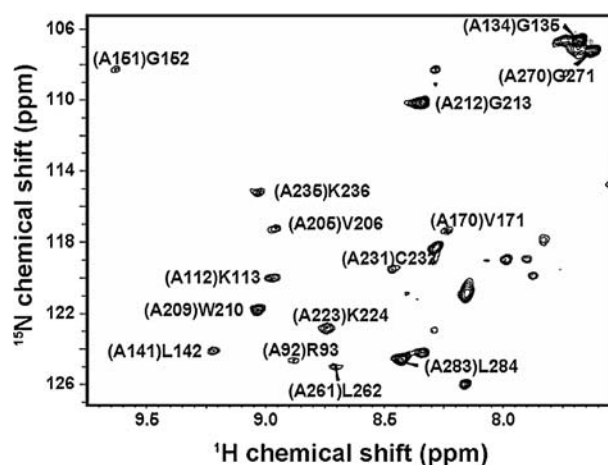


Fig. 6 2D TROSY-HN(CO) spectra of $1\text{-}^{13}\text{C}$ alanine selectively labeled $^2\text{H}^{15}\text{N}$ VDAC1. Assignments for the observed resonances are labeled. The assignments for preceding residues are also shown in parentheses

demonstrating its potential as a tool for analyzing large molecular weight proteins. Amino acid-selective labeling is especially useful for the assignment of large systems that are difficult, and for which standard triple resonance experiments do not yield extensive assignments, such as for calcineurin and micelle-bound VDAC1.

Acknowledgments This work is supported by NIH (grants GM47467 and AI37581). K.T. is supported by the Japan Society for Promotion of Science.

References

- Aramburu J, Yaffe MB, Lopez-Rodriguez C, Cantley LC, Hogan PG, Rao A (1999) Affinity-driven peptide selection of an NFAT inhibitor more selective than cyclosporin A. *Science* 285:2129–2133
- 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
- Fiaux J, Bertelsen EB, Horwich AL, Wuthrich K (2002) NMR analysis of a 900K GroEL-GroES complex. *Nature* 418:207–211
- Fiaux J, Bertelsen EB, Horwich AL, Wuthrich K (2004) Uniform and residue-specific ^{15}N -labeling of proteins on a highly deuterated background. *J Biomol NMR* 29:289–297
- Goddard TD, Kneller DG (2002) Sparky 3. University of California, San Francisco
- Horst R, Bertelsen EB, Fiaux J, Wider G, Horwich AL, Wuthrich K (2005) From the cover: direct NMR observation of a substrate protein bound to the chaperonin GroEL. *PNAS* 102:12748–12753
- 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
- Kato K, Matsunaga C, Igarashi T, Kim H, Odaka A, Shimada I, Arata Y (1991) Complete assignment of the methionyl carbonyl

- carbon resonances in switch variant anti-dansyl antibodies labeled with [1-¹³C]methionine. *Biochemistry* 30:270–278
- Keller R (2004) PhD thesis, Diss ETH Nr 15947
- Kelly MJ, Krieger C, Ball LJ, Yu Y, Richter G, Schmieder P, Bacher A, Oschkinat H (1999) Application of amino acid type-specific ¹H- and ¹⁴N-labeling in a ²H-, ¹⁵N-labeled background to a 47 kDa homodimer: potential for NMR structure determination of large proteins. *J Biomol NMR* 14:79–83
- Kigawa T, Yabuki T, Yoshida Y, Tsutsui M, Ito Y, Shibata T, Yokoyama S (1999) Cell-free production and stable-isotope labeling of milligram quantities of proteins. *FEBS Lett* 442:15–19
- Kigawa T, Yabuki T, Matsuda N, Matsuda T, Nakajima R, Tanaka A, Yokoyama S (2004) Preparation of *Escherichia coli* cell extract for highly productive cell-free protein expression. *J Struct Funct Genomics* 5:63–68
- Koglin A, Klammt C, Trbovic N, Schwarz D, Schneider B, Schafer B, Lohr F, Bernhard F, Dotsch V (2006) Combination of cell-free expression and NMR spectroscopy as a new approach for structural investigation of membrane proteins. *Magn Reson Chem* 44(Spec No): S17–S23
- LeMaster DM (1989) Deuteration in protein proton magnetic resonance. *Methods Enzymol* 177:23–43
- Lohr F, Katsemi V, Hartleib J, Gunther U, Ruterjans H (2003) A strategy to obtain backbone resonance assignments of deuterated proteins in the presence of incomplete amide ²H/¹H back-exchange. *J Biomol NMR* 25:291–311
- Malia TJ, Wagner G (2007) NMR structural investigation of the mitochondrial outer membrane protein VDAC and its interaction with antiapoptotic Bcl-x(L). *Biochemistry* 46:514–525
- McIntosh LP, Dahlquist FW (1990) Biosynthetic incorporation of ¹⁵N and ¹³C for assignment and interpretation of nuclear magnetic resonance spectra of proteins. *Q Rev Biophys* 23:1–38
- Metzler WJ, Wittekind M, Goldfarb V, Mueller L, Farmer BT (1996) Incorporation of ¹H¹³C¹⁵N-{Ile, Leu, Val} into a perdeuterated, ¹⁵N-labeled protein: potential in structure determination of large proteins by NMR. *J Am Chem Soc* 118:6800–6801
- Morita EH, Shimizu M, Ogasawara T, Endo Y, Tanaka R, Kohno T (2004) A novel way of amino acid-specific assignment in (¹H-¹⁵N) HSQC spectra with a wheat germ cell-free protein synthesis system. *J Biomol NMR* 30:37–45
- Ozawa K, Headlam MJ, Schaeffer PM, Henderson BR, Dixon NE, Otting G (2004) Optimization of an *Escherichia coli* system for cell-free synthesis of selectively N-labelled proteins for rapid analysis by NMR spectroscopy. *Eur J Biochem* 271:4084–4093
- Park S, Takeuchi K, Wagner G (2006) Solution structure of the first SRC homology 3 domain of human Nck2. *J Biomol NMR* 34:203–208
- Roehrl MH, Kang S, Aramburu J, Wagner G, Rao A, Hogan PG (2004) Selective inhibition of calcineurin-NFAT signaling by blocking protein-protein interaction with small organic molecules. *Proc Natl Acad Sci USA* 101:7554–7559
- Sawasaki T, Ogasawara T, Morishita R, Endo Y (2002) A cell-free protein synthesis system for high-throughput proteomics. *Proc Natl Acad Sci USA* 99:14652–14657
- 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(Spec No):S2–9
- Torizawa T, Shimizu M, Taoka M, Miyano H, Kainosho M (2004) Efficient production of isotopically labeled proteins by cell-free synthesis: a practical protocol. *J Biomol NMR* 30:311–325
- Tyler RC, Aceti DJ, Bingman CA, Cornilescu CC, Fox BG, Frederick RO, Jeon WB, Lee MS, Newman CS, Peterson FC et al. (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, Lytle BL, Peterson FC, Tyler EM, Volkman BF, Markley JL (2004) Cell-free protein production and labeling protocol for NMR-based structural proteomics. *Nat Methods* 1:149–153
- 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