Selective labeling of a membrane peptide with 15N-amino acids using cells grown in rich medium

ARTICLE *in* BIOPOLYMERS · JANUARY 2006

Impact Factor: 2.39 · DOI: 10.1002/bip.20546 · Source: PubMed

6 AUTHORS, INCLUDING:

Leah S Cohen
City University of New York - College of Staten I…
23 PUBLICATIONS **284** CITATIONS
SEE PROFILE

Jeffrey M Becker
University of Tennessee
290 PUBLICATIONS **6,827** CITATIONS
SEE PROFILE

Fred Naider
City University of New York - College of Staten I…
287 PUBLICATIONS **6,362** CITATIONS
SEE PROFILE

Jacqueline Englander[1]
Leah Cohen[1]
Boris Arshava[1]
Racha Estephan[1,2*]
Jeffrey M. Becker[3]
Fred Naider[1,2†]

[1] *Department of Chemistry and the Macromolecular Assemblies Institute, The College of Staten Island of the City University of New York, Staten Island, NY 10314*

[2] *Graduate School and University Center of the City University of New York, New York, NY 10016*

[3] *Department of Microbiology, University of Tennessee, Knoxville, TN 37996*

Selective Labeling of a Membrane Peptide with ^{15}N-Amino Acids Using Cells Grown in Rich Medium

Received 2 April 2006; revised 19 May 2006; accepted 29 May 2006
Published online 1 June 2006 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bip.20546

Abstract: *Nuclear magnetic resonance spectra of membrane proteins containing multiple transmembrane helices have proven difficult to resolve due to the redundancy of aliphatic and Ser/Thr residues in transmembrane domains and the low chemical shift dispersity exhibited by residues in α-helical structures. Although ^{13}C- and ^{15}N-labeling are useful tools in the biophysical analysis of proteins, selective labeling of individual amino acids has been used to help elucidate more complete structures and to probe ligand–protein interactions. In general, selective labeling has been performed in* Escherichia coli *expression systems using minimal media supplemented with a single labeled amino acid and nineteen other unlabeled amino acids and/or by using auxotrophs for specific amino acids. Growth in minimal media often results in low yields of cells or expression products. We demonstrate a method in which one labeled amino acid is added to a rich medium. These conditions resulted in high expression ($\geq$100 mg/L) of a test fusion protein and milligram quantities of the selectively labeled membrane peptide after cyanogen bromide cleavage to release the peptide from the fusion protein. High levels of ^{15}N incorporation and acceptable levels of cross-labeling into other amino acid residues of the peptide were achieved. Growth in rich media is a simple and*

Correspondence to: Fred Naider; e-mail: naider@mail.csi.cuny.edu
**Present address:* Postdoctoral Fellow at the Department of Physiology and Biophysics, Weill Medical College of Cornell University 1300 York Avenue, New York, NY 10021
†The Leonard and Esther Kurtz Term Professor at the College of Staten Island, CUNY

Contract grant sponsor: National Institutes of Health
Contract grant numbers: GM22086 and GM22087

Biopolymers (Peptide Science), Vol. 84, 508–518 (2006)

convenient alternative to growth in supplemented minimal media and is readily applicable to the expression of proteins selectively labeled with specific amino acids. Biopolymers (Pept Sci) 84: 508–518, 2006

This article was originally published online as an accepted preprint. The "Published Online" date corresponds to the preprint version. You can request a copy of the preprint by emailing the Biopolymers editorial office at biopolymers@wiley.com

Keywords: *transmembrane peptides; ^{15}N-amino acids; rich media; NMR; selective labeling*

INTRODUCTION

Structural studies of proteins are of vital importance to functional analysis and to practical applications such as developing new approaches for drug discovery and understanding disease mechanisms.[1] Elucidation of detailed protein structure has largely been performed by X-ray crystallography and/or nuclear magnetic resonance spectroscopy (NMR). NMR studies on proteins greater than 5–10 kDa require the molecule to be isotopically labeled using ^{13}C and/or ^{15}N incorporation.[2] With certain polypeptides, NMR spectra of the uniformly labeled protein are still too complicated and crowded to make complete assignments and to solve the structure of the molecule. In particular, polytopic helical transmembrane proteins often exhibit poorly resolved spectra and low chemical shift dispersion. This phenomenon is exacerbated by the redundancy of certain amino acid residues in the transmembrane domains. Furthermore, availability of peptides labeled selectively by amino acids is especially useful for resolving spectra of membrane peptides using solid state NMR spectroscopy.[3,4] The use of proteins with only one type of selectively labeled amino acid leads to considerable spectral simplification and may aid in the resolution of a structure or of an active site.[5,6]

Selective labeling of individual amino acids of a protein is a well-established technique and in most cases has been done in *Escherichia coli* cells grown in minimal medium (i.e., M9 medium) supplemented with one labeled amino acid and the remaining amino acids in unlabeled form, and/or with auxotrophic strains that do not possess the enzymatic machinery to biosynthesize a specific amino acid.[7–16] Selective labeling has also been done using *Baculovirus* infected insect cells[17,18] or cell-free translation systems[19,20] with relatively good results. Using these methods, sufficient protein has been produced to conduct biophysical analyses, and little cross-labeling of other amino acids has occurred during expression. However, in minimal medium expression yields are often reduced compared to expression of the same protein in rich medium.

An important aspect of expression studies is to maximize the amount of protein generated per liter of culture to reduce the cost of both uniformly and selectively labeled proteins. We have evaluated a system in which *E. coli* BL21(DE3)pLysS transformed cells are grown in rich medium and selectively labeled with ^{15}N-amino acids added to the medium. Using this system, we were able to generate more than 100 mg/L of a fusion protein containing a multidomain peptide corresponding to the third extracellular loop (EL), the seventh transmembrane domain (TM), and 40 residues of the cytosolic tail (CT) region (EL3–TM7–CT40) of a G-protein-coupled receptor (Ste2p) involved in yeast mating.[21] The fusion protein was isolated, subsequently cleaved to release the desired expression product, and analyzed using NMR spectroscopy. This ariticle discusses our approach to selectively labeling the EL3–TM7–CT40 peptide from Ste2p in rich medium. When compared to labeling in supplemented minimal media, this method may produce higher yields of selectively labeled peptide with significant percent incorporation and minimal cross-labeling of amino acids.

EXPERIMENTAL PROCEDURES

Vectors and Plasmids

The pREJ02M plasmid expresses the TrpΔLE–EL3–TM7–CT40 fusion protein (M7FP). EL3–TM7–CT40 codes for Ste2p(Ser267–Ser339) comprising the third extracellular loop (EL3), the 24-residue seventh transmembrane domain (TM7), and 40 residues of the cytoplasmic tail (CT40) with the mutation Met294Leu. This plasmid was constructed as previously described[21] and uses a T7 expression system, resulting in the desired expressed polypeptide fused to the carboxyl terminus of a portion of the histidine-tagged TrpΔLE1413 polypeptide.[22–24]

Strains

Escherichia coli expression strain BL21(DE3)pLysS was purchased as competent cells from Promega (Madison, WI). *Escherichia coli* DH5α cells used to amplify the engineered plasmids were purchased from Gibco BRL Life Technologies (Grand Island, NY).

Media, Solvents, and Buffers

Luria broth (LB) was prepared by resuspending 25 g of Miller's LB (Sigma, St. Louis, MO) in 1 L of water. LB is a

rich, undefined medium, composed of tryptone, which supplies peptides, amino acids, and other undetermined ingredients generated by degradation of casein by trypsin during tryptone preparation; yeast extract, which supplies vitamins, trace elements, and many undetermined organic compounds; and sodium chloride. The ^{15}N-specifically labeled amino acid rich medium was made by adding 2, 0.5, or 0.1 g of ^{15}N-labeled amino acid (Spectra Stable Isotopes, Columbia, MD, and Cambridge Isotope Laboratories, Andover, MA) to 1 L of LB. The ^{15}N-specifically labeled alanine supplemented M9 minimal medium was made by mixing 200 mL of 5× M9 salts (15 g of KH_2PO_4, 34 g of Na_2HPO_4, 2.5 g of NaCl, 5 g of NH_4Cl), 100 mg of ^{15}N-alanine, 100 mg of the remaining 19 unlabeled amino acids, 8 mL of 50% glucose, 1 mL of 2*M* $MgSO_4$, 0.2 mL of 0.5*M* $CaCl_2$, and 788.3 mL of deionized water.[25] Plates of LB medium were prepared as for the liquid medium with the addition of 15 g of agar. All media and their components were sterilized by autoclaving except the glucose solution, which was filter sterilized and added to the cooled media. After autoclaving, filter-sterilized ampicillin and chloramphenicol were added to the cool media to a final concentration of 200 and 35 μg/mL, respectively.

Lysis solution contained lysis buffer [50 m*M* Tris-HCl, 1 m*M* ethylenediaminetetracetic acid (EDTA)], 1 m*M* phenylmethylsulfonyl fluoride (PMSF), and 300 μg/mL lysozyme, pH 8.7. Inclusion body washing buffer contained 1% Nonidet P-40 (igepal Ca-630) and 1% deoxycholic acid in lysis buffer. Guanidinium buffer was made of 6*M* guanidinium hydrochloride, 10 m*M* Na_2HPO_4, 138 m*M* NaCl, and 2.7 m*M* KCl, pH 7.4. A sample buffer of 4× SDS was prepared from 4 mL of a 20% SDS solution, 2.4 mL glycerol, 400 μL 2-mercaptoethanol, 1 mL 1 M Tris-HCl, pH 6.8, and a small amount of Brilliant Blue G in 20 mL of H_2O.

High performance liquid chromatography (HPLC) grade solvents used for HPLC purification such as methanol, acetonitrile, and H_2O were purchased from J. T. Baker (Phillipsburg, NJ). Trifluoracetic acid (TFA) and trifluoroethanol (TFE) were purchased from Sigma (St. Louis, MO). Deuterated d_2-TFE was purchased from Cambridge Isotope Laboratories (Andover, MA) and $CDCl_3$ and CD_3OH were purchased from Sigma (St. Louis, MO).

Protein Expression

Escherichia coli BL21(DE3)pLysS cells were transformed with the pREJ02M plasmid and two colonies of the transformed cells were inoculated into 5 mL of LB medium containing ampicillin (200 μg/mL) and chloramphenicol (35 μg/mL) and allowed to grow at 37°C and 225 rpm to log phase [optical density (OD_{600}) ~ 0.6]. Two milliliters of this culture were then harvested at 6000 rpm for 3 min, washed in 1 mL ^{15}N-selectively labeled media, and inoculated into 1 L of the same media, LB or M9. The cells were grown to log phase (OD_{600} ~ 0.6), induced with 1 m*M* isopropyl-β-D-thiogalactopyranoside (IPTG), incubated for 6 h, and harvested by centrifugation. To optimize the time of expression, collection of 1 mL samples every hour of induction was performed.

Isolation of Inclusion Bodies, Purification, and CNBr Cleavage

Isolation and purification of the fusion protein inclusion bodies as well as the subsequent CNBr cleavage and purification of the 73-residue EL3–TM7–CT40 peptide from the N-terminal TrpΔLE protein were carried out as described.[21]

NMR Spectroscopy

To prepare the ^{15}N-selectively labeled EL3–TM7–CT40 peptide samples for NMR, 1.2–1.4 mg of peptide was first dissolved in 250 μL d_2-TFE and then 250 μL H_2O + 0.1% TFA was added to give the final ratio of TFE and H_2O (1:1, v/v) and approximately 300 μ*M* peptide concentration. Heteronuclear single quantum correlation (HSQC) spectra of the 73-residue EL3–TM7–CT40 peptides were recorded on a four-channel Varian UNITY INOVA 600-MHz NMR spectrometer (Varian NMR Instruments, Palo Alto, CA) equipped with a *Z*-axis pulsed-field gradient and a Varian 5-mm ^{1}H/^{13}C/^{15}N triple resonance probe at 25°C. The acquisition time of the HSQC spectra ranged from 1 h 20 min to 2 h, collecting 8–12 scans with a relaxation delay of 2 s and 128 complex points in the indirect dimension. Observation of HSQC spectra after a 20-min acquisition time resulted in well-resolved spectra and there was no difficulty in observing the selectively labeled amino acid residues. However, the signal-to-noise ratio was not sufficient for quantitative experiments used to compare percent incorporation and the amount of cross-labeling. Peak assignment in the HSQC spectrum of EL3–TM7–CT40 in TFE/water (1:1) has been reported in our previous publication.[21] After NMR spectra were measured, the peptide solutions were added to a 59% acetonitrile and 41% water solution and relyophilized. The resulting products were then used for NMR in $CDCl_3$:CD_3OH:H_2O (4:4:1 v/v). To prepare these samples, the peptides were first suspended in $CDCl_3$. The peptide did not dissolve completely until CD_3OH, and finally H_2O + 0.1% TFA were added. If too much water was added, the solution separated into two layers. Acidic water was used to aid in water suppression. All peptide concentrations were based on the measured weights of the samples.[21]

Estimation of Percent Cross-Labeling and Percent Incorporation

Mass spectra were measured using an Agilent 1100 LC/MS instrument equipped with an electrospray ion source and an Ion Trap detector. The proteins were dissolved in a 1,1,1,3,3,3-hexafluoro-2-propanol/acetic acid (1:1 v/v) mixture and injected into the LC/MS instrument running with an acetonitrile/water (+0.5% acetic acid) chromatographic eluent. MS data was used as a primary source of information to estimate percentage of isotopic labeling during amino acid incorporation. This method can be considered

absolute as it measures the amount of isotope in every sample and does not need to be related to a sample with a known percentage of incorporation as a standard, as is necessary in using other techniques such as NMR. At the same time, the accuracy of this method is limited when labeling just a few amino acids in a long peptide.

To complement the MS studies, peak volumes in ^{1}H–^{15}N HSQC NMR spectra were used to estimate the relative amount of ^{15}N-isotope label in different samples and to learn about the specificity of amino acid labeling during the incorporation of different labeled amino acids. To correlate NMR peak intensities in the two-dimensional (2D) HSQC spectra of different samples more accurately, we first attempted to equalize the concentrations of different samples by weighing the same amount of peptide. The relative concentrations of the different samples were then determined using the total integrated intensity of all peptide peaks in the aliphatic region of the corresponding one-dimensional (1D) ^{1}H spectra, as this parameter is not effected by ^{15}N labeling. These relative concentrations were then used to normalize the ^{1}H–^{15}N HSQC peak volumes between different samples.

The percent cross-labeling was estimated by dividing the peak volumes due to amino acid residues not added to the growth medium by the peak volumes of the amino acids that we were attempting to selectively label. These values are estimates because the relaxation times and magnetization transfer efficiencies vary with the type and position of an amino acid in the peptide chain and thus the peak volumes are not directly proportional to the number of atoms. To determine the percent incorporation, the experimental molecular weight (MW) of the unlabeled peptide was subtracted from that of the labeled peptide. The difference was then divided by the number of specific amino acid residues (i.e., Leu = 9) contained in EL3-TM7-CT40. The percent incorporation calculated in this manner was corrected for the cross-labeling determined for a given sample.

RESULTS

Expression and Purification of Selectively Labeled Peptides

Examination of the primary structure of the 73-mer peptide indicated that the most useful amino acids to label for NMR clarification based on sequence and location would be leucine, serine, and alanine. Leucine was chosen due to its relative abundance in the TM7 helix, serine residues were replaced because they often overlapped with each other and with threonine, and alanine was chosen because it is a relatively inexpensive labeled amino acid that could be used for the optimization experiments.

Selective labeling was performed by adding 2 g of either ^{15}N-leucine, ^{15}N-serine, or ^{15}N-alanine per 1 L of LB, a rich medium. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) analysis of each of the fusion proteins indicated that, as expected, the addition of 2 g of a ^{15}N-labeled amino acid to the LB medium did not appear to interfere with the expression properties of the fusion protein encoded by the pREJ02M plasmid (Figure 1). Because SDS-PAGE cannot give precise estimates of protein molecular weight, the molecular weight of each fusion protein was confirmed by mass spectrometry (data not shown). The yield of fusion protein was similar to that of non-supplemented LB, approximately 100–145 mg/L of each labeled polypeptide was isolated. After CNBr cleavage, an average of approximately 0.2 mg of pure selectively labeled EL3–TM7–CT40 peptide was obtained per milligram of each fusion protein ($\sim$53% yield on a weight basis). The difference between the experimental mass of the cleaved ^{15}N-Ser and ^{15}N-Ala peptides and the mass calculated for 100% insertion of the isotopically labeled amino acid into the peptide was less than 1 Da, indicating that the labeling was successful with a high incorporation of ^{15}N into the respective peptides (Table I). In contrast, the experimental MW of the ^{15}N-Leu peptide was more than 2 Da lower than the expected mass for full incorporation of ^{15}N-Leu into the peptide, indicating that the percent incorporation was not as high (Table I).

NMR Analysis of Selectively Labeled Peptides

Selective labeling of certain amino acid residues in a membrane protein can be very valuable in deciphering its NMR spectrum. In particular, simplification of the ^{1}H–^{15}N HSQC spectrum of the uniformly labeled EL3–TM7–CT40 peptide was helpful for confirming the correct assignment of NMR peaks in chloroform : methanol : water (4:4:1 v/v). Selective labeling will also be useful in future work when attempting to resolve the NMR spectra of fragments of Ste2p containing double or triple transmembrane domains. The HSQC spectra of ^{15}N-Leu and ^{15}N-Ser labeled EL3–TM7–CT40 contained the expected number of leucine (Figure 2B) or serine (Figure 2D) cross peaks, respectively, as well as some peaks due to incorporation of the ^{15}N moiety into other amino acids (as described below). Conversely, only the expected cross peaks were observed for ^{15}N-Ala labeled EL3–TM7–CT40 (Figure 2C).

The mass of the specifically labeled EL3–TM7–CT40 peptide reflected contributions from incorporation of the specifically labeled amino acid and from ^{15}N that was scrambled into other residues. We used NMR spectroscopy to assess the fidelity of the selective labeling and the percent of the ^{15}N that cross-la-

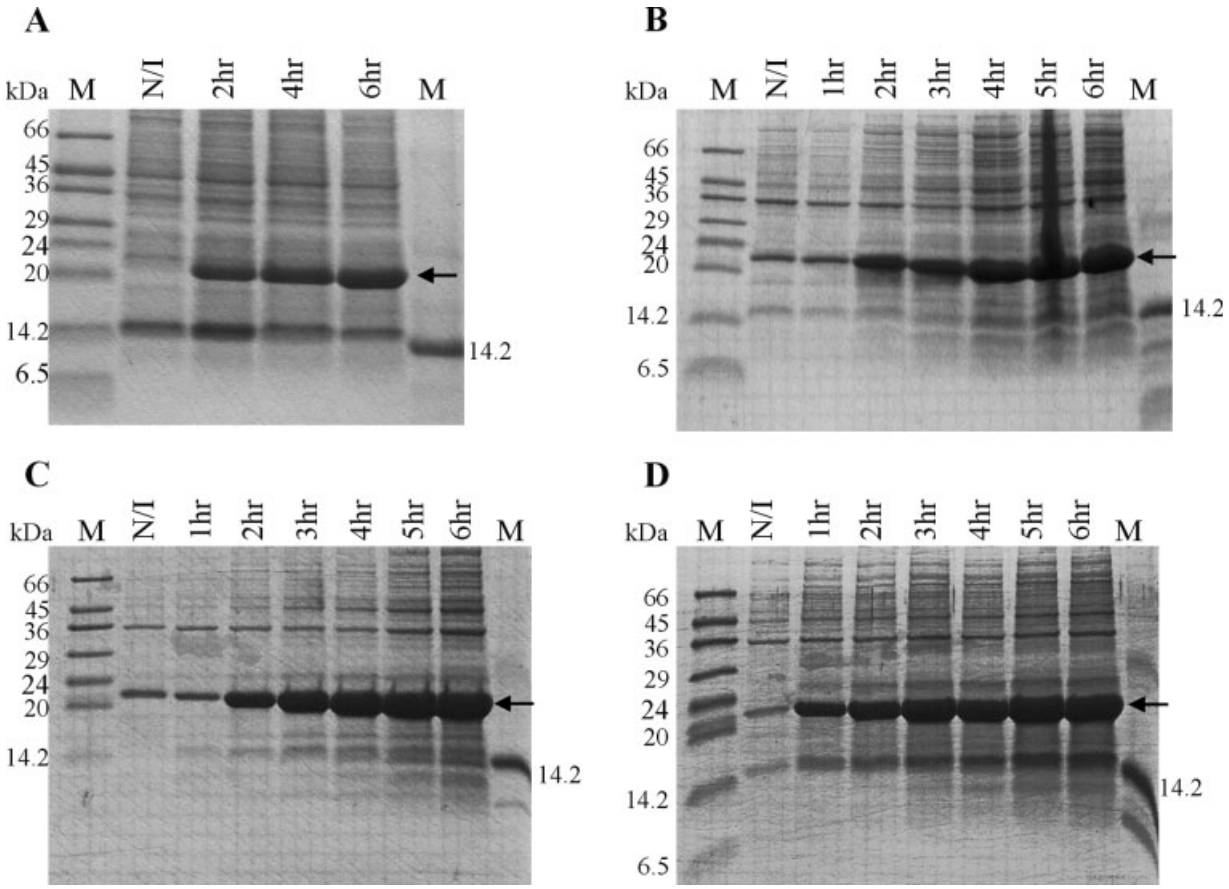

FIGURE 1 Expression of isotopically labeled TrpΔLE–EL3–TM7–CT40 fusion protein from the pREJ02M plasmid after induction with 1 m*M* IPTG in LB. The expression of the M7FP (MW $\sim$ 21.2 kDa) was monitored every hour for up to 6 h and visualized via SDS-PAGE. M = marker, N/I = noninduced, *x* h = time after addition of IPTG. The arrow indicates the protein of interest. (A) Unlabeled M7FP (as seen in Ref. 21), (B) ^{15}N-Leu labeled M7FP, (C) ^{15}N-Ala labeled M7FP, and (D) ^{15}N-Ser labeled M7FP.

beled other residues. To conduct this experiment we used the peak areas in the aliphatic region of the one dimensional proton spectrum to adjust the concentration of all samples to the same value. We then measured the cross-peak volumes of the ^{15}N-HSQC spectrum to determine the relative incorporation under

Table I Label Incorporation and Cross-Labeling in EL3-TM7-CT40 Peptides

Sample Name	Expected MW (Da)	Experimental MW (Da)	% ^{15}N Label Incorporation	Cross-Labeling
Unlabeled	7670.5	7670.4	NA[b]	NA[b]
[2 g ^{15}N Leu]	7679.5	7677.3	72	Yes 7%
[2 g ^{15}N Ser]	7682.5	7683.0	96	Yes 11%
[2 g ^{15}N Ala]	7677.5	7677.0	94	No
[0.5 g ^{15}N Ala]	7677.5	7675.3	70	No
[0.1 g ^{15}N Ala]	7677.5	7672.4	29	No
[MM[a] ^{15}N Ala]	7677.5	7673.1	38	No

[a] MM = supplemented M9 media with [^{15}N]-Ala and 19 other unlabeled amino acids.
[b] NA = not applicable.

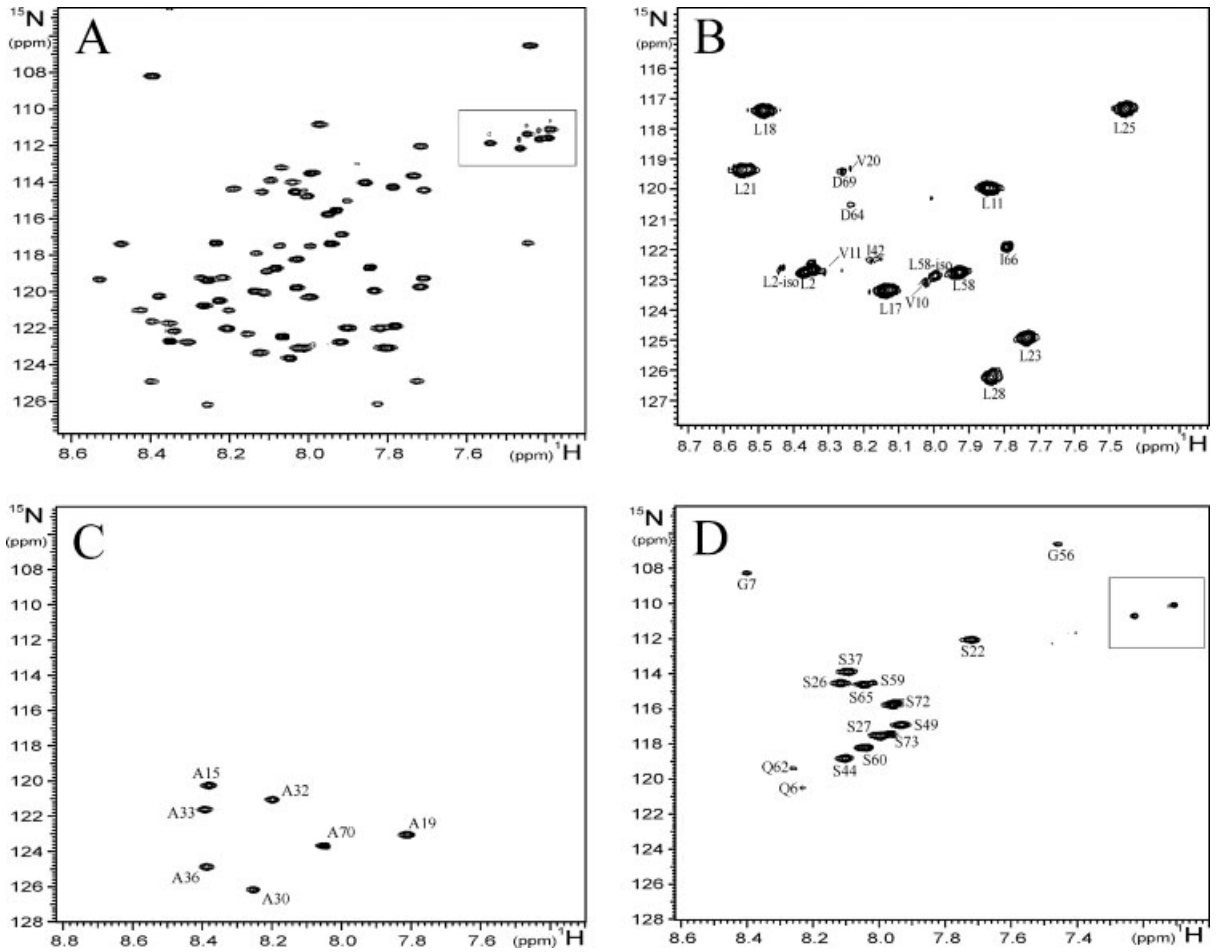

FIGURE 2 ^{1}H–^{15}N HSQC spectra of EL3–TM7–CT40 in TFE/H_2O (1:1). (A) ^{15}N-uniformly labeled EL3–TM7–CT40 (as seen in Ref. 21), (B) 2 g/L ^{15}N-Leu EL3–TM7–CT40, (C) 2 g/L ^{15}N-Ala EL3–TM7–CT40, and (D) 2 g/L ^{15}N-Ser EL3–TM7–CT40. The boxed regions indicate labeled side chains of Asn and/or Gln.

different culture conditions (Table II). By comparing the volumes of cross peaks due to the labeled amino acids (Leu or Ser) with the volumes of the other peaks, we could estimate the percent cross-labeling (Table I). In the latter calculation, we determined the total volume of all peaks and then divided the volume due to the non-Leu or non-Ser peaks by the total volume of all cross peaks in the HSQC spectrum. To determine the percent incorporation, we corrected the expected molecular weight for the percent of label that was scrambled into other residues. Details of this calculation are given in the Methods section.

Table II ^{1}H–^{15}N HSQC Cross-Peak Volumes for EL3–TM7–CT40 Peptides Labeled with Individual Amino Acids

2 g/L [^{15}N]-Leu in LB		2 g/L [^{15}N]-Ser in LB		2 g/L [^{15}N]-Ala in LB	
Amino Acid	Peak Volume	Amino Acid	Peak Volume	Amino Acid	Peak Volume
9 Leu	574.15	11 Ser	1092.57	7 Ala	638.88
2 Ile	25.55	2 Gly	25.56		
3 Val	9.26	3 Asp	13.02		
2 Asp	7.61	1 Asn	3.03		
		Asn/Gln side chains	89.9		
Total peak volume	616.57		1225.08		638.88

Leucine Incorporation

Based on the MS results described above, the ^{15}N-Leu appeared to be incorporated with less efficiency than the ^{15}N-Ser or ^{15}N-Ala. This conclusion considers the 2 Da difference observed between the expected MW, if all of the Leu residues were labeled, and the observed MW (Table I). As expected, nine ^{15}N-Leu peaks were observed. In addition, two of the leucines showed multiple peaks due to the presence of conformational isomerism likely reflecting *cis–trans* isomerism of peptide bonds involving the proline nitrogen. There was also some cross-labeling of Ile, Val, and Asp residues, each with a peak volume less than 20% of the average peak volume of a Leu residue (Figure 2 and Table II). The cross-labeling of Ile and Val may be explained by the fact that these residues are all in the same biosynthetic pathway.[26] We are less sure about the origin of the Asp residues that were cross-labeled, although interconversions of amino acids may occur by transamidations and transcarboxylations in *E. coli* as well as other catabolic and anabolic transformations of amino acids and their precursors.[27] When both the MS and the NMR analyses are taken into account, the percent of ^{15}N incorporated in EL3–TM7–CT40 prepared as described in Experimental Procedures is 72% (Table I).

Alanine Incorporation

When supplemented at a level of 2 g/L, ^{15}N-Ala was incorporated into the peptide to a level of 94% of theoretical (Table 1), as observed by MS. Seven out of seven ^{15}N-Ala residues were observed in the NMR spectrum (Figure 2C). There was no cross-labeling to other amino acids and the peaks were clearly defined. *Escherichia coli* is highly demanding of this amino acid, which appears to be converted quickly into proteins and peptidoglycan before catabolism may occur.[26,27]

Serine Incorporation

Based on the MS analysis, alone it would appear that almost all of the Ser residues in the peptide were labeled to completion with the ^{15}N-moiety (see Table I). There are 12 serine residues in the primary structure of the peptide, but only 11 residues are expected to be observed by NMR due to the location of one of the residues at the N-terminus of the peptide, where the proton exchange is very fast and the peak of a labeled amine group would be too broad to observe. The 11 expected residues were observed as well as some cross-labeled amino acid residues (Figure 2D and Table II). The presence of two Gly residues probably is due to the breakdown of ^{15}N-Ser into ^{15}N-Gly by glycine hydroxyl methyltransferase, GlyA in *E. coli*.[26,27] There are also 3 Asp peaks and multiple peaks of Asn and/or Gln backbone and side-chain residues. These latter peak volumes are low and contribute very little to the amount of cross-labeling. The percent incorporation of ^{15}N from ^{15}N-Ser was 96%, with approximately 11% cross-labeling (Table I).

Optimizing the Amount of Labeled Amino Acid in LB Medium

The addition of 2 g/L of ^{15}N-labeled amino acid could be seen as cost prohibitive when compared with the cost of addition of 0.1 g of labeled amino acid per liter that is widely employed in labeling in supplemented minimal media. To evaluate the amount of selective amino acid required for effective labeling in rich medium, the labeling was performed by supplementing 0.5 or 0.1 g/L of ^{15}N-alanine to LB-rich media and by supplementing M9 minimal medium with 0.1 g of ^{15}N-alanine and 0.1 g of the other 19 unlabeled amino acids. PAGE analysis indicated that under these conditions expression levels were not noticeably changed (data not shown) and yields of the fusion protein per 1 L of growth were between 100 and 133 mg/L of culture.

The ^{1}H–^{15}N HSQC spectra of the above peptides all produced the same 7 alanine peaks with similar chemical shifts as that seen in the HSQC spectrum of the 2 g [^{15}N-Ala]–EL3–TM7–CT40 peptide (Figure 3). Because there was no cross-labeling of other amino acid peaks, MS analysis alone was sufficient to determine percent incorporation of the ^{15}N-Ala. MS analysis of the cleaved and purified EL3–TM7–CT40 peptide product indicated that the experimental mass of these peptides decreased as the amount of ^{15}N-Ala decreased, which correlates with a loss in ^{15}N incorporation (Table I). Supplementation with 0.5 g ^{15}N Ala/L resulted in 70% of the Ala labeled with ^{15}N incorporation, while supplementation with 0.1 g ^{15}N Ala/L gave 29% isotopic labeling. This compares with 38% isotopic labeling found using the supplemented minimal medium and 94% labeling found using 2 g ^{15}N-Ala per liter (Table I).

Even at the lowest amount of incorporation, the peaks are intense enough to be used as aides in resolving NMR spectra of a uniformly labeled peptide (Figure 3). In fact, the preparation from 0.1 g ^{15}N-Ala gave a clean, well-resolved spectrum that was measured in just 2 h. The low amount of isotopic labeling in the minimal media is surprising, but could be indicative that the amount of ^{15}N-Ala was not sufficient to label 100 mg of fusion protein as well as the

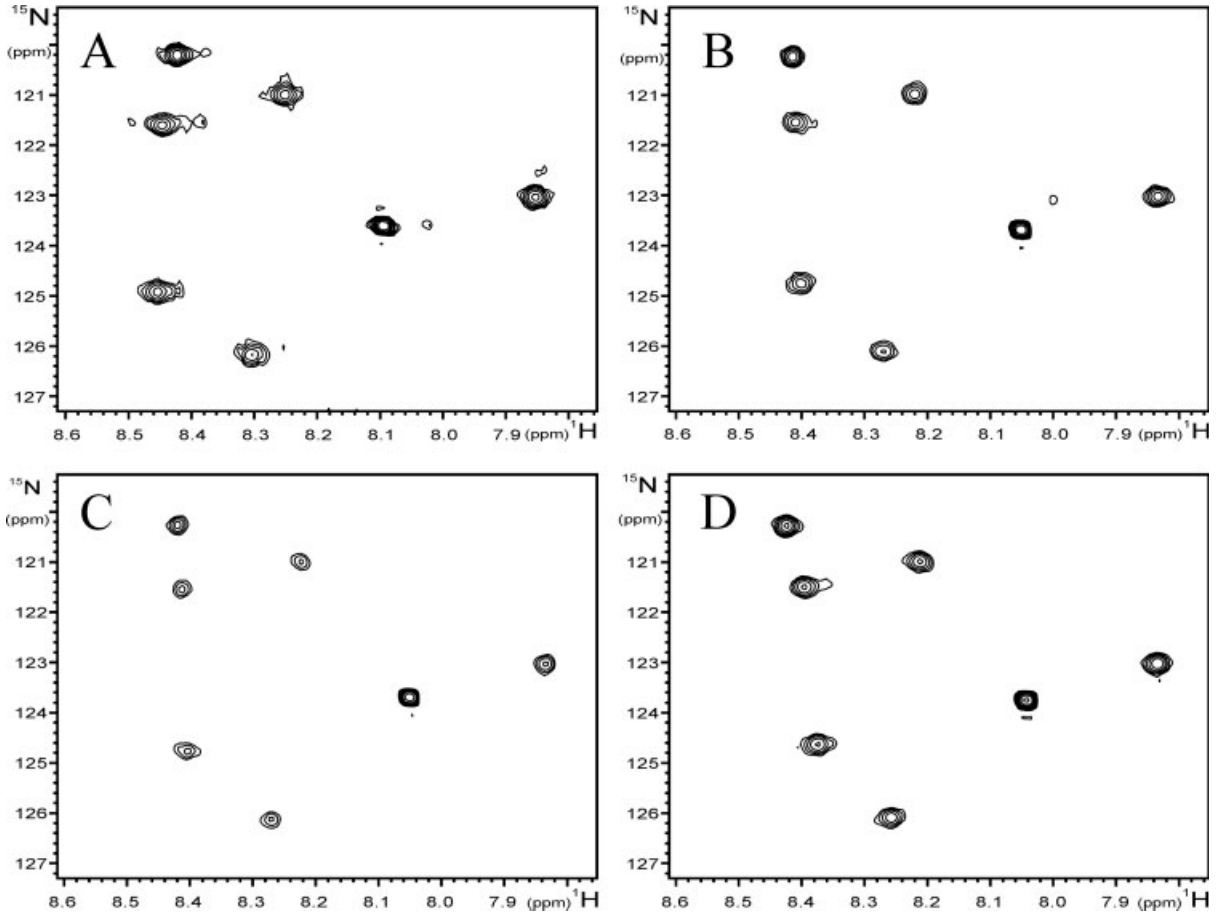

FIGURE 3 ^{1}H–^{15}N HSQC spectra of EL3–TM7–CT40 selectively labeled with ^{15}N-Ala using different levels of amino acid in the growth medium. (**A**) LB with 2 g of ^{15}N-Ala/L of medium, (B) LB with 0.5 g of ^{15}N-Ala/L of medium, (C) LB with 0.1 g of ^{15}N-Ala per liter of medium, and (D) M9 supplemented with 19 naturally occurring amino acids and 0.1 g of ^{15}N-Ala/L of medium. All spectra were on a solution of $\sim$1.3 mg peptide in 0.5 mL of trifluorethanol-d_2/water (1:1 v/v). The HSQC spectra were collected under identical NMR instrument conditions and represent 2 h of acquisition time. The printouts were adjusted to the same vertical gain in an attempt to represent the actual strength of the signal.

other cellular proteins. If this is the case, then the Ala would have to become available through the breakdown of other, nonlabeled, amino acids supplemented in the media. For example, Val, Cys, and Gln all play a role in the biosynthetic pathway leading to Ala synthesis,[26] so if there were any excess of any of these amino acids and a deficiency of alanine, unlabeled Ala could be made and incorporated into the fusion protein. In fact, when the minimal media was supplemented with 0.1 g ^{15}N-Ala and 18 other amino acids (excluding Tyr), less protein was made (75 mg), but the incorporation of ^{15}N was higher (60%). Our observation was that the 0.5 g ^{15}N-Ala supplemented LB resulted in a stronger HSQC spectrum for EL3–TM7–CT40 than the 0.1 g ^{15}N-Ala supplemented M9 medium, whereas this latter medium gave a stronger spectrum for this peptide than from expression in 0.1 g ^{15}N-Ala supplemented LB (Figure 3). Nevertheless, the spectrum from growth in 0.1 g ^{15}N-Ala supplemented LB (Figure 3C) was still more than adequate for NMR investigations and it is comparable to that measured on a sample prepared using supplemented minimal medium (Figure 3D).

DISCUSSION

The NMR spectrum of a peptide can provide a wealth of information about its structure, molecular interactions, and even its functions. However, many times a peptide's NMR spectrum is very complex, making it difficult to assign specific resonances to particular

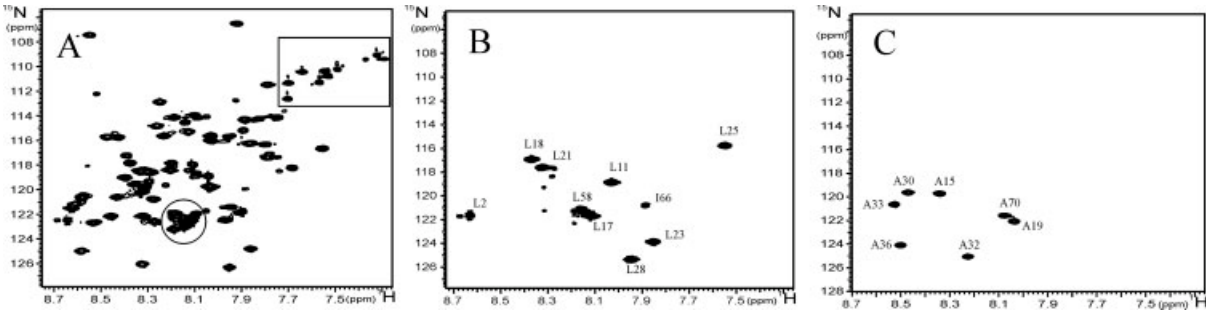

FIGURE 4 ^{1}H–^{15}N HSQC spectra of EL3–TM7–CT40 in chloroform : methanol : water (4:4:1 v/v). (A) ^{15}N-uniformly labeled EL3–TM7–CT40 (data from reference [21]), (B) 2 g/L ^{15}N-Leu EL3–TM7–CT40, (C) 2 g/L ^{15}N-Ala EL3–TM7–CT40. The overlapped region 8.1/122 ppm discussed in the text is indicated by a circle. The boxed regions indicate labeled side chains of Asn and/or Gln.

atoms of the peptide. This is especially true for highly helical proteins, where the chemical shift dispersion is rather small. Moreover, the problem is exacerbated for integral membrane proteins, which often contain significant numbers of similar amino acid residues in their helical domains. Finally, even in globular proteins selective labeling results in simplification that can prove beneficial in studies of ligand–protein or protein–protein interaction.

One way to circumvent this problem is through selective labeling of specific amino acids in the peptide. This can be done in bacterial cells through supplementation of minimal media[5,8,9,12,14] or through use of an auxotroph[11,16,28] or both.[6,10,13,15] Usually, unlabeled amino acids are added to the minimal media along with the ^{15}N-labeled amino acid. Most laboratories use a 100 mg/L ratio of all amino acids based on the work of Muchmore et al.,[7] but smaller and larger amounts have also been used.[15,29] Both the use of minimal media and bacterial auxotrophs have been successful in that 20–100 mg of a selectively labeled protein per liter of growth have been isolated with little amino acid scrambling.[5,8,9,14,16] Because minimal medium contains nutrients the cell needs for growth in simple form (i.e., ammonium or glucose), the cell has to expend energy to synthesize essential amino acids. If the cell is provided with a surplus of all the amino acids, it should be able to shut down all its amino acid biosynthetic pathways and just use the amino acids provided. It is thought that with the repression of the pathways the chance of cross-labeling significantly decreases.

Consistent with results reported by other laboratories, we had previously observed that the yields of fusion protein are much greater in rich medium than in unsupplemented minimal medium.[21] As reported in Estephan et al.,[21] recovery of the TrpΔLE–EL3–TM7–CT40 fusion protein was ~100–125 mg/L for cells grown in LB and 30–50 mg/L in unsupplemented minimal medium. In a parallel study, no expression of EL3–TM7–CT40 as a thioredoxin fusion protein was observed in minimal medium.[21] Therefore, in this latter case, isotopic labeling could only have been performed in LB. Given the high yield found in LB, we attempted to develop a method to selectively label protein in this rich medium. This strategy is straightforward and easily applied in that we simply add an excess of one ^{15}N-labeled amino acid to the rich growth medium in an attempt to suppress its individual biosynthetic pathway, thereby forcing the cell to incorporate only the exogenously added labeled amino acid into the expression product. Based on the components of LB medium, a rough calculation was made that the addition of 2 g of ^{15}N-leucine to 1 L of LB would result in approximately 80% of the leucine in the expressed peptide having a ^{15}N-label. It was decided to use this amount of amino acid in expectation that this would result in selective labeling with the amino acid of interest without significant scrambling.

The ^{1}H–^{15}N HSQC spectrum of the uniformly labeled [^{15}N]–EL3–TM7–CT40 in TFE/water showed good dispersion for a membrane peptide (Figure 2A). However, additional studies using this peptide performed in chloroform:methanol:water (Figure 4) and in the presence of dodecylphosphocholine micelles (data not shown) proved far more difficult to interpret. For example, in the HSQC spectrum measured for [^{15}N]–EL3–TM7–CT40 in chloroform:methanol:water, 8 amino acid residues appear in an unresolved cluster of cross peaks near 8.1/122 ppm (Figure 4A). Four of these residues (Ala19, Leu17, Leu58, Ala70) are readily resolved in the HSQC spectra of the selectively labeled polypeptides (Figures 4B and 4C) and this approach aided the peak assignments.[21] Thus, in both TFE-based and chloroform/methanol-based organic

aqueous media, selective labeling of EL3–TM7–CT40 with leucine, alanine, and serine in rich medium was used to simplify the spectra and confirm the amino acid assignments (Figures 2 and 4). These particular amino acids were chosen because of their abundant presence in TM7, the tendency of some of these residues to exhibit degenerate chemical shifts, and their relatively low cost.

Addition of the 2 g of labeled amino acid to rich medium was very successful in producing large amounts of pure selectively labeled EL3–TM7–CT40 peptide with high percent incorporations calculated for each amino acid (Table I). Labeling with ^{15}N-alanine produced the seven expected peaks in the ^{15}N-HSQC spectrum and no cross-labeling of other amino acids was observed. However, the HSQC spectra of the ^{15}N-leucine and ^{15}N-serine labeled EL3–TM7–CT40 peptides indicated that growth in rich medium resulted in cross-labeling of other amino acids by the ^{15}N moiety. It appears that most of the amino acids that are cross-labeled tend to be within a biosynthetic pathway. For instance, selective labeling with ^{15}N-leucine led to a small amount of labeled valine and isoleucine, both members of the pyruvate family of amino acids.[26] Also, some of the "extra" peaks observed after labeling with ^{15}N-serine were glycine residues. Glycine is a direct product of serine in its biosynthetic pathway.[26] The low amount of cross-labeling that we observed using rich medium supplemented with 2 g/L of ^{15}N-leucine or ^{15}N-serine may decrease if limiting amounts of these labeled amino acid were added per liter of LB. It should also be noted that serine has a strong tendency to cross-label even in supplemented minimal medium (K. E. Tortolani, Spectra Stable Isotopes, personal communication). Because the goal of the selective labeling in this investigation was to simplify the analysis of the NMR spectrum of polypeptides, a few extra peaks with decidedly lower intensities than the labeled amino acid should be tolerable.

The main drawback to the method described above is that adding 2 g/L of ^{15}N-labeled amino acid to selectively label a peptide in rich media compared to the 0.1 g/L typically supplemented into minimal media can be very costly. Therefore, it was decided to see if the excess ^{15}N-labeled amino added to the rich medium could be reduced enough to produce results similar to that of minimal medium supplemented with 0.1 g/L of labeled amino acid. To test this proposal, M7FP was expressed in LB medium containing 0.5 g ^{15}N Ala/L of growth, LB containing 0.1 g ^{15}N-Ala per liter of growth, and in M9 minimal medium supplemented with 0.1 g ^{15}N-Ala plus 0.1 g of the additional amino acids per liter. Percent incorporation of the labeled amino acid decreased as the amount of amino acid added to rich media decreased, 0.5 g/L was 70% incorporated and 0.1 g/L was 29% incorporated (Table I). The peptide grown in minimal medium supplemented with the 20 naturally occurring amino acids also showed a relatively low incorporation (38%) of isotopically labeled alanine (Table I). This may be because so much peptide was produced that the cell simply exhausted the supply of the labeled amino acid during cell growth. To compensate for the decrease in incorporation, one could consider adding more labeled and unlabeled amino acids (i.e., 0.5 g/L) to the supplemented minimal media.[5,13,15,29] This would potentially increase the percent incorporation while possibly decreasing the cross-labeling. However, increasing the amount of label from amounts normally used (0.05–0.1 g/L)[7,9,14] would be costly. The HSQC spectra indicate that even at 29% incorporation the peaks are still intense enough to aid in the analysis of the spectra for the uniformly labeled peptide (Figure 3C). Given the simplicity of using rich medium and the improved growth observed for many expression strains, it would seem that simply adding one labeled amino acid would be advantageous to preparing minimal medium containing supplementation with all amino acids. The fact that the percent labeling obtained using 0.1 g/L of ^{15}N-alanine was only marginally reduced compared to labeling in the 0.1 g/L amino acid supplemented minimal medium and the likelihood that growth in rich medium will be more efficient and will generate increased protein expression, suggests to us that the approach evaluated in this article should be considered when attempting to selectively label proteins with a specific amino acid residue.

This work was supported by research grants GM22086 and GM22087 from the National Institutes of Health.

REFERENCES

1. Congreve, M.; Murray, C. W.; Blundell, T. L. Drug Discov Today 2005, 10, 895–907.
2. Gardner, K. H.; Kay, L. E. Annu Rev Biophys Biomol Struct 1998, 27, 357–406.
3. Opella, S. J.; Marassi, F. M. Chem Rev 2004, 104, 3587–606.
4. Nielsen, N.; Malmendal, A.; Vosegaard, T. Mol Membr Biol 2004, 21, 129–141.
5. Park, S. H.; Mrse, A. A.; Nevzorov, A. A.; Mesleh, M. F.; Oblatt-Montall, M.; Montal, M.; Opella, S. J. J Mol Biol 2003, 333, 409–424.
6. Vance, C. K.; Kang, Y. M.; Miller, A. F. J Biomol NMR 1997, 9, 201–206.

7. Muchmore, D. C.; McIntosh, L. P.; Russell, C. B.; Anderson, D. E.; Dahlquist, F. W. Methods Enzymol 1989, 177, 44–73.
8. Ramesh, V.; Frederick, R. O.; Syed, S. H. E.; Gibson, C. F.; Yang, J.-C.; Roberts, G. C. Eur J Biochem 1994, 225, 601–608.
9. Gettins, P. G. W.; Backovic, M.; Peterson, F. C. Methods 2004, 32, 120–129.
10. Studts, J. M.; Fox, B. G. Protein Expr Purif 1999, 16, 109–119.
11. McIntosh, L. P.; Griffey, R. H.; Muchmore, D. C.; Nielson, C. P.; Redfield, A. G.; Dahlquist, F. W. Proc Natl Acad Sci USA 1987, 84, 1244–1248.
12. Kalus, W.; Zweckstetter, M.; Renner, C.; Sanchez, Y.; Georgescu, J.; Grol, M.; Demuth, D.; Schumacher, R.; Dony, C.; Lang, K.; Holak, T. A. EMBO J 1998, 17, 6558–6572.
13. Ou, H.; Lai, H. C.; Serber, Z.; Dotsch, V. J Biomol NMR 2001, 21, 269–273.
14. Peterson, F. C.; Gordon, N. C.; Gettins, P. G. Biochemistry 2001, 40, 6275–6283.
15. Cheng, H.; Westler, W. M.; Xia, B.; Oh, B.; Markley, J. L. Arch Biochem Biophys 1995, 316, 619–634.
16. Badii, R.; Basran, J.; Casarotto, M. G.; Roberts, G. C. Protein Expr Purif 1995, 6, 237–243.
17. Strauss, A.; Bitsch, F.; Cutting, B.; Fendrich, G.; Graff, P.; Liebetanz, J.; Zurini, M.; Jahnke, W. J Biomol NMR 2003, 26, 367–372.
18. Brüggert, M.; Rehm, T.; Shanker, S.; Georgescu, J.; Holak, T. A. J Biomol NMR 2003, 25, 335–348.
19. Parker, M. J.; Aulton-Jones, M.; Hounslow, A. M.; Craven, C. J. J Am Chem Soc 2004, 126, 5020–5021.
20. Yabuki, T.; Kigawa, T.; Dohmae, N.; Takio, K.; Terada, T.; Ito, Y.; Laue, E.; Cooper, J. A.; Kainosho, M.; Yokoyama, S. J Biomol NMR 1998, 11, 295–306.
21. Estephan, R.; Englander, J.; Arshava, B.; Samples, K. L.; Becker, J. M.; Naider, F. Biochemistry 2005, 44, 11795–11810.
22. Miozzari, G. F.; Yanofsky, C. J Bacteriol 1978, 133, 1457–1466.
23. Kleid, D. G.; Yansura, D.; Small, B.; Dowbenko, D.; Moore, D. M.; Grubman, M. J.; McKercher, P. D.; Morgan, D. O.; Robertson, B. H.; Bachrach, H. L. Science 1981, 214, 1125–1129.
24. Staley, J.; Kim, P. Protein Sci 1994, 3, 1822–1832.
25. Marley, J.; Lu, M.; Bracken, C. J Biomol NMR 2001, 20, 71–75.
26. Keseler, I. M.; Collado-Vides, J.; Gama-Castro, S.; Ingraham, J.; Paley, S.; Paulsen, I. T.; Peralta-Gil, M.; Karp, P. D. Nucleic Acids Res 2005, 33, D334–337.
27. Neidhardt, F. C.; Curtiss, R., 3rd; Ingraham, J. L.; Lin, E. C. C.; Low, K. B.; Magasanik, B.; Reznikoff, W. S.; Rileym M.; Schaechter, M.; Umbarger, H. E., Eds. Escherichia coli and Salmonella: Cellular and Molecular Biology, 2nd ed.; ASM Press: Washington, DC, 1996.
28. Petkova, A. T.; Baldus, M.; Belenky, M.; Hongs, M.; Griffin, R. G.; Herzfeld, J. J Magn Resona 2003, 160, 1–12.
29. Chaykovsky, M. M.; Bae, L. C.; Cheng, M.-C.; Murray, J. H.; Tortolani, K.; Zhang, R.; Seshadri, K.; Findlay, J. H. B. C.; Hsieh S.-Y.; Kalverda, A. P.; Homans, S. W.; Brown, J. M. J Am Chem Soc 2003, 125, 15767–15771.