Strategies for Protein NMR in *Escherichia coli*

Guohua Xu,‡Yansheng Ye,‡Xiaoli Liu,‡ Shufen Cao,‡ Qiong Wu,‡ Kai Cheng,‡§ Maili Liu,‡ Gary J. Pielak,§ and Conggang Li*†

1Key Laboratory of Magnetic Resonance in Biological Systems, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Center for Magnetic Resonance, Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences, Wuhan 430071, P. R. China.
‡Graduate University of Chinese Academy of Sciences, Beijing 100029, P. R. China
§Department of Chemistry and Department of Biochemistry and Biophysics, University of North Carolina—Chapel Hill, Chapel Hill, North Carolina 27599-3290, United States

**ABSTRACT:** In-cell NMR spectroscopy provides insight into protein conformation, dynamics, and function at atomic resolution in living cells. Systematic evaluation of isotopic-labeling strategies is necessary to observe the target protein in the sea of other molecules in the cell. Here, we investigate the detectability, sensitivity, and resolution of in-cell NMR spectra of the globular proteins GB1, ubiquitin, calmodulin, and bcl-xl-cutloop, resulting from uniform 15N enrichment (with and without deuteration), selective 15N-Leu enrichment, 13C-methyl enrichment of isoleucine, leucine, valine, and alanine, fractional 13C enrichment, and 19F labeling. Most of the target proteins can be observed by 19F labeling and 13C enrichment with direct detection because selectively labeling suppresses background signals and because deuteration improves in-cell spectra. Our results demonstrate that the detectability of proteins is determined by weak interactions with intercellular components and that choosing appropriate labeling strategies is critical for the success of in-cell protein NMR studies.

In cells, most proteins function where the concentration of macromolecules is as high as 400 g/L. This crowded environment can affect protein structure, dynamics, and function. Investigating biomolecules in their natural environment is now a part of modern biochemistry, and NMR (nuclear magnetic resonance) spectroscopy is an increasingly popular approach for gaining atom-level information in living cells.

Historically, the term *in vivo* NMR has been used to describe studies of small molecule metabolites, such as ATP and creatine, by means of one-dimensional 1H and 31P spectroscopies because hydrogen and phosphorus nuclei are abundant in organisms. “In-cell protein NMR spectroscopy”, however, focuses on enriching or labeling proteins with NMR-active nuclei such as 15N, 13C, and 19F and using one- and multidimensional NMR experiments to characterize their conformation, dynamics, and interactions. For instance, the high-resolution structure of the 66-residue heavy-metal binding protein TTHA1718 has been determined in living *Escherichia coli*, and the maturation of human superoxide dismutase 1 (SOD1) has been investigated in live human cells.

Enrichment or labeling of the target protein is a prerequisite for in-cell NMR in *E. coli*. To date, several strategies have been investigated. Serber et al. uniformly enriched calmodulin with 13C and NmerA with 15N by overexpression. They observed a high background spectrum from metabolites in the 13C-1H spectrum of calmodulin, suggesting that uniform 13C labeling alone is not a good labeling strategy. Lower backgrounds were observed in the 15N-1H heteronuclear single quantum coherence (HSQC) spectrum of NmerA, suggesting that 15N enrichment may be preferable.

Li et al. conducted in-cell NMR studies of uniformly 15N-enriched and 19F-labeled α-synuclein, chymotrypsin inhibitor 2, ubiquitin, and calmodulin. The 15N-1H HSQC spectra of the three globular proteins, the inhibitor, ubiquitin, and calmodulin, could not be detected in cells; only the disordered protein, α-synuclein, gave a reasonable spectrum. However, 19F resonances of all four proteins were detectable in *E. coli* cells, suggesting 19F labeling is a useful strategy.

Serber et al. conducted in-cell NMR studies of selectively [methyl-13C]-methionine-enriched calmodulin and [methyl-13C]-alanine-enriched NmerA. Their results indicate that selective site-specific enrichment is a useful strategy to minimize background of in-cell spectra. These authors also selectively enriched calmodulin with 15N-lysine and obtained a high-resolution in-cell spectrum, which contradicts the observation of Li et al. from uniform labeling described above.

**Received:** January 19, 2014
**Revised:** March 3, 2014
**Published:** March 5, 2014

Supporting Information
The ubiquitous nature of weak, nonspecific interactions in cells is perhaps the main challenge for multidimensional in-cell NMR. These interactions retard the rotational motion of globular proteins and the exchange dynamics, resulting in severe resonance broadening.

In summary, although enrichment and labeling approaches have made valuable contributions to in-cell protein NMR, some results contradict one another, and their advantages and disadvantages have not been systematically investigated. Here, we evaluate several enrichment and labeling strategies to help guide investigators in selecting the best one for their application. Uniform $^{15}$N enrichment, selective $^{15}$N-Leu enrichment, fractional $^{13}$C enrichment, $^{13}$C-methyl enrichment of Ile, Leu, Val, and Ala methyl groups in a deuterated background (ILVA enrichment), deuteration, and $^{19}$F labeling were evaluated by comparing the detectability, sensitivity, and resolution of in-cell spectra of four model proteins: human ubiquitin (8.5 kDa, pI 6.6), human calmodulin (16.7 kDa, pI 6.8), the B1 domain of protein G (GB1, 6.2 kDa, pI 4.5), and bcl-xl-cutloop (19.4 kDa, pI 4.6). Our results suggest that detectability, sensitivity, and resolution in cells are highly influenced by weak interactions and that choosing an appropriate strategy is critical for obtaining high-quality data.

**MATERIALS AND METHODS**

**Protein Expression.** The plasmids containing the genes for GB1, calmodulin, ubiquitin, and bcl-xl-cutloop were transformed into *E. coli* BL21(DE3) competent cells. Plasmid-containing cells for GB1, ubiquitin, or bcl-xl-cutloop were selected with 100 μg/mL of ampicillin. Plasmid-containing cells for calmodulin were selected with 50 μg/mL of kanamycin. The cells were grown at 37 °C in a rotary shaker at 220 rpm and induced at a final concentration of 1 mM isopropyl β-D-thiogalactopyranoside after reaching an absorbance of 600 nm of 0.6–0.8.

**Enrichment and Labeling.** Reagents were used without further purification. $^{2,13}$C- and $^{[U-^{2}H]}$-glucose, $^{15}$N-enriched NH$_4$Cl, D$_2$O, $^{15}$N-enriched Leu, and precursors for ILVA $^{13}$C-methyl enrichment ({U-[13H]; Ala$^{[13}$CH$_3$]; Ile$^{[13}$CH$_3$; Leu,Val-[13CH$_3$/12CD$_3$]) were from Cambridge Isotope Laboratories. Fluorine-labeled aromatic amino acid analogues were from Sigma-Aldrich.

$^{15}$N-enrichment and $^{19}$F-labeling procedures have been described. Fractional $^{13}$C labeling was achieved as described by using $[2,13]$C-glucose as the sole carbon source in M9 medium. Selective $^{15}$N-Leu labeling was carried out in a manner similar to that described by Muchmore et al. ILVA methyl labeling was achieved as described. Briefly, 1 h prior to inducing protein expression, 800 mg of $\{2-d_3$-$^{13}$C$_1\}$-l-alanine, 2.5 g of succinate-$d_4$, 120 mg of 2-keto-3-methyl-$d_3$-$d_4$-$^{13}$C-butrate, or 60 mg of 2-keto-$d_3$-$d_4$-$^{13}$C-butrate were added to 1 L of D$_2$O-based M9 growth medium containing $[U-^{2}H]$-d-glucose.

**Sample Preparation.** In-cell samples were prepared as described. Briefly, cultures (usually 200 mL) were harvested by centrifugation at 2000g for 10 min at room temperature. The cell pellets were gently resuspended in 0.5 mL of M9 buffer (6.5 g/L of NaHPO$_4$, 3.0 g/L of KH$_2$PO$_4$, and 1.0 g/L of NaCl in 90:10 v/v H$_2$O/D$_2$O) and transferred to a 5 mm tube. Supernatants were collected immediately after each experiment by centrifugation (2000g, 10 min) to assess leakage. To prepare lysate samples, these pellets were resuspended in 1 mL of M9 buffer and sonicated on ice for 20 min with a duty cycle of 3 s on and 5 s off. The lysates were clarified by centrifugation (12 000g for 10 min) prior to NMR.

GB1 was purified as described. Calmodulin was purified by using a Ca$^{2+}$-dependent Phenyl-Sepharose column (GE Life Sciences). His-tagged bcl-xl-cutloop was purified by using a Ni-NTA column. Ubiquitin was purified using a procedure similar to the one previously described. Briefly, cells harvested by centrifugation were resuspended in 50 mM sodium acetate, pH 5.0, followed by the addition of an equal volume of CHCl$_3$ with vortexing. The sample was centrifuged, and the supernatant was loaded onto a Fast Flow SP Sepharose column (GE Life Sciences) and eluted with a linear gradient of 0 to 1 M

![Figure 1. $^{1}$H–$^{15}$N TROSY-HSQC spectra of the uniformly$^{15}$N-enriched proteins in buffer, *E. coli* (BL21(DE3)), and cell lysates. The ubiquitin and calmodulin data have been published.](dx.doi.org/10.1021/bi500079u)
NaCl in the same buffer. Fractions containing ubiquitin were pooled and concentrated. Size-exclusion chromatography on Superdex 75 was used as the final purification step.

NMR. Experiments were performed at 298 K on Bruker 800 and 600 MHz NMR spectrometers equipped with a triple-resonance HCN cryoprobe and a HFCN cryoprobe, respectively. One-dimensional $^{13}$C spectra (256 transients) were acquired with a $90^\circ$ pulse and a 2.0 s relaxation delay. $^1$H−$^{15}$N and $^1$H−$^{13}$C HSQC spectra were collected with 2048 (12 transients) and 256 complex points in the direct and indirect dimensions, respectively. One-dimensional $^{19}$F spectra (up to 2048 transients) were acquired with a 2 s relaxation delay and a sweep width of 32 kHz. $^{19}$F chemical shifts are referenced to trifluorotoluene at $-63.72$ ppm. The data were processed with NMRPipe$^{52}$ and Topspin 2.1.

**RESULTS**

Uniform $^{15}$N Enrichment. Backbone amide groups are valuable probes of protein structure and dynamics. Uniform $^{15}$N enrichment is a standard strategy for solution protein NMR. We expressed four proteins, GB1, calmodulin, ubiquitin, and bcl-xl-cutloop, in *E. coli* grown in media containing $^{15}$N ammonium salts and acquired their in-cell $^1$H−$^{15}$N transverse relaxation optimized spectroscopy ($^1$H−$^{15}$N TROSY-HSQC spectra) (Figure 1). Detectability was assessed by comparing the in-cell spectra to those of the purified proteins in buffer. Cells expressing GB1 yielded a well-resolved in-cell NMR spectrum similar to that of the purified protein and consistent with prior results.$^{22,35}$ The other three proteins did not yield high-quality in-cell spectra. For calmodulin, only strong background signals from small molecule metabolites or highly mobile groups were observed. Ikeya et al. reported that good-quality $^1$H−$^{15}$N HSQC spectra can be obtained in *E. coli* JM109 (DE3) and HMS174 (DE3) cells, whereas cross peaks of rat calmodulin were broadened in BL21(DE3) cells.$^{53}$ We did not observe the spectrum of human calmodulin when the protein was expressed in JM109 (DE3) (Figure S1) or BL21(DE3) (Figure 1).$^{20}$ Ubiquitin was also invisible; only background signals were observed, consistent with prior studies$^{31,35,48}$ but inconsistent with the work of Burz et al.,$^{54,55}$ who observed high-quality spectra by using a protocol involving freezing and then thawing the cells before data acquisition.

In summary, of the four uniformly $^{15}$N-enriched proteins, only GB1 yielded an in-cell NMR spectrum of reasonable resolution. The molecular mass of ubiquitin is only $\sim 2$ kDa larger than that of GB1, suggesting that molecular mass is not
the limiting factor. These observations led us to apply other strategies for obtaining in-cell NMR spectra.

**Uniform $^{15}$N Enrichment with Deuteration.** Replacing hydrogen atoms with deuterons reduces dipolar relaxation, which increases the resolution and sensitivity of the $^1$H–$^{15}$N TROSY-HSQC experiment.\(^{56}\) To investigate the effects of deuteration, we expressed proteins GB1 and calmodulin in BL21(DE3) cells grown in D\(_2\)O-based medium containing [U-2H]-D-glucose and acquired in-cell spectra in M9 buffer made with H\(_2\)O (Figures 2 and 3). The quality of the GB1 spectrum improved, as shown by comparing cross sections. The width at half-height of the amide $^1$H resonance decreased from an average of 60 Hz in the nondeuterated sample to 41 Hz in the deuterated sample.

The impact of deuteration on the in-cell NMR spectrum of calmodulin is remarkable (Figure 3), although the cross peaks remain broad (average \(\sim 63\) Hz). Furthermore, the in-cell spectrum of deuterated calmodulin is similar to that of the Ca\(^{2+}\)-free apo form. During expression, the labile amide protons on calmodulin were replaced by deuterons because expression was carried out in D\(_2\)O-containing medium. However, the cells were harvested and resuspended in H\(_2\)O-containing buffer about 1 h prior to acquiring the in-cell NMR spectra to allow back exchange of hydrogen. The time may need to be optimized on a case-by-case basis to allow sufficient back exchange.

![Figure 4. One-dimensional $^{13}$C and two-dimensional $^1$H–$^{13}$C HSQC spectra of fractionally $^{13}$C-enriched GB1, calmodulin, ubiquitin, and bcl-xl-cutloop in buffer, E. coli cells, and cell lysates.](image-url)
exchange, especially for amide protons in the hydrophobic core. Furthermore, care must be taken when interpreting data from $^{15}$N relaxation experiments in cells if deuteration is used. HSQC-TROSY spectra of ubiquitin and bcl-xl-cutloop could not be obtained in cells even with deuteration (data not shown). In summary, in-cell NMR spectra of deuterated GB1...
and calmodulin were improved by deuteration, consistent with the observations of Serber et al., who applied it to NmerA.

**Fractional ^13C Enrichment.** This approach produces isolated ^13C spins (i.e., the absence of ^13C−^13C one bond scalar couplings) by growth in media containing either [1-^13C]- or [2-^13C]-glucose. The simplified relaxation mechanism of the isolated spins facilitates interpretation of relaxation data. The approach has been employed to probe protein dynamics in solution, and it provides information that is complementary to that provided by ^15N relaxation studies. Importantly, fractional enrichment facilitates dynamics studies of aromatic side chains.

To investigate whether this strategy is suitable for in-cell NMR, we expressed proteins GB1, calmodulin, ubiquitin, and bcl-xl-cutloop in *E. coli* BL21(DE3) cells grown in medium containing [2-^13C]-glucose and acquired one-dimensional ^13C and two-dimensional ^1H−^13C spectra in cells, cell lysates, and buffer (Figure 4). The cells expressing calmodulin, ubiquitin, and bcl-xl-cutloop yielded similar two-dimensional spectra, which are different from the spectra of the purified proteins and the cell lysates. The result suggests that the protein cross peaks either are obscured in or absent from the in-cell spectra. Overlap is especially notable in the aliphatic region. For GB1, a large number of expected cross peaks are observed in the in-cell spectrum. In summary, fractional ^13C labeling, like uniform ^13C labeling, yields strong background signals that overlap with the target protein signals, and neither method is suitable for most in-cell protein NMR studies.

**^15N-Leu Enrichment.** Enriching specific residue types with ^15N is a potential strategy for overcoming resonance broadening. The approach is often used to simplify spectra and to confirm resonance assignments. We expressed GB1, calmodulin, ubiquitin, and bcl-xl-cutloop in *E. coli* BL21(DE3) cells grown in media containing ^15N-Leu and acquired in-cell ^1H−^15N TROSY-HSQC spectra (Figure 5). As expected, ^15N-Leu GB1 yielded a high-quality in-cell spectrum. Cross peaks from its three Leu residues are visible, and there are no obvious background signals, suggesting negligible scrambling of the ^15N-Leu. ^15N-Leu-enriched calmodulin, ubiquitin, and bcl-xl-cutloop, however, failed to give interpretable in-cell spectra, a result similar to what we observed for uniform ^15N enrichment. Our result is inconsistent with that of Serber et al., who observed high-resolution in-cell ^15N−^1H HSQC spectra of ^15N-Lys-enriched calmodulin. In addition, like GB1, the

---

**Figure 7.** Size-exclusion chromatography of *E. coli* cell lysates containing bcl-xl-cutloop at several NaCl concentrations as monitored by SDS-PAGE and absorbance at 280 nm. M and CL stand for protein marker and cell lysate, respectively. Numbers below the SDS-PAGE gel images correspond to the fraction number in the chromatogram. The bcl-xl-cutloop band is indicated by the horizontal arrow.
background in the in-cell spectra of 15N-Leu calmodulin, 
ubiquitin, and bcl-xl-cutloop are low, especially compared to 
the background observed with uniform 15N and 13C fractional 
labeling. In summary, selective 15N enrichment of specific 
residue types can overcome interference from background 
signals when in-cell 15N−1H HSQC spectra are visible.

19F Labeling. 19F NMR is popular for in-cell protein 
NMR for several reasons. The nucleus is 100% abundant, 
has a spin of 1/2 and can be detected with high sensitivity (83% 
of 1H). Furthermore, fluorinated amino acid analogues can be 
incorporated biosynthetically, and spectral background is not a 
problem because natural proteins lack fluorine.

We incorporated the analogues 3-fluorotyrosine (3-FY) and 
3-fluorophenylalanine (3-FF) into the four test proteins and 
acquired one-dimensional in-cell 19F spectra (Figure 6). GB1 
has three tyrosine and two phenylalanine residues. Three 
distinguishable 19F resonances from 3FY GB1 in cells are 
observed, although they are broadened significantly compared to 
the spectrum in buffer. Resonances from two of the three 3-
FY residues partly overlap in cells, and the remaining one 
overlaps with the resonance of free 3-FY. The in-cell spectrum 
of 3-FF-labeled GB1 shows a sharp free amino acid resonance 
and a broad protein peak. This broad peak contains 19F 
resonances from both phenylalanine residues, consistent with 
the spectrum of the purified protein. Compared to the other 
three proteins, the GB1 spectra exhibited the least broadening 
and yielded good-quality in-cell data.

Calmodulin contains two tyrosine residues and eight 
phenylalanine residues. 19F signals from the two 3-FY residues 
overlap partially with the free amino acid resonance in the 
in-cell spectrum. 19F resonances from the eight 3-FF residues yield 
two broad envelopes in cells. The resolution is improved in the 
lysates, but the resonances remained unresolved. Although the 
broadening observed with calmodulin is greater than it is for
GB1, the situation is even worse for ubiquitin and bcl-xl-cutloop.

Ubiquitin contains one tyrosine residue and two phenylalanine residues. The resonance from its 3-FY is severely broadened in cells (~1.5 ppm width at half-height) compared to that observed in lyase and buffer (~0.1 ppm). Considering that the molecular weight of ubiquitin is only ~2 kDa larger than GB1, this observation implies that ubiquitin interacts more strongly than GB1 with other macromolecules in cells. The resonances from 3FF–ubiquitin are also broad, and the two peaks overlap, but the chemical shifts are almost the same as those from the protein in buffer.

Bcl-xl-cutloop contains six tyrosine residues and 10 phenylalanine residues. The resonances from the 3-FY-labeled protein in cells are an extremely broad peak, and resolution is not observed even in the lyase. This is in stark contrast to the narrow resonances observed in buffer. This observation implies that bcl-xl-cutloop strongly interacts with endogenous macromolecules or other large cellular components.

To understand why the bcl-xl-cutloop resonances are so broad in cells, we used size-exclusion chromatography to examine the interaction of the protein with other components in cell lysates, as suggested by Crowley et al.\textsuperscript{22} Without added NaCl, the majority of bcl-xl-cutloop co-eluted with high-molecular-weight proteins and/or complexes (Figure 7), which corresponds to particles with molecular weights exceeding 150 kDa.\textsuperscript{22} At an NaCl concentration of 200 mM, the amount of bcl-xl-cutloop that eluted early decreased, and more eluted at the expected volume. Nevertheless, some bcl-xl-cutloop eluted in fractions corresponding to both high- and low-molecular-weight particles even at 1 M NaCl. These results support our conclusion that bcl-xl-cutloop does not exist in a free form in cells. Most of the protein forms complexes with other macromolecules, and these complexes are at least partially stabilized by attractive charge–charge interactions.

In summary, 3F-GB1 and 3FF-GB1 yield the narrowest in-cell 19F resonances, although resonance envelopes could be observed for the other three proteins. Overall, 19F labeling yielded more useful data than uniform 15N enrichment and fractional 13C enrichment. When a protein contains only one or two aromatic residues and resolution is not problematic, fluorinated aromatic residues are useful for obtaining in-cell NMR spectra. Genetic incorporation of a 19F-labeled non-natural amino acid, such as trifluoromethyl-phenylalanine, is also a useful strategy, especially for large proteins.\textsuperscript{20,24}

### Table 1. Summary of In-Cell NMR Spectral Quality of GB1, Calmoduline(CaM), Ubiquitin (UBQ), and bcl-xl-cutloop, with 15N Enrichment (with and without Deuteration), 13C-Methyl Enrichment of Isoleucine, Leucine, Valine, and Alanine in a Deuterated Background (ILVA Methyl), Selective 15N-Leu Enrichment, Fractional 13C Enrichment, and 19F Labeling\textsuperscript{a}

<table>
<thead>
<tr>
<th>protein</th>
<th>15N</th>
<th>15N, 1H</th>
<th>fractional 13C (1D 13C detection)</th>
<th>fractional 13C (2D)</th>
<th>15N-Leu</th>
<th>19F</th>
<th>ILVA methyl (1D 13C detection)</th>
<th>ILVA methyl (2D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GB1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>UBQ</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CAM</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>bcl-xl-cutloop</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

|      | +, detectable in-cell spectra; −, undetectable in-cell spectra; ■, very low background; □, low background; ■■, medium background; and ■■■, high background. |

\textbf{Methyl Enrichment.} Methyl groups are often located in hydrophobic cores, protein–protein interaction interfaces, ligand binding pockets, and active sites. Methyl-TROSY experiments have been applied to proteins and protein complexes with megadalton molecular weights.\textsuperscript{38} There are potential advantages of methyl enrichment for in-cell NMR. First, methyl resonances have higher signal-to-noise ratios than amide resonances. Second, transverse relaxation is slowed, resulting in increased resolution and sensitivity. Furthermore, the background in 1H–13C correlation spectra of methyl-enriched protein is low when all protons, except those on the methyl groups, are replaced by deuterons.

To determine whether methyl labeling is well-suited for in-cell NMR, methyl groups from Ile, Leu, Val, and Ala residues of the four proteins were enriched using published protocols.\textsuperscript{46,47} Two-dimensional 1H–13C correlation spectra and one-dimensional 13C spectra were acquired for each protein in purified form in buffer, cells, and cell lysates (Figure 8). For Val and Leu methyl groups, the number of resonances and cross peaks observed were larger than the number of residues present because enrichment is not stereospecific.

In-cell spectra of GB1 cells are of reasonable quality, and the chemical shifts are similar to those of the purified protein. Although there are background resonances, the protein resonances can be distinguished because the background signals are narrow, arising as they do from the small molecule precursors.

Methyl-enriched calmodulin exhibited better in-cell spectra than in-cell spectra from 15N, 15N-Leu, and 13C enrichment, although the methyl resonances are broader than those for GB1. The methyl resonances in cells are similar to those of the purified Ca\textsuperscript{2+}-free apoprotein, suggesting that the protein exists mainly in the apo form in cells, in agreement with our 15N–1H TROSY-HSQC data on deuterated calmodulin and the report of Serber et al.\textsuperscript{18}

For ubiquitin and bcl-xl-cutloop, broad peaks were observed in one-dimensional in-cell spectra, and no protein cross peaks could be identified in the two-dimensional spectra, although strong background signals were observed. This observation indicates that methyl enrichment will not improve in-cell spectra for proteins whose resonances are severely broadened by nonspecific interactions.

Specific enrichment of methyl groups combined with one-dimensional direct carbon detection may hold promise for in-cell studies for two reasons. First, one-dimensional experiments require shorter acquisition time compared to two-dimensional
experiments. The time savings can be key for in-cell studies because long acquisitions risk leakage of target proteins. Second, protein cross peaks from in-cell experiments that are undetectable in two-dimensional spectra using $^1$H detection may be visible in simpler one-dimensional $^{13}$C spectra, even when the resonances are broad (e.g., Figure 8).

**DISCUSSION**

Four proteins, ubiquitin, calmodulin, GB1, and bcl-xl-cutloop, were used to evaluate commonly used isotopic enrichment and labeling schemes for in-cell protein NMR. Spectra were acquired in cells, cell lysates, and buffer. Supernatants were collected and analyzed after acquisition of in-cell data to confirm that the test protein did not leak during the experiment. The quality of the in-cell spectra for the four proteins is summarized in Table 1.

Our results suggest that fractional $^{13}$C enrichment and uniform $^{13}$C enrichment are less suitable because they result in a high background. $^{15}$N enrichment would seem to be more useful because it produces a less intense background. Deuterium reduces the width of amide $^1$H resonances (by ~20 Hz for GB1) and improves the quality of in-cell spectra. $^{13}$C methyl-TROSY has sufficient sensitivity to allow solution studies on quite large systems, such as the 670 kDa 20S proteasome, but the quality of in-cell spectra from ILVA methyl enrichment was not as high as expected considering the improvement seen in buffer, possibly because of chemical exchange with other intracellular constituents. One-dimensional $^{13}$C direct detection of ILVA methyl-enriched proteins in cells can be useful when cross peaks in two-dimensional spectrum cannot be observed. $^{19}$F labeling is highly recommended for in-cell studies when other approaches fail. Additionally, $^{19}$F labeling is less expensive than ILVA methyl enrichment.

On the basis of these observations, we propose a strategy for selecting the most appropriate methods for protein in-cell NMR studies in *E. coli* that saves both time and expense. Uniform $^{15}$N enrichment should be tried first to assess detectability. If uniform $^{15}$N enrichment yields in-cell $^1$H–$^{13}$N HSQC spectra of reasonable quality (e.g., GB1), then selective $^{15}$N enrichment, $^{19}$F, methyl $^{13}$C enrichment, and ILVA methyl enrichment in a deuterated background will probably yield high-quality in-cell spectra. However, if uniform $^{15}$N enrichment fails (i.e., resonances are invisible in cells, as seen, for example, for calmodulin, ubiquitin, and bcl-xl-cutloop), then selective $^{15}$N enrichment will probably not lead to high-quality data. When faced with this failure, $^{19}$F labeling, which costs less than ILVA methyl enrichment, is encouraged. If $^{19}$F yields severely broadened in-cell spectra (e.g., ubiquitin, bcl-xl-cutloop), then ILVA methyl enrichment will also probably yield broad and uninformative one-dimensional $^{13}$C spectra and $^{13}$C–$^1$H correlation spectra. If the broadening of $^{19}$F resonances is moderate (e.g., calmodulin), then ILVA methyl enrichment is worth trying. If $^{19}$F resonances are severely broadened (e.g., bcl-xl-cutloop), then the test protein has strong interactions with other biomolecules in cells, and high-resolution NMR studies will be challenging if not impossible. In summary, uniform $^{15}$N and $^{19}$F labeling should be tried first. If successful, other more complicated and expensive enrichment strategies are worth pursuing.

**CONCLUSIONS**

We conducted a systematic investigation of the applicability of uniform $^{15}$N enrichment, selective $^{15}$N-Leu enrichment, deuteration, $^{19}$F labeling, ILVA $^{13}$C methyl enrichment, and fractional $^{13}$C enrichment for protein in-cell NMR. Most of the target proteins can be observed by $^{19}$F or $^{13}$C direct detection. Selective enrichment can suppress background signals, and deuteration improves spectral quality. On the basis of our data, we proposed an economical strategy for selecting the appropriate enrichment or labeling method for protein in-cell NMR studies.

**ASSOCIATED CONTENT**

Supporting Information
In-cell spectra of CaM in *E. coli* JM109(DE3), and enlarged versions of Figures 1, 4, 5, and 8. This material is available free of charge via the Internet at http://pubs.acs.org.

**AUTHOR INFORMATION**

Corresponding Author
E-mail: conggangli@wipm.ac.cn.

**Funding**
This work was supported by the Ministry of Science and Technology of China (grant no. 2013CB910200 to C.L.), the 1000 Young Talents Program (C.L.), the National Natural Sciences Foundation of China (grant nos. 21075134 and 21173258 to C.L.), and National Science Foundation (grant no. MCB-1051819 to G.J.P.).

**Notes**
The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

We thank Prof. Bin Xia for providing the bcl-xl-cutloop plasmid and Elizabeth Pielak for helpful comments.

**REFERENCES**


