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*,[†]

[†]Key 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

S *Supporting Information*

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 ^{15}N enrichment (with and without deuteration), selective ^{15}N-Leu enrichment, ^{13}C-methyl enrichment of isoleucine, leucine, valine, and alanine, fractional ^{13}C enrichment, and ^{19}F labeling. Most of the target proteins can be observed by ^{19}F labeling and ^{13}C enrichment

Labeling   In-cell NMR

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.
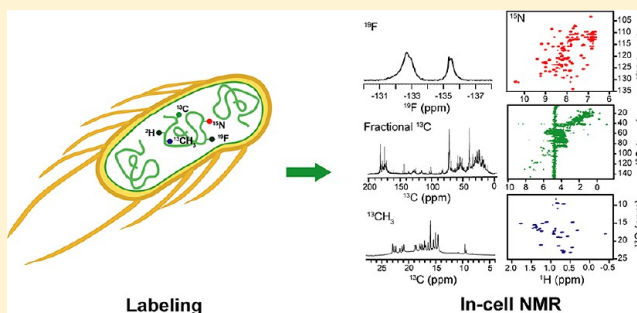
In cells, most proteins function where the concentration of macromolecules is as high as 400 g/L.[1,2] This crowded environment can affect protein structure, dynamics, and function.[3-7] Investigating biomolecules in their natural environment is now a part of modern biochemistry, and NMR (nuclear magnetic resonance) spectroscopy is an increasingly popular approach for gaining atom-level information in living cells.[8-24]

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 ^{31}P spectroscopies because hydrogen and phosphorus nuclei are abundant in organisms.[25,26] "In-cell protein NMR spectroscopy", however, focuses on enriching or labeling proteins with NMR-active nuclei such as ^{15}N, ^{13}C, and ^{19}F and using one- and multidimensional NMR experiments to characterize their conformation, dynamics, and interactions.[27,28] For instance, the high-resolution structure of the 66-residue heavy-metal binding protein TTHA1718 has been determined in living *Escherichia coli*,[17] and the maturation of human superoxide dismutase 1 (SOD1) has been investigated in live human cells.[29]

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 ^{13}C and NmerA with ^{15}N by overexpression. They observed a high background spectrum from metabolites in the $^{13}C-^1H$

spectrum of calmodulin, suggesting that uniform ^{13}C labeling alone is not a good labeling strategy. Lower backgrounds were observed in the $^{15}N-^1H$ heteronuclear single quantum coherence (HSQC) spectrum of NmerA, suggesting that ^{15}N enrichment may be preferable.[18,30]

Li et al. conducted in-cell NMR studies of uniformly ^{15}N-enriched and ^{19}F-labeled α-synuclein, chymotrypsin inhibitor 2, ubiquitin, and calmodulin. The $^{15}N-^1H$ 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, ^{19}F resonances of all four proteins were detectable in *E. coli* cells, suggesting ^{19}F labeling is a useful strategy.[30,31]

Serber et al. conducted in-cell NMR studies of selectively [methyl-^{13}C]-methionine-enriched calmodulin and [methyl-^{13}C]-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 ^{15}N-lysine and obtained a high-resolution in-cell spectrum,[18,30] which contradicts the observation of Li et al.[31] from uniform labeling described above.
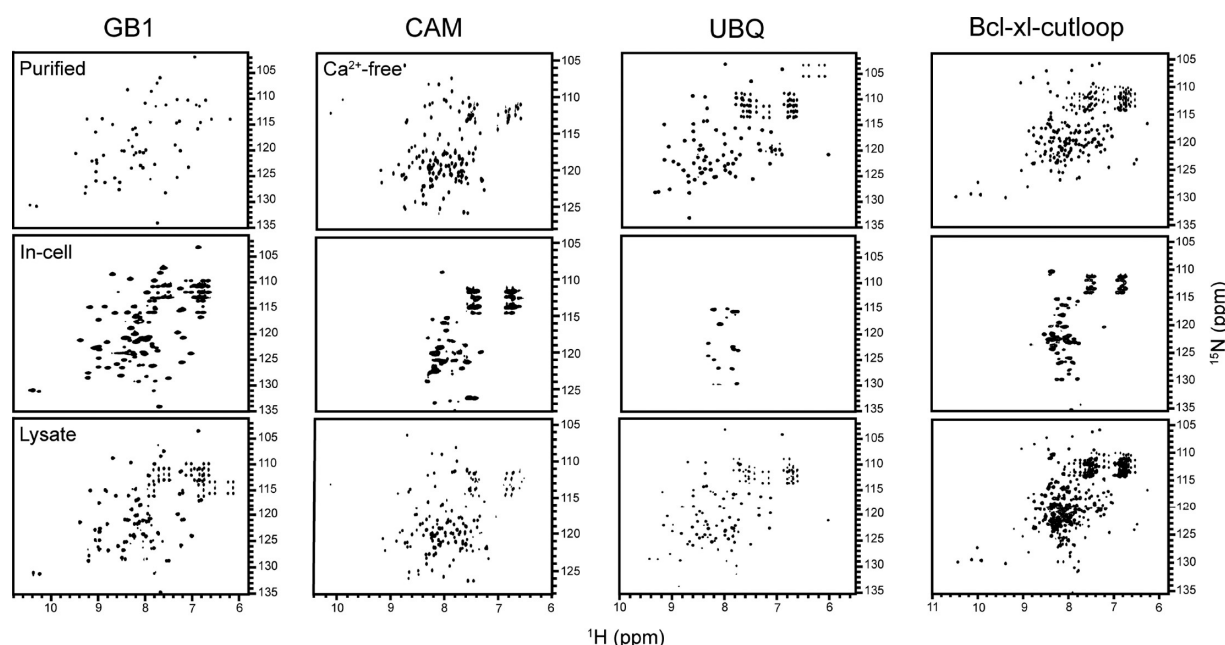
Figure 1. $^1H-^{15}N$ TROSY-HSQC spectra of the uniformily ^{15}N-enriched proteins in buffer, *E. coli* [BL21(DE3)], and cell lysates. The ubiquitin and calmodulin data have been published.[20]

The ubiquitous nature of weak, nonspecific interactions in cells[32−34] 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.[13,22,23,35−38]

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),[39] human calmodulin (16.7 kDa, pI 4.1),[40] the B1 domain of protein G (GB1, 6.2 kDa, pI 4.5),[41] and bcl-xl-cutloop (19.4 kDa, pI 4.6).[42] 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 at 600 nm of 0.6−0.8.

Enrichment and Labeling. Reagents were used without further purification. [2-^{13}C]- and [U−^2H]-glucose, ^{15}N-enriched NH_4Cl, D_2O, ^{15}N-enriched Leu, and precursors for ILVA ^{13}C-methyl enrichment ({U-[^2H]; Ala$^\beta$-[$^{13}CH_3$]; Ile$^{\delta 1}$-[$^{13}CH_3$]; Leu,Val-[$^{13}CH_3/^{12}CD_3$]}) 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.[31] Fractional ^{13}C labeling was achieved as described[43,44] 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.[45] ILVA methyl labeling was achieved as described.[46,47] Briefly, 1 h prior to inducing protein expression, 800 mg of {2-d_1,3-$^{13}C_1$}-L-alanine, 2.5 g of succinate-d_4, 120 mg of 2-keto-3-methyl-d_3-3-d_1-4-^{13}C-butyrate, or 60 mg of 2-keto-3-d_2-4-^{13}C-butyrate were added to 1 L of D_2O-based M9 growth medium containing [U-^2H]-D-glucose.

Sample Preparation. In-cell samples were prepared as described.[31] Briefly, cultures (usually 200 mL) were harvested by centrifugation at 2000*g* for 10 min at room temperature. The cell pellets were gently resuspended in 0.5 mL of M9 buffer (6.5 g/L of Na_2HPO_4, 3.0 g/L of KH_2PO_4, and 1.0 g/L of NaCl in 90:10 v/v H_2O/D_2O) and transferred to a 5 mm tube. Supernatants were collected immediately after each experiment by centrifugation (2000*g*, 10 min) to assess leakage.[48] 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 (12 000*g* for 10 min) prior to NMR.

GB1 was purified as described.[49] Calmodulin was purified by using a Ca^{2+}-dependent Phenyl-Sepharose column (GE Life Sciences).[50] 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.[51] 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
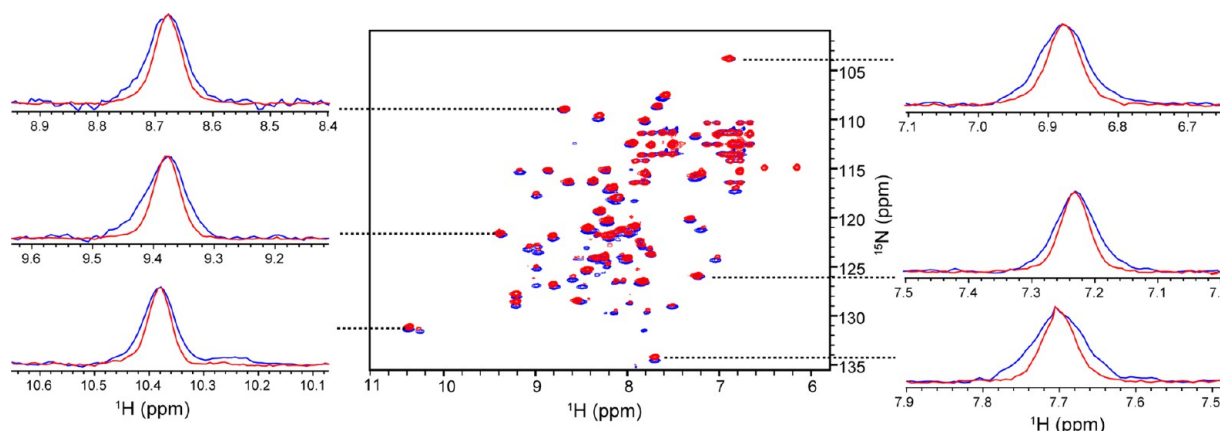
Figure 2. In-cell spectra of ^{15}N-enriched GB1 with (red) and without deuteration. Cross sections from the position indicated by the dotted lines are shown.
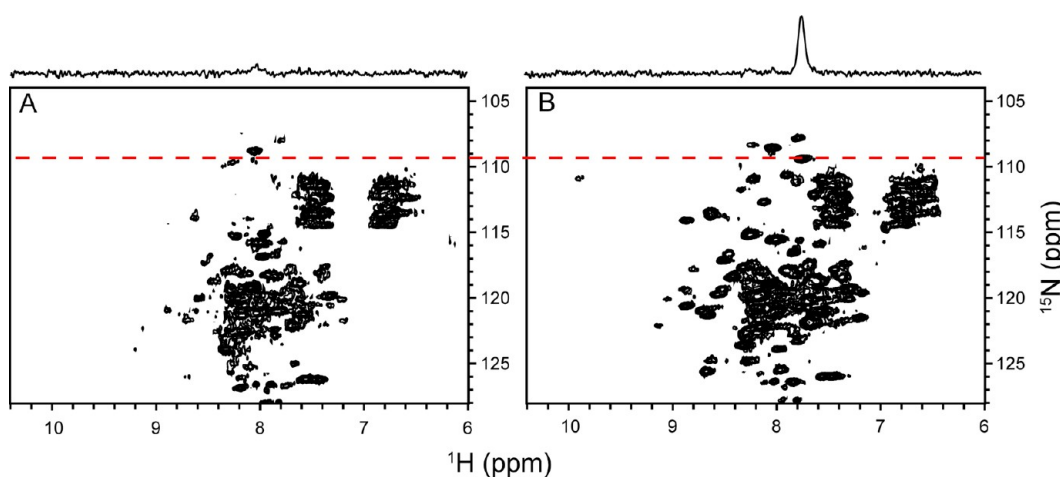

Figure 3. In-cell spectra of ^{15}N-enriched calmodulin without (A) and with (B) deuteration. One-dimensional cross sections taken at the position indicated by the dashed line are shown above the corresponding 2D spectrum. Both spectra comprise 12 transients and are plotted at equal contour levels.

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° 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 ∼2 kDa larger than that of GB1, suggesting that molecular mass is not
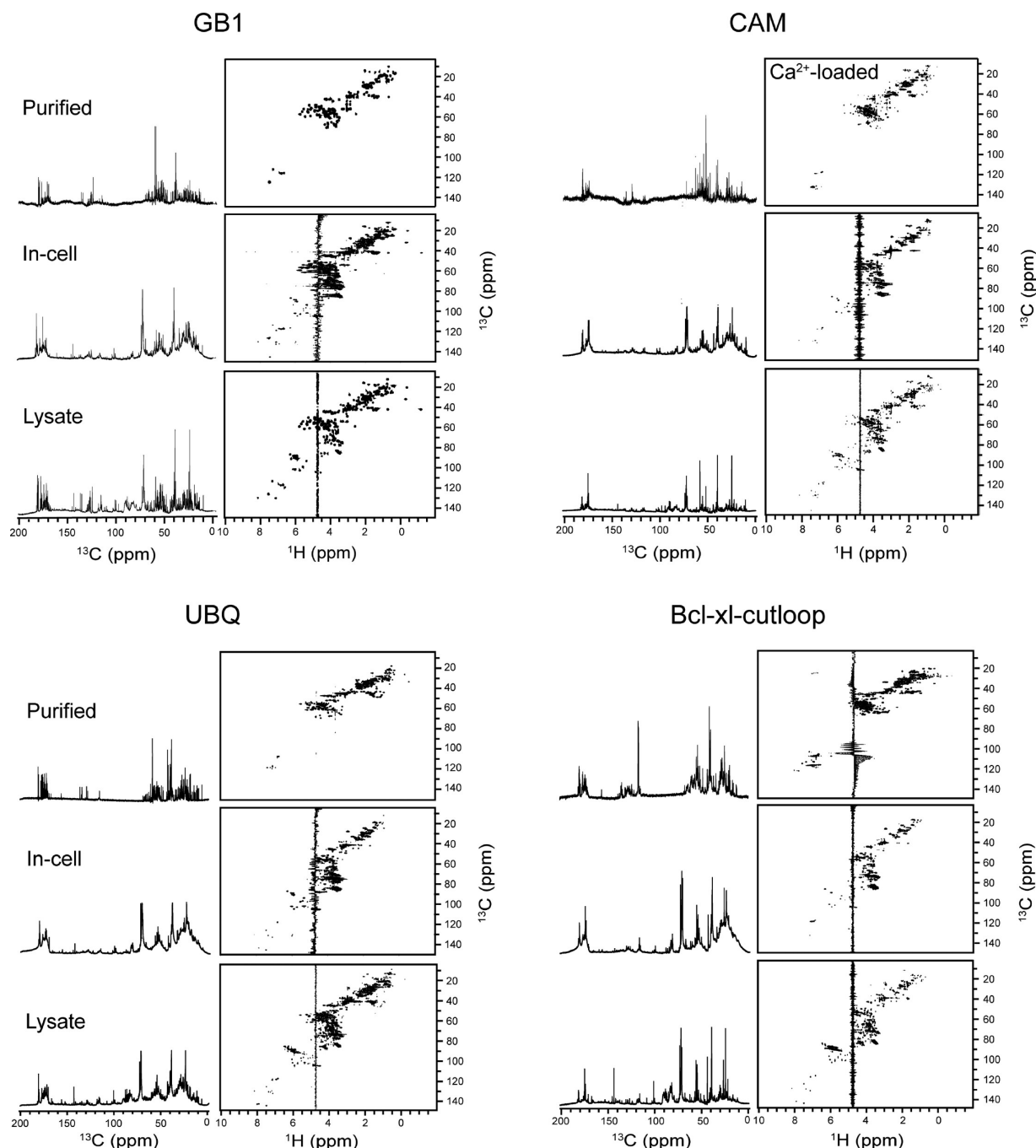
Figure 4. One-dimensional ^{13}C and two-dimensional $^{1}H-^{13}C$ HSQC spectra of fractionally ^{13}C-enriched GB1, calmodulin, ubiquitin, and bcl-xl-cutloop in buffer, *E. coli* cells, and cell lysates.

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_2O-based medium containing [U-^{2}H]-D-glucose and acquired in-cell spectra in M9 buffer made with H_2O (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 ~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_2O-containing medium. However, the cells were harvested and resuspended in H_2O-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
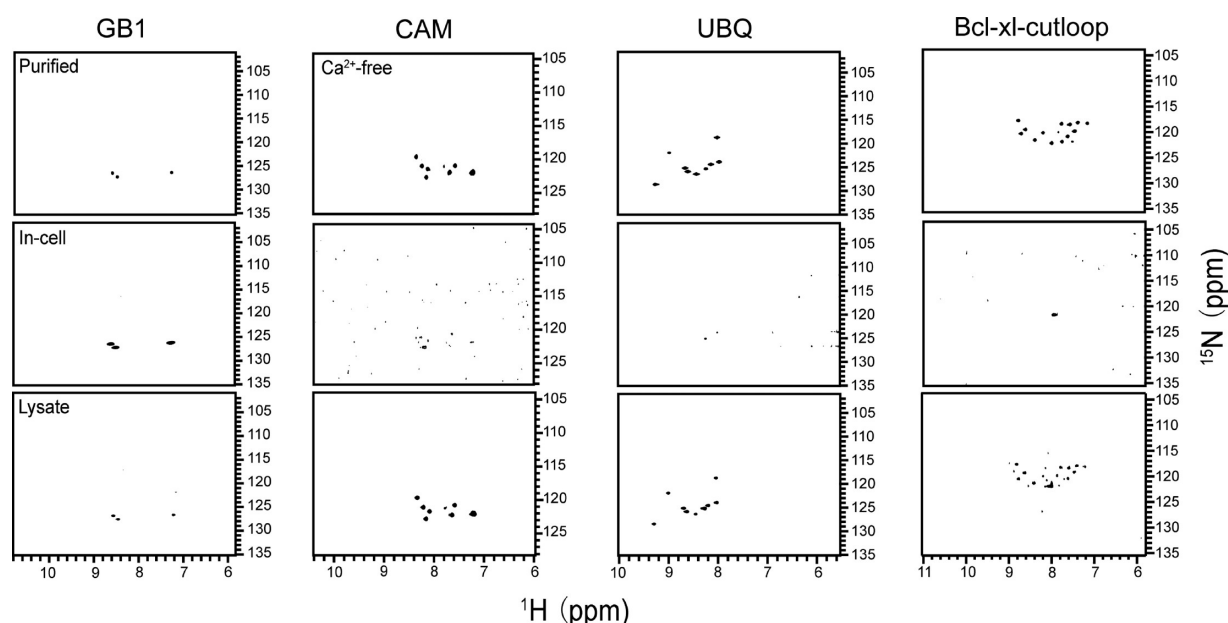
Figure 5. $^1H-^{15}N$ TROSY-HSQC spectra of GB1, calmodulin, ubiquitin, and bcl-xl-cutloop selectively enriched with ^{15}N-Leu in buffer, *E. coli* cells, and cell lysates.
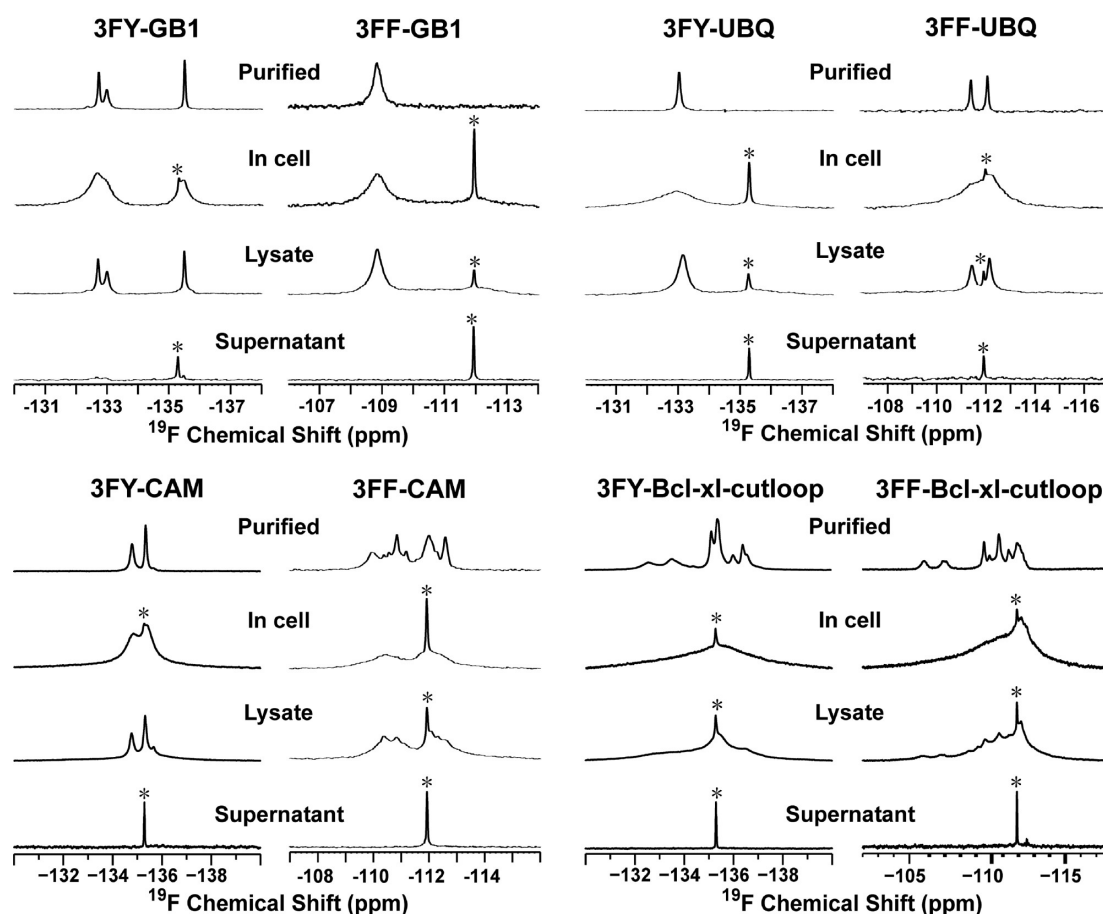

Figure 6. Spectra of 3-fluorotyrosine (3FY)- and 3-fluorophenylalanine (3FF)-labeled GB1, calmodulin (CAM), ubiquitin (UBQ), and bcl-xl-cutloop in buffer, *E. coli* cells, and cell lysates. The spectra from the supernatant of the in-cell experiments are also shown. The asterisk indicates the resonance from the free labeled amino acids. The 3FY GB1, ubiquitine and calmodulin spectra have been published.[20,23]

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
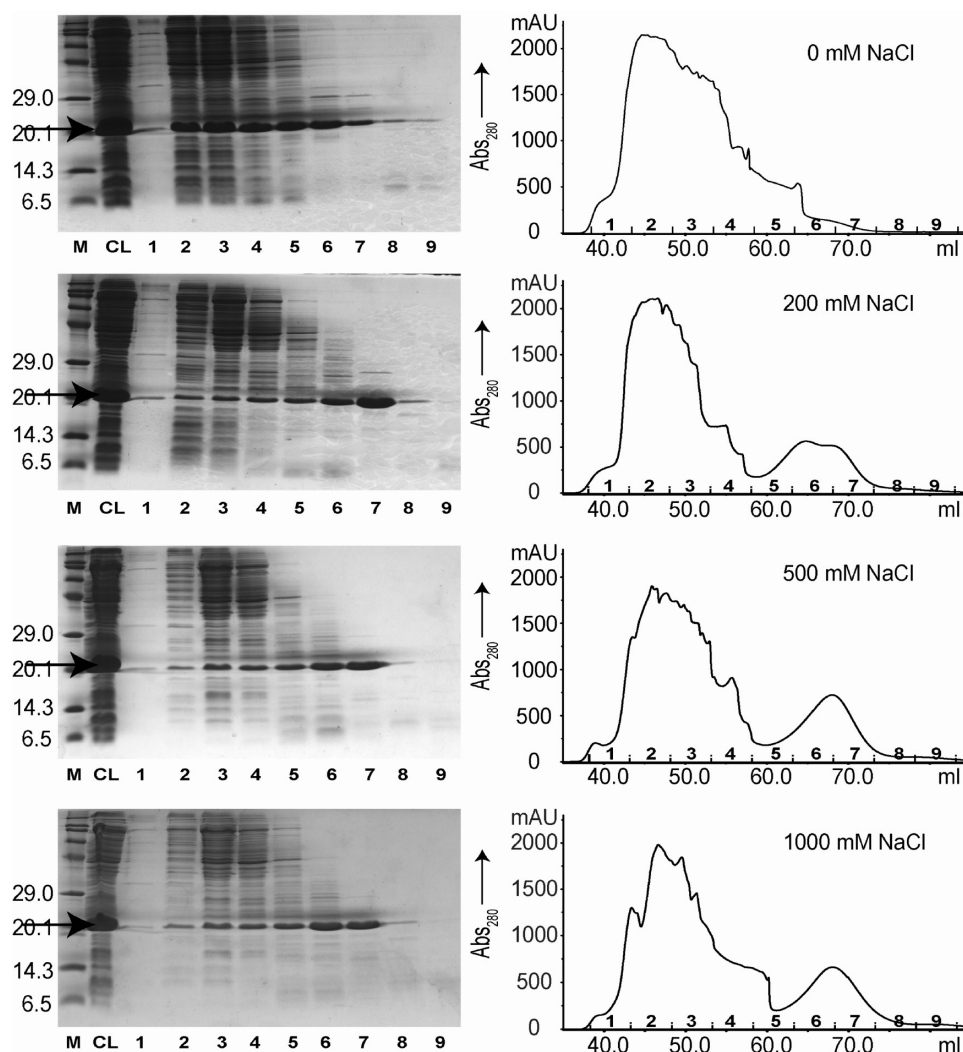
Figure 7. Size-exclusion chromatography of *E. coli* cell lysates containing bcl-xl-cutloop at several NaCl concentrations as monitored by SDS-PAGE and absorbance at 280 nm. M and CL stand for protein marker and cell lysate, respectively. Numbers below the SDS-PAGE gel images correspond to the fraction number in the chromatogram. The bcl-xl-cutloop band is indicated by the horizontal arrow.

and calmodulin were improved by deuteration, consistent with the observations of Serber et al.,[30] who applied it to NmerA.

Fractional [13]C Enrichment. This approach produces isolated ^{13}C spins (i.e., the absence of [13]C−[13]C one bond scalar couplings) by growth in media containing either [1-^{13}C]- or [2-^{13}C]-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 ^{15}N relaxation studies.[44] Importantly, fractional enrichment facilitates dynamics studies of aromatic side chains.[43,57]

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-^{13}C]-glucose and acquired one-dimensional ^{13}C and two-dimensional [1]H−[13]C 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 ^{13}C labeling, like uniform ^{13}C labeling,[18] yields strong background signals that overlap with the target protein signals, and neither method is suitable for most in-cell protein NMR studies.

[15]N-Leu Enrichment. Enriching specific residue types with ^{15}N 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 ^{15}N-Leu and acquired in-cell [1]H−^{15}N TROSY-HSQC spectra (Figure 5). As expected, ^{15}N-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 ^{15}N-Leu. ^{15}N-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 ^{15}N enrichment. Our result is inconsistent with that of Serber et al., who observed high-resolution in-cell [15]N−^{1}H HSQC spectra of ^{15}N-Lys-enriched calmodulin.[18,30,51] In addition, like GB1, the
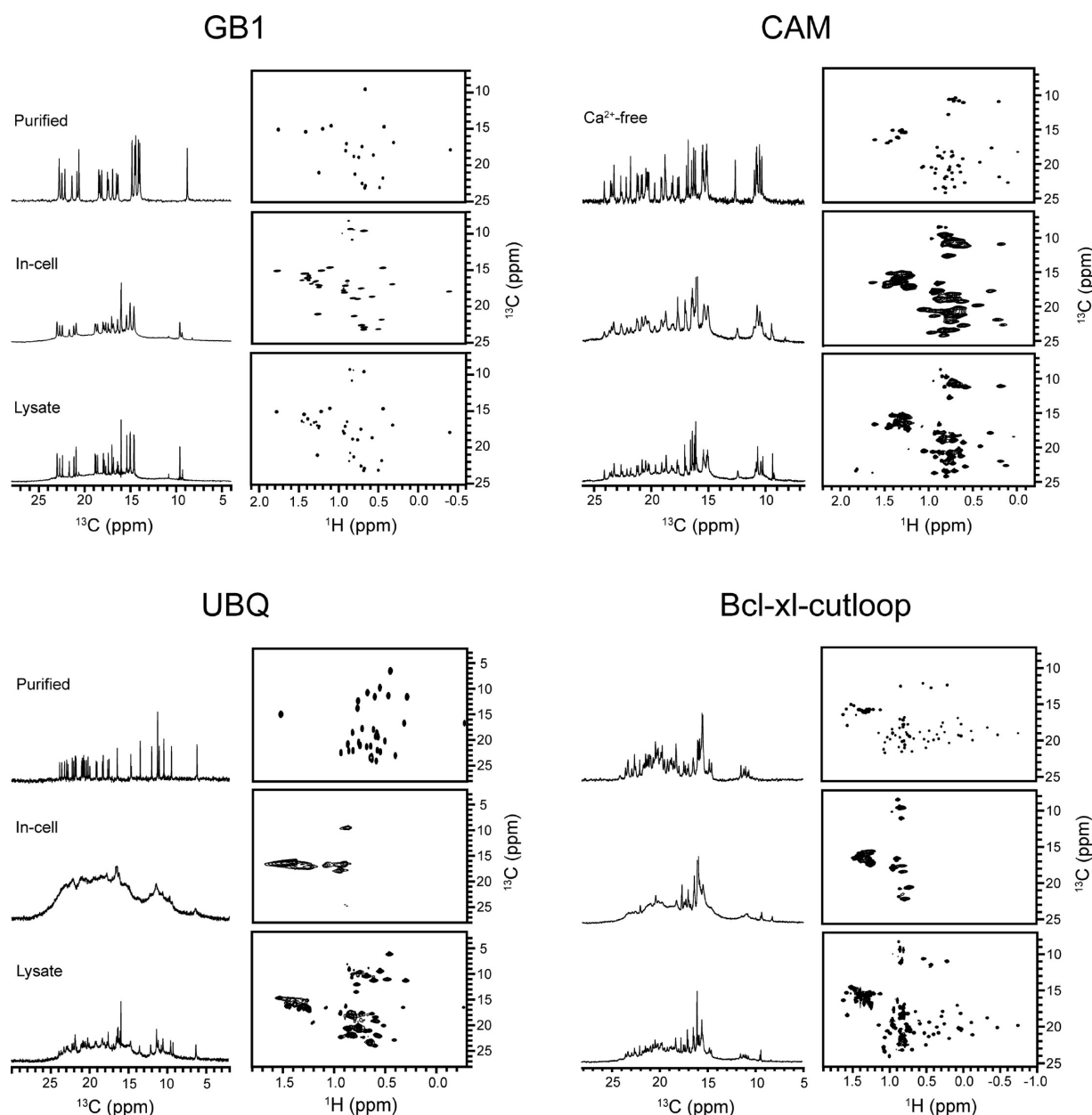
Figure 8. One-dimensional ^{13}C and two-dimensional $^{1}H-^{13}C$ TROSY spectra of the ILVA methyl enriched ($\{U-[^{2}H]; Ala^{\beta}-[^{13}CH_{3}]; Ile^{\delta 1}-[^{13}CH_{3}]; Leu,Val-[^{13}CH_{3}/^{12}CD_{3}]\}$) proteins in buffer, *E. coli*, and cell lysates.

background in the in-cell spectra of ^{15}N-Leu calmodulin, ubiquitin, and bcl-xl-cutloop are low, especially compared to the background observed with uniform ^{15}N and ^{13}C fractional labeling. In summary, selective ^{15}N enrichment of specific residue types can overcome interference from background signals when in-cell $^{15}N-^{1}H$ HSQC spectra are visible.

^{19}F Labeling. ^{19}F NMR is popular for in-cell protein NMR[20,23,24] for several reasons. The nucleus is 100% abundant, has a spin of $^{1}/_{2}$, and can be detected with high sensitivity (83% of ^{1}H). 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 ^{19}F spectra (Figure 6). GB1 has three tyrosine and two phenylalanine residues. Three distinguishable ^{19}F 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 ^{19}F 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. ^{19}F signals from the two 3-FY residues overlap partially with the free amino acid resonance in the in-cell spectrum. ^{19}F 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

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

protein	[15]N	[15]N, [2]H	fractional [13]C (2D)	fractional [13]C (1D ^{13}C detection)	^{15}N-Leu	[19]F	ILVA methyl (2D)	ILVA methyl (1D ^{13}C detection)
GB1	+	+	+	+	+	+	+	+
	■■	■■	■■■	■■■	■	□	■■	■■
UBQ	−	−	−	+	−	+	−	+
	■■	■■	■■■	■■■	■	□	■■	■■
CAM	−	+	−	+	−	+	+	+
	■■■	■■■	■■■	■■■	■	□	■■	■■
bcl-xl-cutloop	−	−	−	+	−	+	−	+
	■■■	■■■	■■■	■■■	■	□	■■	■■

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

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 lysate 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 lysate. 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.[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.[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, 3FY-GB1 and 3FF-GB1 yield the narrowest in-cell ^{19}F resonances, although resonance envelopes could be observed for the other three proteins. Overall, ^{19}F labeling yielded more useful data than uniform ^{15}N enrichment and fractional ^{13}C 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 ^{19}F-labeled non-natural amino acid, such as trifluoromethyl-phenylalanine, is also a useful strategy, especially for large proteins.[20,24]

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.[58] 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 $^1H-^{13}C$ 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.[46,47] Two-dimensional $^1H-^{13}C$ correlation spectra and one-dimensional ^{13}C 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 ^{15}N, ^{15}N-Leu, and ^{13}C enrichment, although the methyl resonances are broader than those for GB1. The methyl resonances in cells are similar to those of the purified Ca^{2+}-free apoprotein, suggesting that the protein exists mainly in the apo form in cells, in agreement with our $^{15}N-^{1}H$ TROSY-HSQC data on deuterated calmodulin and the report of Serber et al.[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. Deuteration 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,[58] 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.[37,38] 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−^{15}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.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Zimmerman, S. B., and Trach, S. O. (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*. *J. Mol. Biol. 222*, 599−620.

(2) Luby-Phelps, K. (1999) Cytoarchitecture and physical properties of cytoplasm: Volume, viscosity, diffusion, intracellular surface area. *Int. Rev. Cytol. 192*, 189−221.

(3) Hatters, D. M., Minton, A. P., and Howlett, G. J. (2002) Macromolecular crowding accelerates amyloid formation by human apolipoprotein C-II. *J. Biol. Chem. 277*, 7824−7830.

(4) Dong, H., Qin, S., and Zhou, H. (2010) Effects of macro-molecular crowding on protein conformational changes. *PLoS Comput. Biol. 6*, e1000833-1−e1000833-10.

(5) Hong, J., and Gierasch, L. M. (2010) Macromolecular crowding remodels the energy landscape of a protein by favoring a more compact unfolded state. *J. Am. Chem. Soc. 132*, 10445−10452.

(6) Wang, Y., Sarkar, M., Smith, A. E., Krois, A. S., and Pielak, G. J. (2012) Macromolecular crowding and protein stability. *J. Am. Chem. Soc. 134*, 16614−16618.

(7) Stagg, L., Zhang, S., Cheung, M. S., and Wittung-Stafshede, P. (2007) Molecular crowding enhances native structure and stability of α/β protein flavodoxin. *Proc. Natl. Acad. Sci. U.S.A. 104*, 18976−18981.

(8) Inomata, K., Ohno, A., Tochio, H., Isogai, S., Tenno, T., Nakase, I., Takeuchi, T., Futaki, S., Ito, Y., and Hiroaki, H. (2009) High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. *Nature 458*, 106−109.

(9) Burz, D. S., Dutta, K., Cowburn, D., and Shekhtman, A. (2006) In-cell NMR for protein-protein interactions (STINT-NMR). *Nat. Protoc. 1*, 146−152.

(10) Serber, Z., Selenko, P., Hänsel, R., Reckel, S., Löhr, F., Ferrell, J. E., Wagner, G., and Dötsch, V. (2007) Investigating macromolecules inside cultured and injected cells by in-cell NMR spectroscopy. *Nat. Protoc. 1*, 2701–2709.

(11) Serber, Z., Keatinge-Clay, A. T., Ledwidge, R., Kelly, A. E., Miller, S. M., and Dötsch, V. (2001) High-resolution macromolecular NMR spectroscopy inside living cells. *J. Am. Chem. Soc. 123*, 2446–2447.

(12) Sakai, T., Tochio, H., Tenno, T., Ito, Y., Kokubo, T., Hiroaki, H., and Shirakawa, M. (2006) In-cell NMR spectroscopy of proteins inside *Xenopus laevis* oocytes. *J. Biomol. NMR 36*, 179–188.

(13) Li, C., Charlton, L. M., Lakkavaram, A., Seagle, C., Wang, G., Young, G. B., Macdonald, J. M., and Pielak, G. J. (2008) Differential dynamical effects of macromolecular crowding on an intrinsically disordered protein and a globular protein: Implications for in-cell NMR spectroscopy. *J. Am. Chem. Soc. 130*, 6310–6311.

(14) Hamatsu, J., O'Donovan, D., Tanaka, T., Shirai, T., Hourai, Y., Mikawa, T., Ikeya, T., Mishima, M., Boucher, W., and Smith, B. O. (2013) High-resolution heteronuclear multidimensional NMR of proteins in living insect cells using a baculovirus protein expression system. *J. Am. Chem. Soc. 135*, 1688–1691.

(15) Bertrand, K., Reverdatto, S., Burz, D. S., Zitomer, R., and Shekhtman, A. (2012) Structure of proteins in eukaryotic compartments. *J. Am. Chem. Soc. 134*, 12798–12806.

(16) Dedmon, M. M., Patel, C. N., Young, G. B., and Pielak, G. J. (2002) FlgM gains structure in living cells. *Proc. Natl. Acad. Sci. U.S.A. 99*, 12681–12684.

(17) Sakakibara, D., Sasaki, A., Ikeya, T., Hamatsu, J., Hanashima, T., Mishima, M., Yoshimasu, M., Hayashi, N., Mikawa, T., and Wälchli, M. (2009) Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature 458*, 102–105.

(18) Serber, Z., Straub, W., Corsini, L., Nomura, A. M., Shimba, N., Craik, C. S., Ortiz de Montellano, P., and Dötsch, V. (2004) Methyl groups as probes for proteins and complexes in in-cell NMR experiments. *J. Am. Chem. Soc. 126*, 7119–7125.

(19) Selenko, P., Serber, Z., Gadea, B., Ruderman, J., and Wagner, G. (2006) Quantitative NMR analysis of the protein G B1 domain in *Xenopus laevis* egg extracts and intact oocytes. *Proc. Natl. Acad. Sci. U.S.A. 103*, 11904–11909.

(20) Li, C., Wang, G., Wang, Y., Creager-Allen, R., Lutz, E. A., Scronce, H., Slade, K. M., Ruf, R. A., Mehl, R. A., and Pielak, G. J. (2010) Protein ^{19}F NMR in *Escherichia coli*. *J. Am. Chem. Soc. 132*, 321–327.

(21) Crowley, P. B., Kyne, C., and Monteith, W. B. (2012) Simple and inexpensive incorporation of ^{19}F-tryptophan for protein NMR spectroscopy. *Chem. Commun. 48*, 10681–10683.

(22) Crowley, P. B., Chow, E., and Papkovskaia, T. (2011) Protein interactions in the *Escherichia coli* cytosol: An impediment to in-cell NMR spectroscopy. *ChemBioChem 12*, 1043–1048.

(23) Ye, Y., Liu, X., Zhang, Z., Wu, Q., Jiang, B., Jiang, L., Zhang, X., Liu, M., Pielak, G. J., and Li, C. (2013) ^{19}F NMR spectroscopy as a probe of cytoplasmic viscosity and weak protein interactions in living cells. *Chem.—Eur. J. 19*, 12705–12710.

(24) Jackson, J. C., Hammill, J. T., and Mehl, R. A. (2007) Site-specific incorporation of a ^{19}F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity. *J. Am. Chem. Soc. 129*, 1160–1166.

(25) Cohen, J. S., Lyon, R. C., and Daly, P. F. (1989) Monitoring intracellular metabolism by nuclear magnetic resonance. *Methods Enzymol. 177*, 435–452.

(26) Szwergold, B. S. (1992) NMR spectroscopy of cells. *Annu. Rev. Physiol. 54*, 775–798.

(27) Serber, Z., and Dötsch, V. (2001) In-cell NMR spectroscopy. *Biochemistry 40*, 14317–14323.

(28) Hänsel, R., Foldynová-Trantírková, S., Löhr, F., Buck, J., Bongartz, E., Bamberg, E., Schwalbe, H., Dötsch, V., and Trantírek, L. (2009) Evaluation of parameters critical for observing nucleic acids inside living *Xenopus laevis* oocytes by in-cell NMR spectroscopy. *J. Am. Chem. Soc. 131*, 15761–15768.

(29) Banci, L., Barbieri, L., Bertini, I., Luchinat, E., Secci, E., Zhao, Y., and Aricescu, A. R. (2013) Atomic-resolution monitoring of protein maturation in live human cells by NMR. *Nat. Chem. Biol. 9*, 297–299.

(30) Serber, Z., Ledwidge, R., Miller, S. M., and Dötsch, V. (2001) Evaluation of parameters critical to observing proteins inside living *Escherichia coli* by in-cell NMR spectroscopy. *J. Am. Chem. Soc. 123*, 8895–8901.

(31) Li, C., Lutz, E. A., Slade, K. M., Ruf, R. A., Wang, G., and Pielak, G. J. (2009) ^{19}F NMR studies of α-synuclein conformation and fibrillation. *Biochemistry 48*, 8578–8584.

(32) Sarkar, M., Li, C., and Pielak, G. J. (2013) Soft interactions and crowding. *Biophys. Rev. 5*, 187–194.

(33) Sarkar, M., Smith, A. E., and Pielak, G. J. (2013) Impact of reconstituted cytosol on protein stability. *Proc. Natl. Acad. Sci. U.S.A. 110*, 19342–19347.

(34) Li, C., and Pielak, G. J. (2009) Using NMR to distinguish viscosity effects from nonspecific protein binding under crowded conditions. *J. Am. Chem. Soc. 131*, 1368–1369.

(35) Wang, Q., Zhuravleva, A., and Gierasch, L. M. (2011) Exploring weak, transient protein—protein interactions in crowded *in vivo* environments by in-cell nuclear magnetic resonance spectroscopy. *Biochemistry 50*, 9225–9236.

(36) Wang, Y., Li, C., and Pielak, G. J. (2010) Effects of proteins on protein diffusion. *J. Am. Chem. Soc. 132*, 9392–9397.

(37) Latham, M. P., and Kay, L. E. (2012) Is buffer a good proxy for a crowded cell-like environment? A comparative NMR study of calmodulin side-chain dynamics in buffer and *E. coli* lysate. *PLoS One 7*, e48226-1–e48226-13.

(38) Latham, M. P., and Kay, L. E. (2013) Probing non-specific interactions of Ca^{2+}-calmodulin in *E. coli* lysate. *J. Biomol. NMR 55*, 239–247.

(39) Hochstrasser, M. (2009) Origin and function of ubiquitin-like proteins. *Nature 458*, 422–429.

(40) Mal, T. K., and Ikura, M. (2006) NMR investigation of calmodulin, In *Modern Magnetic Resonance* (Webb, G. A., Ed.) pp 503–516, Springer, Dordrecht, The Netherlands.

(41) Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T., and Clore, G. (1991) A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science 253*, 657–661.

(42) Wysoczanski, P., Mart, R. J., Loveridge, E. J., Williams, C., Whittaker, S. B., Crump, M. P., and Allemann, R. K. (2012) NMR solution structure of a photoswitchable apoptosis activating Bak peptide bound to Bcl-x$_L$. *J. Am. Chem. Soc. 134*, 7644–7647.

(43) Teilum, K., Brath, U., Lundström, P., and Akke, M. (2006) Biosynthetic ^{13}C labeling of aromatic side chains in proteins for NMR relaxation measurements. *J. Am. Chem. Soc. 128*, 2506–2507.

(44) Lundström, P., Teilum, K., Carstensen, T., Bezsonova, I., Wiesner, S., Hansen, D. F., Religa, T. L., Akke, M., and Kay, L. E. (2007) Fractional ^{13}C enrichment of isolated carbons using [1-^{13}C]-or [2-^{13}C]-glucose facilitates the accurate measurement of dynamics at backbone Cα and side-chain methyl positions in proteins. *J. Biomol. NMR 38*, 199–212.

(45) Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., and Dahlquist, F. W. (1989) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. *Methods Enzymol. 177*, 44–73.

(46) Ayala, I., Sounier, R., Usé, N., Gans, P., and Boisbouvier, J. (2009) An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. *J. Biomol. NMR 43*, 111–119.

(47) Godoy-Ruiz, R., Guo, C., and Tugarinov, V. (2010) Alanine methyl groups as NMR probes of molecular structure and dynamics in high-molecular-weight proteins. *J. Am. Chem. Soc. 132*, 18340–18350.

(48) Barnes, C. O., and Pielak, G. J. (2011) In-cell protein NMR and protein leakage. *Proteins 79*, 347–351.

(49) Byeon, I. J., Louis, J. M., and Gronenborn, A. M. (2004) A captured folding intermediate involved in dimerization and domain-swapping of GB1. *J. Mol. Biol. 340*, 615–625.

(50) Putkey, J., Slaughter, G., and Means, A. (1985) Bacterial expression and characterization of proteins derived from the chicken calmodulin cDNA and a calmodulin processed gene. *J. Biol. Chem. 260*, 4704−4712.

(51) Lazar, G. A., Desjarlais, J. R., and Handel, T. M. (1997) De novo design of the hydrophobic core of ubiquitin. *Protein Sci. 6*, 1167−1178.

(52) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277−293.

(53) Ikeya, T., Sasaki, A., Sakakibara, D., Shigemitsu, Y., Hamatsu, J., Hanashima, T., Mishima, M., Yoshimasu, M., Hayashi, N., and Mikawa, T. (2010) NMR protein structure determination in living *E. coli* cells using nonlinear sampling. *Nat. Protoc. 5*, 1051−1060.

(54) Burz, D. S., and Shekhtman, A. (2010) The STINT-NMR method for studying in-cell protein-protein interactions. *Curr. Protoc. Protein Sci.*, 17.11-1−17.11-20.

(55) Burz, D. S., Dutta, K., Cowburn, D., and Shekhtman, A. (2006) Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR). *Nat. Methods 3*, 91−93.

(56) Fernández, C., and Wider, G. (2003) TROSY in NMR studies of the structure and function of large biological macromolecules. *Curr. Opin. Struct. Biol. 13*, 570−580.

(57) Weininger, U., Diehl, C., and Akke, M. (2012) ^{13}C relaxation experiments for aromatic side chains employing longitudinal- and transverse-relaxation optimized NMR spectroscopy. *J. Biomol. NMR 53*, 181−190.

(58) Sprangers, R., and Kay, L. E. (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature 445*, 618−622.