Published in final edited form as:
Biochemistry. 2013 June 25; 52(25): 4285–4287. doi:10.1021/bi4007034.

Quantitation of Recombinant Protein in Whole Cells and Cell Extracts with Solid-State NMR Spectroscopy

Erica P. Vogel and **David P. Weliky***
Department of Chemistry, Michigan State University, East Lansing, Michigan 48824, United States

Abstract

Recombinant proteins (RPs) are commonly expressed in bacteria followed by solubilization and chromatography. Purified RP yield can be diminished by losses at any step with very different changes in methods needed to try to improve yield. Time and labor can therefore be saved by first identifying the specific reason for low yield. The present study describes a new solid-state NMR approach to RP quantitation in whole cells or cell pellets without solubilization or purification. The approach is straightforward and inexpensive and only requires ~50 mL culture and a low-field spectrometer.

A common approach to produce recombinant protein (RP) begins with incorporation of recombinant DNA (rDNA) into bacteria followed by cell growth, expression and lysis, and finally chromatography to obtain pure RP. Assessment of RP quantity and purity after the expression, solubilization, and/or chromatography steps is typically done using SDS-PAGE that separates proteins by molecular weight (MW). For several different RPs in our laboratory, the RP gel band was not clearly observed after expression or solubilization and the final RP purified yield was unacceptably low, eg. 0.1 mg RP/L culture.[1] One hypothesis to explain this result is poor RP expression followed by high-yield solubilization and chromatography. A second distinct hypothesis is high RP expression followed by poor solubilization and highyield chromatography. A third hypothesis is high RP expression and solubilization followed by chromatographic loss of RP. Distinguishing between these hypotheses is important because: (1) the corrective changes to the experimental protocol to improve RP yield are very different for each hypothesis; and (2) implementing these changes is often time- and labor-intensive. For example, low protein expression might be improved by codon changes in the rDNA or by varying induction time whereas poor solubilization might be improved by comprehensive screening of lysis buffers which differ in additives such as denaturants and detergents.

The present study focuses on distinguishing between the first low expression and the second poor solubilization hypotheses. The third chromatographic loss hypothesis is typically straightforwardly tested by comparing the relative RP gel band intensities of washes vs elutions from the chromatographic column. RP expression is typically examined by first boiling an aliquot of cells in buffer containing SDS buffer with subsequent SDS-PAGE of solubilized protein. The RP quantity is estimated by comparison of the intensity of the RP

*Corresponding Author Phone: (517)355-9715. weliky@chemistry.msu.edu.

ASSOCIATED CONTENT

Supporting Information

Additional NMR spectra, cell growth and labeling, sample preparation, plasmids and protein sequences, NMR methods and analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

The authors declare no competing financial interests.

band to the intensities of bands of native bacterial proteins. There are a few reports of more accurate quantitation. [2] This approach relies on a RP MW which is fortuitously different from the MWs of any of the abundant bacterial proteins. Alternatively, the quantity of the solubilized RP could be much higher than the quantities of any of these native proteins, ie. high RP expression and solubility.

An assumption of the approach is that most of the RP is solubilized by boiling. However, the largest RP fraction in cells is typically solid inclusion body (IB) aggregates that can be difficult to solublilize. It is therefore important to develop alternative approaches for RP quantitation in either whole cells or cell extracts enriched in IB solids. One potential method is IR spectroscopy of IBs and is based on the hypothesis of an increased fraction of β sheet for the RP in IBs relative to the native structure, perhaps because of partial amyloid structure in the IB.[3] However, the fractional increase in β sheet structure is likely highly variable among RPs in IBs with at least one RP in IBs showing retention of a large fraction of native helical structure.[4]

The present study describes an alternate solid-state NMR (SSNMR) approach to quantify RP in whole bacterial cells and cell extracts enriched in IBs. The approach does not depend on the structure(s) of the RPs in IBs. We note that there have been earlier applications of SSNMR to whole bacterial cells and cell extracts with a typical goal of elucidation of details of atomicresolution structure.[5–11] The new method has been tested with five different RPs whose amino acid sequences are given in the SI. The generality of the approach is supported by use of different plasmid and *Escherichia coli* (*E. coli*) strain types.

One RP is Human proinsulin (HPI) which is the precursor to the hormone insulin.[12] Folded HPI is a monomer with an α helical core.[13] Three RPs (Hairpin, Fgp41, and Fgp41+) are different ectodomain segments of the HIV gp41 protein.[1,14,15] Gp41 is an integral HIV membrane protein and the N-terminal ~175 residues of gp41 are the ectodomain that lies outside the virus. The ectodomain is subdivided into the ~20 N-terminal fusion peptide (FP) residues that bind to membranes and the larger C-terminal region that folds as a helical hairpin with a 180° turn.[16] There is further assembly of three hairpins to form a molecular trimer with sixhelix-bundle (SHB) structure that is hyperthermostable. Hairpin, Fgp41, and Fgp41+ likely all form SHB structure with sequence differences among constructs as well as lack of FP and most of the loop in Hairpin. The fifth RP (FHA2) is the full ectodomain of the HA2 subunit of the hemagglutinin protein of the influenza virus.[17–19] HA2 has similar topology and membrane fusion function as gp41 but there is little sequence homology between HA2 and gp41.[20] Membrane-associated FHA2 has folded SHB structure. Previous efforts to solubilize each of these RPs from bacteria were consistent with a large fraction of RP in IBs.

We present “HC” and “HCN” SSNMR approaches to RP quantitation which respectively require double-resonance $^{1}H/^{13}C$ and triple-resonance $^{1}H/^{13}C/^{15}N$ SSNMR spectrometers and probes. For the HC approach, two samples are prepared that are denoted “RP^{+}_{lab}” and “RP^{-}_{lab}” and the bacteria respectively have a plasmid with or with the RP rDNA insert. The preparation of either sample includes addition of a ^{13}CO-labeled amino acid to the expression medium. For the present study, this is ^{13}CO-Leu. The SSNMR (RP^{+}_{lab}–RP^{-}_{lab}) ^{13}CO difference intensity should therefore be the signal of the labeled (*lab*) ^{13}CO nuclei of the RP. Comparison with a standard curve of ^{13}CO intensity vs mole ^{13}CO allows for conversion to mass RP/L bacterial culture which is a common metric of RP expression. Variation in cell mass between the RP^{+}_{lab} and RP^{-}_{lab} samples is accounted for by matching the intensities of the two samples in the 0–90 ppm aliphatic region. This aliphatic ^{13}C signal serves as an internal standard because it is due to natural abundance (*na*) nuclei whose numbers should be comparable in a RP^{+} and a RP^{-} cell.

The second HCN approach applies rotational-echo double-resonance (REDOR) SSNMR to one RP^{+}_{lab} sample labeled with either a ^{13}CO,^{15}N-amino acid or a ^{13}CO-amino acid and a ^{15}N-amino acid.[21] Separate S_0 and S_1 REDOR data are acquired with 1 ms dephasing. All ^{13}C signals are detected in S_0 whereas there is specific attenuation in S_1 of signals of directly-bonded ^{13}CO-^{15}N spin pairs because of the ~1 kHz dipolar coupling. The $\Delta S = S_0 - S_1$ ^{13}CO spectrum is therefore dominated by these pairs.[22] The ΔS ^{13}CO signal intensity is converted to RP mass/L culture using a method analogous to that of the HC approach. Relative to HC, the HCN RP quantitation has the advantage of one rather than two samples. The HC variant has the advantage of requiring a double-rather than triple-resonance SSNMR spectrometer and probe. Triple-resonance SSNMR instruments are less common and can have lower ^{13}C sensitivity.

The SI provides detailed protocols for sample preparation and SSNMR. The ~2-day experiment is mostly unattended. The approach is inexpensive with only ~50 ml culture volume and ~10 mg labeled amino acid. There is also a protocol to suppress scrambling of the ^{13}CO and/or ^{15}N labels to other amino acid types.[23] A “whole cell” (WC) sample is the centrifugation pellet of the bacterial culture. Cell lysis is done prior to centrifugation for the “insoluble cell pellet” (ICP) sample. The RP^{+}_{lab} ICP is therefore enriched in IB RP. The 9.4 T magnetic field, 8 kHz MAS frequency, and ~50 kHz rf fields are moderate and accessible for many NMR facilities including those with a SSNMR probe (~$100,000 cost) on an otherwise liquid-state NMR instrument.

Fig. 1 displays results from HC RP quantitation. Panel a displays ^{13}C spectra of the ^{13}CO-Leu RP^{-}_{lab} and RP^{+}_{lab} samples. Although there are differences in plasmid and *E. coli* strain types among the samples, the corresponding spectra have similar aliphatic ^{13}C signal intensities in the 0–90 ppm region. The isotropic ^{13}CO signals of the spectra are near 175 ppm with much weaker spinning sideband ^{13}CO signals near 95 and 255 ppm. Relative to RP^{-}_{lab}, there are much larger ^{13}CO signals in the RP^{+}_{lab} spectra which supports significant expression of all the RPs. There are also differences among the ^{13}CO intensities of the different RP^{+}_{lab} samples which support RP-dependent variation in expression. Panel b displays the expression levels in mg RP/L culture as determined from: (1) measurements of I_{Al}^{-} and $I_{CO}-$, respectively the integrated aliphatic and isotropic ^{13}CO intensities of the RP^{-}_{lab} spectrum; (2) measurement of the corresponding I_{Al}^{+} and I_{CO}^{+} of the RP^{+}_{lab} spectrum; and (3) calculation of the ^{13}CO intensity from ^{13}CO-Leus in the RP using $I_{Al}^{\rho} \times [(I_{CO}^{+}/I_{Al}^{+}) - (I_{CO}^{-}/I_{Al}^{-})]$ where the I_{Al}^{ρ} is the value for a typical sample and $I_{Al}^{\rho} \times (I_{CO}^{+}/I_{Al}^{+})$ and $I_{Al}^{\rho} \times (I_{CO}^{-}/I_{Al}^{-})$ are ^{13}CO intensities normalized to NMR sample mass. The expression level is calculated using (3) and (i) an experimentally-determined μmole ^{13}CO/I_{CO} conversion factor; (ii) MW_{RP}/N_{Leu} where N_{Leu} is the number of Leus in the RP sequence; and (iii) NMR sample is from cells in ~25 mL culture volume.

The SSNMR-determined expression levels (panel b) are 100–450 mg RP/L culture. These levels are very high relative to the reported ~5 mg/L purified yields for Fgp41, FHA2, and HPI.[1,12,18] The most common current approach to assess RP expression is SDS-PAGE. Panel c displays SDS-PAGE of boiled ICPs. Relative to the back ground, there are clear bands for HPI and Hairpin and much fainter and more ambiguous bands for FHA2 and Fgp41. The variation of the RP band intensities in the SDS-PAGE is more reflective of differences in RP IB solubilization than differences in expression levels. FHA2 and Fgp41 are membrane proteins while HPI and Hairpin are not, so the membrane RP IBs appear to be less well-solubilized. The SSNMR approach has the important advantage of being independent of IB solubilization.

The HCN approach is based on the $\Delta S = S_0 - S_1$ ^{13}CO REDOR difference spectrum of the RP^{+}_{lab} ICP sample. This spectrum is dominated by directly-bonded *lab* ^{13}CO-^{15}N spin pairs

in the IB RP. For Fig. 2a, RP≡Fgp41, *lab*≡^{13}CO,^{15}N-Leu, and the ΔS spectrum is mostly due to the *N*-terminal ***L***s of the six ***LL*** dipeptides in the Fgp41 sequence. One control is the RP^-_{lab} ΔS spectrum which is dominated by ***LL*** dipeptides of proteins other than Fgp41 produced during expression. However, there is no $\Delta S(RP^-_{lab})$ signal, cf. SI, or equivalently, Fig. 2b shows a $\Delta S(RP^+_{lab})$–$\Delta S(RP^-_{lab})$ spectrum very similar to the $\Delta S(RP^+_{lab})$ spectrum that must therefore be dominated by the IB Fpg41 signals. Another control is the $\Delta S(RP^+_{na})$ spectrum of a sample prepared with unlabeled Leu and reflecting signals of *na* ^{13}CO-^{15}N spin pairs. However, there is little $\Delta S(RP^+_{na})$ signal as reflected in Fig. 2c $\Delta S(RP^+_{lab})$–$\Delta S(RP^+_{na})$ spectrum that is similar to the $\Delta S(RP^+_{lab})$ spectrum.

The HCN approach to quantitation of RP expression is detailed in the SI. For a particular RP^+_{lab} sample, the HC and HCN expression levels typically agree to within a factor of 2. Quantitative labeling of the RP is assumed for both approaches so the levels are likely lower limits of expression but probably within a factor of ~2. Incomplete labeling will have a larger effect on HCN quantitation because the ΔS signal is only observed from dipeptides with both residues labeled.

Most of folded Fgp41 is a thermostable six-helix bundle which includes the six ***LL*** dipeptides.[16] The ΔS spectrum was previously obtained for ^{13}CO,^{15}N-Leu Fgp41 that had been purified, refolded, and reconstituted in membranes.[1] There was a single peak with 178 ppm shift and 3 ppm width which is consistent with folded helical structure. The Fig. 2 $\Delta S(RP^+_{lab})$ spectrum of Fgp41 in IBs is very similar and supports formation of folded Fgp41 structure in the IBs. For other RPs in IBs, the ΔS spectral widths are sometimes much broader, eg. ~7 ppm for HPI, cf. SI. This breadth is consistent with unfolded RP structure in the IBs. SSNMR quantitation of RP expression by either the HC or HCN approaches is independent of the degree of RP folding in the IBs.

For all the RPs of the present study, the SSNMR spectra demonstrated high expression, ie. ≥100 mg IB RP/L culture, so the main obstacle to purified RP is solubilization of the IBs. For other RPs that are produced at much lower levels, SSNMR could also be applied to optimize RP production. ^{13}CO-RP^+_{lab} samples would be prepared with different growth and/or expression parameters and expression levels determined from the ^{13}CO intensities. In summary, this paper describes general, inexpensive, rapid, and straightforward SSNMR approaches to RP quantitation in whole cells and cell extracts without purification.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Drs. R. Mackin , Y-K Shin, and K. Sackett provided plasmids.

Funding Sources

National Institutes of Health Grant A147153

ABBREVIATIONS

Fgp41	Fgp41+, and Hairpin, HIV gp41 ectodomains
FHA2	Influenza hemagglutinin ectodomain
HPI	human proinsulin

IB	inclusion body
ICP	insoluble cell pellet
REDOR	rotational-echo double-resonance
RP	recombinant protein
SHB	six-helix bundle
SSNMR	solid-state NMR
WC	Whole cell

REFERENCES

1. Vogel EP, Curtis-Fisk J, Young KM, Weliky DP. Biochemistry. 2011; 50:10013–10026. [PubMed: 21985645]
2. Miles, AP.; Saul, A. Protein Protocols Handbook. 3rd ed. Totowa: Humana; 2009. p. 487-496.
3. Gross-Selbeck S, Margreiter G, Obinger C, Bayer K. Biotech. Prog. 2007; 23:762–766.
4. Curtis-Fisk J, Spencer RM, Weliky DP. J. Am. Chem. Soc. 2008; 130:12568–12569. [PubMed: 18759389]
5. Wang J, Balazs YS, Thompson LK. Biochemistry. 1997; 36:1699–1703. [PubMed: 9048553]
6. Kim SJ, Cegelski L, Stueber D, Singh M, Dietrich E, Tanaka KSE, Parr TR, Far AR, Schaefer J. J. Mol. Biol. 2008; 377:281–293. [PubMed: 18258256]
7. Ieronimo M, Afonin S, Koch K, Berditsch M, Wadhwani P, Ulrich AS. J. Am. Chem. Soc. 2010; 132:8822–8823. [PubMed: 20550126]
8. Fu RQ, Wang XS, Li CG, Santiago-Miranda AN, Pielak GJ, Tian F. J. Am. Chem. Soc. 2011; 133:12370–12373. [PubMed: 21774553]
9. Reckel S, Lopez JJ, Lohr F, Glaubitz C, Dotsch V. Chembiochem. 2012; 13:534–537. [PubMed: 22298299]
10. Zhou XX, Cegelski L. Biochemistry. 2012; 51:8143–8153. [PubMed: 22974326]
11. Wang SL, Ishii Y. J. Am. Chem. Soc. 2012; 134:2848–2851. [PubMed: 22280020]
12. Mackin RB, Choquette MH. Prot. Expr. Purif. 2003; 27:210–219.
13. Yang YW, Hua QX, Liu J, Shimizu EH, Choquette MH, Mackin RB, Weiss MA. J. Biol. Chem. 2010; 285:7847–7851. [PubMed: 20106974]
14. Sackett K, Nethercott MJ, Shai Y, Weliky DP. Biochemistry. 2009; 48:2714–2722. [PubMed: 19222185]
15. Sackett K, Nethercott MJ, Epand RF, Epand RM, Kindra DR, Shai Y, Weliky DP. J. Mol. Biol. 2010; 397:301–315. [PubMed: 20080102]
16. Yang ZN, Mueser TC, Kaufman J, Stahl SJ, Wingfield PT, Hyde CC. J. Struct. Biol. 1999; 126:131–144. [PubMed: 10388624]
17. Curtis-Fisk J, Preston C, Zheng ZX, Worden RM, Weliky DP. J. Am. Chem. Soc. 2007; 129:11320–11321. [PubMed: 17718569]
18. Curtis-Fisk J, Spencer RM, Weliky DP. Prot. Expr. Purif. 2008; 61:212–219.
19. Kim CS, Epand RF, Leikina E, Epand RM, Chernomordik LV. J. Biol. Chem. 2011; 286:13226–13234. [PubMed: 21292763]
20. Chen J, Skehel JJ, Wiley DC. Proc. Natl. Acad. Sci. U.S.A. 1999; 96:8967–8972. [PubMed: 10430879]
21. Gullion T, Schaefer J. J. Magn. Reson. 1989; 81:196–200.
22. Yang J, Parkanzky PD, Bodner ML, Duskin CG, Weliky DP. J. Magn. Reson. 2002; 159:101–110. [PubMed: 12482688]
23. Tong KI, Yamamoto M, Tanaka T. J. Biomol. NMR. 2008; 42:59–67. [PubMed: 18762866]

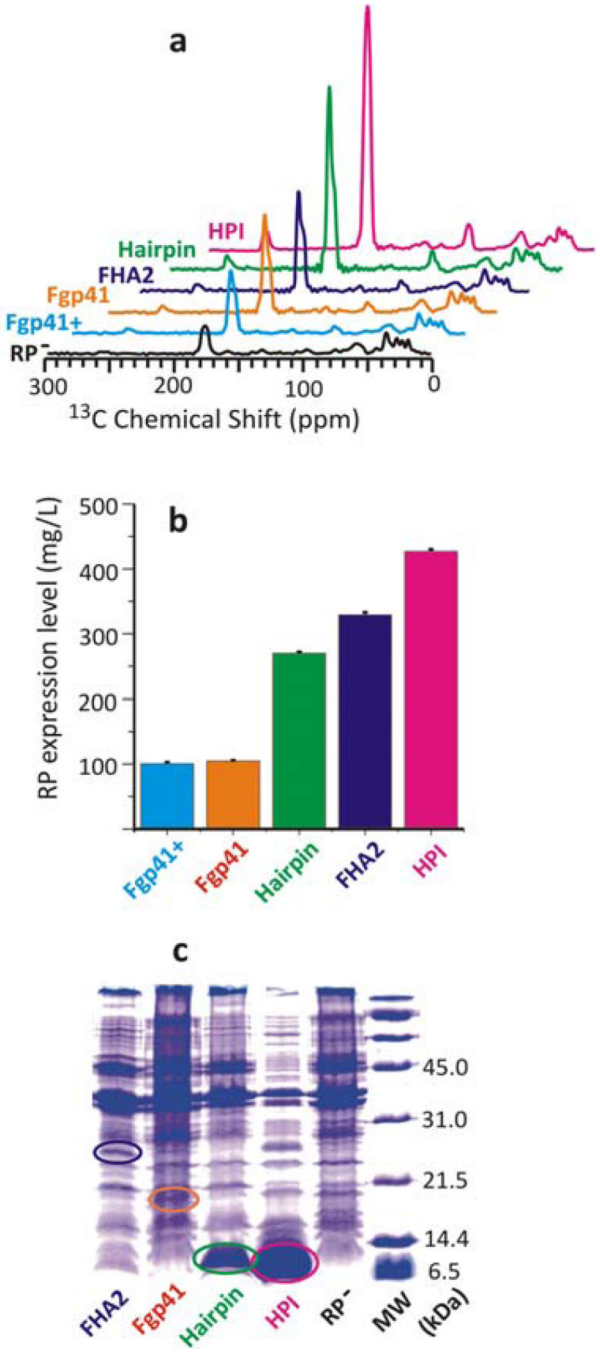

Figure 1.

HC variant of RP quantitation. Panel a displays ^{13}C SSNMR spectra of ^{13}CO-Leu ICP samples. Panel b displays the RP expression levels calculated from the (RP^{+}_{lab}–RP^{-}_{lab}) difference intensities of the ^{13}CO (~175 ppm) region. The differences between the ordering of spectral intensities in panel a and the ordering of expression levels in panel b are largely due to different ratios of N_{Leu}/N_{tot} in the RP sequences. Besides the displayed uncertainties based on spectral noise, there is ~10% sample-to-sample variation in RP expression, cf. SI. Panel c displays SDS-PAGE of most of the samples and the marked bands may correspond to the RPs.

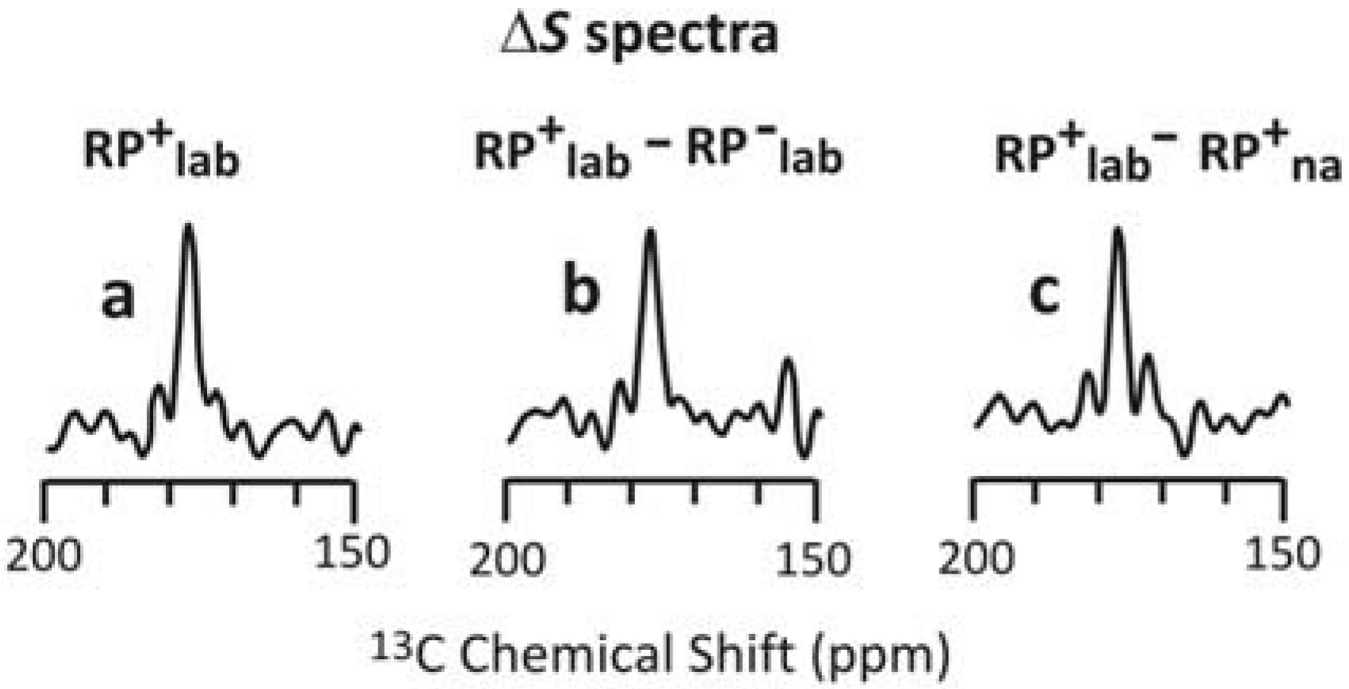

Figure 2.

^{13}CO Δ*S* spectra based on the S_0 and S_1 spectra of three different ICP samples. The RP^+_{lab} and RP^-_{lab} plasmids respectively had and lacked the Fgp41 insert. The *lab* and *na* expression media respectively contained ^{13}CO, ^{15}N-Leu and unlabeled Leu. Panel a ΔS (RP^+_{lab}=S_0–S_1 signal represents directly-bonded ^{13}CO-^{15}N spin pairs of the RP^+_{lab} sample. Panel b ΔS=ΔS(RP^+_{lab})–ΔS(RP^-_{lab}) is from spin pairs of IB Fgp41. Panel c ΔS=ΔS(RP^+_{lab})–ΔS(RP^+_{na}) is from *lab* spin pairs of the RP^+_{lab} sample. The similar peak intensities, shifts, and lineshapes of the three spectra support a dominant contribution to the ΔS(RP^+_{lab}) spectrum of ^{13}CO signals of labeled ***L***s of the ***LL*** dipeptides of IB Fgp41.