HHS Public Access

Author manuscript

J Biomol NMR. Author manuscript; available in PMC 2024 June 15.

Published in final edited form as:

J Biomol NMR. 2024 June ; 78(2): 125–132. doi:10.1007/s10858-023-00434-3.

A comprehensive assessment of selective amino acid ^{15}N-labeling in human embryonic kidney 293 cells for NMR spectroscopy

Ganesh P. Subedi[1], **Elijah T. Roberts**[2], **Alexander R. Davis**[3], **Paul G. Kremer**[3], **I. Jonathan Amster**[2], **Adam W. Barb**[1,2,3,4]

[1]Roy J. Carver Dept of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, Iowa

[2]Department of Chemistry, University of Georgia, Athens, GA

[3]Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA

[4]Complex Carbohydrate Research Center, University of Georgia, Athens, GA

Summary

A large proportion of human proteins contain post-translational modifications that cannot be synthesized by prokaryotes. Thus, mammalian expression systems are often employed to characterize structure/function relationships using NMR spectroscopy. Here we define the selective isotope labeling of secreted, post-translationally modified proteins using human embryonic kidney (HEK)293 cells. We determined that alpha-[^{15}N]- atoms from 10 amino acids experience minimal metabolic scrambling (C,F,H,K,M,N,R,T,W,Y). Two more interconvert to each other (G,S). Six others experience significant scrambling (A,D,E,I,L,V). We also demonstrate that tuning culture conditions suppressed V and I scrambling. These results define expectations for ^{15}N-labeling in HEK293 cells.

Keywords

glycoprotein; post-translational modification; CD16a; FcγRIIIa

Protein NMR spectroscopy was previously highly focused on structure determination [1]. However, with the rapid increase in high-resolution structures determined by X-ray crystallography, cryogenic electron microscopy, and computational predictions [2,3], NMR is poised to capitalize on these models and provide measurements to define how proteins function in solution without the need to determine all atom structural models. The advent of techniques over the past 25 years to promote data collection on large, slowly tumbling systems likewise promote the application of NMR to large, complex systems [4–6].

Corresponding author: abarb@uga.edu, Address: 120 E. Green St., Athens, GA, 30602.

Conflict of Interest

The authors declare that they have no conflicts of interest with the contents of this article.

With this fortuitous coalescence of supporting techniques, expanding NMR capabilities to investigate previously inaccessible proteins requires significant methods development efforts. For example, protein NMR studies have largely utilized proteins expressed in prokaryotic systems due to the ease of isotope labeling from simple metabolic precursors [7]. One significant limitation of this expression technology is an inability to produce proteins with often essential mammalian post-translation modifications including glycosylation. For example, glycoproteins are found in the ER and Golgi, at the cell surface, and secreted outside the cell. A large fraction of the human proteome, 22% or 4,512 proteins, contain at least one N-glycosylation site (average = 3.6). Often, proper glycans are essential for protein function (including, but not limited to ref. [8–10]). Glycoproteins, and other post-translationally modified proteins represent important drugs, including monoclonal antibodies, cytokines, and blood clotting factors, to name just a few. Thus, there is a clear need to improve the protein expression capabilities to define structure/function relationships in this cohort of dynamic and important proteins.

Post-translationally modified proteins, including glycoproteins, are often expressed in mammalian cells, including human embryonic kidney (HEK)293 cells. HEK293 cells grow in complex medium, though chemically defined media are available [11]. Appending an appropriate signal peptide to the N-terminus of the protein directs it to the ER translocon complex for import in the ER, providing an oxidative folding environment and a host of processing machinery including glycosyltransferases [12]. The disadvantage of mammalian cells, in contrast to *Escherichia coli*, is an inability to introduce ^{15}N and ^{13}C labels from inexpensive metabolic precursors like ammonium chloride and glucose. Furthermore, mammalian cells do not tolerate deuterium oxide and thus ^{2}H labeling is prohibited. Mammalian cell growth medium often contains a mixture of many (if not all) amino acids, thus uniform labeling requires supplementation with a labeled version of each amino acid.

Commercial media for HEK293 labeling are available, but are very expensive [13]. Labeling in HEK293 cells is described [14–17], however, a thorough evaluation of selective amino acid labeling is not known. Such a description is expected to provide a compendium for the behavior of individual ^{15}N labels during protein expression as well as provide guidance for labeling strategies. It is important to note that Prestegard and coworkers described multiple strategies to assign selectively-labeled amino acids, thus peaks can be assigned in a manner that avoids expensive uniform labeling or onerous single site mutagenesis [18–21]. Here we investigate the labeling of proteins using HEK293F cells grown in suspension.

The protein expression system investigated here was developed by Dr. Kelley Moremen, and utilizes transient transfection that permits proteins to be rapidly expressed without the weeks of time required to isolate stably-transfected cell lines [22,23]. Furthermore, commercial chemically-defined medium can be prepared without amino acids and sugars, allowing for medium to be reconstituted with a single labeled amino acid. This approach dramatically reduces costs compared to existing commercial products or uniform labeling. Here we investigated methods to generate amino acid labeling during protein expression with HEK293F cells and identified conditions for reducing metabolic scrambling and multiple amino acid labeling to increase the number of observable signals.

We expressed proteins using a culture medium supplemented with selective ^{15}N-labeled amino acids to determine how HEK293 cells metabolize the ^{15}N atoms, using 18 different amino acid types. Glutamine was excluded from the analysis due to the high concentration required for HEK cell growth and viability which thus introduces a significant expense. Proline was likewise excluded due to the absence of an amide proton leading to invisibility in standard^{1}H-detected ^{1}H-^{15}N correlation experiments. Culture medium contained each labeled and unlabeled amino acid at a concentration of 100 mg/L (except Q as noted; Kato and coworkers also recently developed a method to label glutamine [24]). NMR spectra of each protein, following purification, revealed a characteristic spectrum as shown in Figure 1. We evaluated whether the alpha amine of each amino acid was removed and incorporated into other amino acids, a process termed “scrambling,” by comparing the number of peaks in a 2d HSQC spectrum to the number of that type expected for the protein. Labeling with amino acids that experience minimal scrambling are expected to produce a spectrum with the same number of peaks compared to count of residues of that type found in the protein sequence, with few if any additional peaks for these well-behaved proteins.

Ten amino acids showed minimal metabolic scrambling (Figure 1), including six of the nine amino acids essential to human nutrition (F,H,K,M,T,W) as well as four additional residues (C,N,R,Y). Two amino acids, G and S, converted only to each other. Thus, a total of twelve amino acids exhibited limited or no observable scrambling (C,F,G,H,K,M,N,R,S,T,W,Y). Six amino acids exhibited a high degree of scrambling, including A,D,E,I,L and V. Interestingly, I,L and V are the three remaining essential amino acids, and show similar spectra indicating similar scrambling pathways. The metabolism of the alpha amine from P and Q in this system remain undefined.

We examined whether reducing the concentration of the labeled amino acid while keeping the unlabeled amino acids at 100 mg/L would decrease scrambling. This approach is expected to increase the percentage of the exogenous labeled amino acid that is incorporated into the protein and decrease the amount available for metabolism into other amino acid types. We examined the labeling of both V and I at 100 mg/L, 50 mg/L and 25 mg/L. Protein expressed from culture medium containing 100 mg/L of each amino acid showed a high degree of scrambling as expected, though labeling with 25 mg/L showed minimal scrambling with clear enrichment of intense peaks corresponding to the labeled amino acid type (Figure 2). The intermediate condition with 50 mg/L showed a spectrum of intermediate quality with substantial contribution from additional peaks, though at a reduced intensity compared to the strongest peaks in the spectrum also found with the 25 mg/L concentration.

It is likewise of interest to identify cost-effective labeling strategies that provide a great number of peaks than is available from single amino acid labeling. Combining multiple labeled amino acids into a single expression medium is expected to preserve the benefits of selective labeling, notably isolated peaks with minimal overlap. We first attempted supplementing the medium with 100 mg/L K and S plus 25 mg/L V with no labeled or unlabeled glycine. Based on the previous experiments, this condition is expected to reduce scrambling and provide signal for K,V,G and S residues. This strategy failed to produce protein and the cells showed reduced viability. We next explored supplementing the medium

with 100 mg/L K,G and S and recovered a high yield of protein. The spectrum revealed minimal scrambling with 38 peaks observed of 38 expected (Figure 3A).

We next evaluated V, I and L labeling, with the expectation that these likely interconverted with one another, and supplementing all three at 100 mg/L will increase observable signals in a single sample. Indeed, a spectrum of the ^{15}N-VIL CD16a sample showed 42 strong peaks with 32 weaker peaks, compared to 35 expected VIL peaks (Figure 3B). This number would increase to 43 if A is likewise labeled through metabolic scrambling. Thus, this labeling strategy provided strong labeling of VIL and A residues, with weaker labeling of other amino acids. Here we demonstrated that VIL labeling provides a low-cost labeling option.

Lastly, we quantified the percentage of ^{15}N incorporation for the ^{15}N-VIL and the ^{15}N-KGS labeled samples. Though we established metabolic scrambling of the ^{15}N atoms with the NMR experiments, it is possible, in fact probable for some residues, that ^{14}N incorporation into the target amino acids occurs using other ^{14}N sources in the medium as it is only selectively ^{15}N labeled. MALDI-FTICR-MS spectra of trypsinized peptides recovered from secreted proteins revealed clear shifts in the peak multiplet, indicating isotope incorporation (Figure 3CD). The abundances of the isotope peaks in the MS data from four peptides for each sample was fitted to statistically calculated values to determine the amount of peptide that contained at least one additional heavy isotope label. This analysis provides an estimate of isotope incorporation of 52 ±4% and 30 ±14% for the ^{15}N-KGS and ^{15}N-VIL samples, respectively. The labeling percentage is significantly below 100%, indicating that even though the amino acids of residues like Lys, Gly and Ser are not scrambled into other amino acids, it is likely that a significant pool of ^{14}N amino acids exist. This pool could arise from multiple sources. Grzesiek and coworkers recently noted that Gln is scrambled to Ala, Asp and Glu in insect cells, and the high unlabeled Gln concentration may contribute ^{14}N incorporation we observed [25]. One additional source may be the residual pool of amino acids present prior to transferring the expressing cells into the labeling medium. It is also possible that the alpha amino groups of these residues are being added from unlabeled sources present in the growth medium. Despite this reduced labeling efficiency, it is important to note that HEK293 cells support robust labeling and provide high quality NMR spectra.

It is interesting to note that the ^{15}N-lysine included both ^{15}N-alpha and ^{15}N-epsilon atoms, but the detected peptides contained predominantly one ^{15}N atom. It is notable that we observed minimal scrambling of the ^{15}N alpha atom by NMR spectroscopy, suggesting the sidechain epsilon ^{15}N was exchanged. Saccharopine dehydrogenase catalyzed deamination is the first step in lysine catabolism in humans but is reversible and can exchange the epsilon nitrogen with the alpha nitrogen from glutamate[26], which would not be labeled in the ^{15}N-KGS labeling medium and readily exchanging with other ^{14}N-amino acids.

In conclusion, these experiments demonstrate selective amino acid ^{15}N labeling using the HEK293F protein expression system for 18 amino acid types. The availability of amino acid selective labeling provides identity information to each peak, and here we showed minimal metabolic scrambling for 12 residues, along with feasible strategies to reduce scrambling for two more. Selective amino acid labeling represents an accessible strategy

to introduce NMR-observable nuclei into heavily post-translationally modified proteins, including glycoproteins, for study by NMR spectroscopy.

Experimental Procedures

Materials.

All materials were purchased from Sigma unless otherwise noted.

Protein expression.

We expressed both IgG1 Fc and GFP-CD16a from the pGen2 vector [22]. Both vectors include a signal sequence to import the nascent polypeptide into the ER, at which point N-glycosylation and folding occurs. These peptides are then secreted into the medium. The sequences of the secreted proteins are:

GFP-CD16a ("//**//" denotes the TEV cleavage site):

MHHHHHHHHMSGLNDIFEAQKIEWHEMSKGEELFTGVVPILVELDGDVNGHKFSV
RGEGEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKRHDFFKS
AMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEY
NFNSHNVYITADKQKNGIKANFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNH
YLSTQSVLSKDPNEKRDHMVLLEFVTAAGITHGEFSSENLYFQ//**//
GRTEDLPKAVVFLEPQWYRVLEKDSVTLKCQGAYSPEDNSTQWFHNESLISSQASSY
FIDAATVDDSGEYRCQTNLSTLSDPVQLEVHIGWLLLQAPRWVFKEEDPIHLRCHSW
KNTALHKVTYLQNGKGRKYFHHNSDFYIPKATLKDSGSYFCRGLVGSKNVSSETVN
ITITQG

IgG1 Fc:
AKSEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHED
PEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVS
NKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWE
SNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQ
KSLSLSPGK

pGen2 vectors encoding these two proteins were used to transiently transfect HEK293F cells according to a previously published method [23], except that the medium used during the transfection and the dilution (24h after transfection) differed. The custom Freestyle293 culture medium utilized here was purchased from Life Technologies, and did not include any amino acids (including Glutamax) or carbohydrates. Furthermore, the osmolarity of this medium was not adjusted by the manufacturer. For each expression, one or more [^{15}N]-labeled amino acids (Isotec) were added to the medium followed by every remaining amino acid at a concentration of 100 mg/L; these remaining amino acids did not contain [^{15}N] enrichment. The only exception was glutamine, which was added at 1 g/L. Unlabeled glucose was added at a final concentration of 3 g/L. After these additions, osmolarity was adjusted to 260-280 mOsm/kg using a 5004 Micro Osmette machine (Precision Systems), the volume adjected by adding distilled/deionized water, and the medium was passed through a sterile 0.20 μm filter (VWR). Following the culture incubation (4-6 d), cells

and cell debris were removed by centrifugation (2x 5 min 1000 x g) and the used medium filtered before protein purification.

Protein Purification.

Clarified spent culture medium containing IgG1 Fc was diluted 1:1 with 20 mM sodium phosphate, 150 mM sodium chloride, pH 7.2, then applied to a Protein A Sepharose column on an Åkta Go system (Cytiva). IgG1 Fc was eluted with a 100 mM glycine, pH 3.0, in 2 mL fractions into tubes containing 1 mL of 1.0 M Tris, pH 8.0. IgG1 Fc was then exchanged into a buffer containing 20 mM sodium phosphate, 100 mM potassium chloride, and 0.1 mM 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS), 5% D_2O, pH 7.2 using an Amicon centrifugal concentration unit with a 10 kDa molecular weight cutoff. Protein samples were analyzed at a 100-400 μM concentration.

Spent culture medium containing GFP-CD16a was diluted 1:1 with 50 mM MOPS, 200 mM potassium chloride, 25 mM imidazole, pH 7.2, then passed over a Ni-NTA column (Qiagen). GFP-CD16a was eluted with 50 mM MOPS, 200 mM potassium chloride, 250 mM imidazole, pH 7.2. GFP-CD16a was then exchanged into a buffer containing 20 mM sodium phosphate, 100 mM potassium chloride, and 0.5 mM 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS), 5% D_2O, pH 7.2 using an Amicon centrifugal concentration unit with a 10 kDa molecular weight cutoff.

CD16a was cleaved from GFP using tobacco etch virus protease (TEV). First, GFP-CD16a was exchanged into a buffer containing 50 mM Tris, 0.5 mM EDTA, 1 mM dithiothreitol, pH 8.0. One gram of TEV was added for every 50 grams of GFP-CD16a, and the reaction was incubated at 23 °C for 24 h in the dark without mixing. The TEV mixture was next diluted 1:10 with buffer A containing 25 mM Tris, pH 7.4, then applied to a QFF column on an Åkta Go FPLC (Cytiva). The column was eluted stepwise with 10%, 11%, 12% of buffer B (25 mM Tris, 1 M potassium chloride, pH 7.4) in buffer A. CD16a eluted in fractions corresponding to 10% buffer B. GFP is then eluted at 25% buffer B in buffer A. CD16a or GFP were then exchanged into a buffer containing 20 mM sodium phosphate, 100 mM potassium chloride, and 0.5 mM 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS), 5% D_2O, pH 7.2, using an Amicon centrifugal concentration unit with a 10 kDa molecular weight cutoff

NMR spectroscopy.

NMR spectra were collected at a 30 °C (GFP, CD16a) or 50 °C (IgG1 Fc) sample temperature on one of three solution NMR systems, including: 1) 18.8 T spectrometer equipped with a Bruker NEO console and 1.7 mm TXO cryoprobe, 2) 18.8 T spectrometer equipped with a Bruker Avance III console and 5 mm TXO cryoprobe, and 3) 16.4 T spectrometer equipped with a Bruker Avance II console and 5 mm TXO cryoprobe. HSQC-TROSY spectra were processed with NMRPipe [27] and analyzed in NMRViewJ [28].

FT-ICR Mass Spectrometry Data Collection and Analysis: CD16a and GFP were trypsinized, then desalted using a C18 Zip-tip. Mass spectra were collected on a Bruker SolariX XR 12 T Fourier Transform Ion Cyclotron Resonance (FT-ICR) mass spectrometer

equipped with a dual ESI/MALDI source. Calibration used electrospray ionization of a 5 mg/mL solution of Cesium Iodide (Aldrich 99.9999%) in positive mode. Mass spectra for peptides were acquired using the MALDI source equipped with a SmartBeam II laser. The laser power was set to between 50 – 60% and between 5 – 50 laser shots were used. Mass spectra were collected between 500 – 5000 m/z. 512 k data points were collected with a transient length of 0.70 s which gave a resolution of 155,000 at 500 m/z. Initial analysis of mass spectra was performed using Bruker Data Analysis 5.3. Isotopic enrichment levels were determined by calculating the best fit to the observed intensity data for two isotopologue patterns: one with the standard natural isotope abundance, and a second with one ^{15}N label, by varying the coefficient for each distribution. Best-fit values representing the most probably isotope incorporation were those with the lowest Root-means squared error values.

Acknowledgement

We thank Prof. Kelley Moremen (UGA) for suggesting this inquiry area and for use of the instrument to measure osmolality.

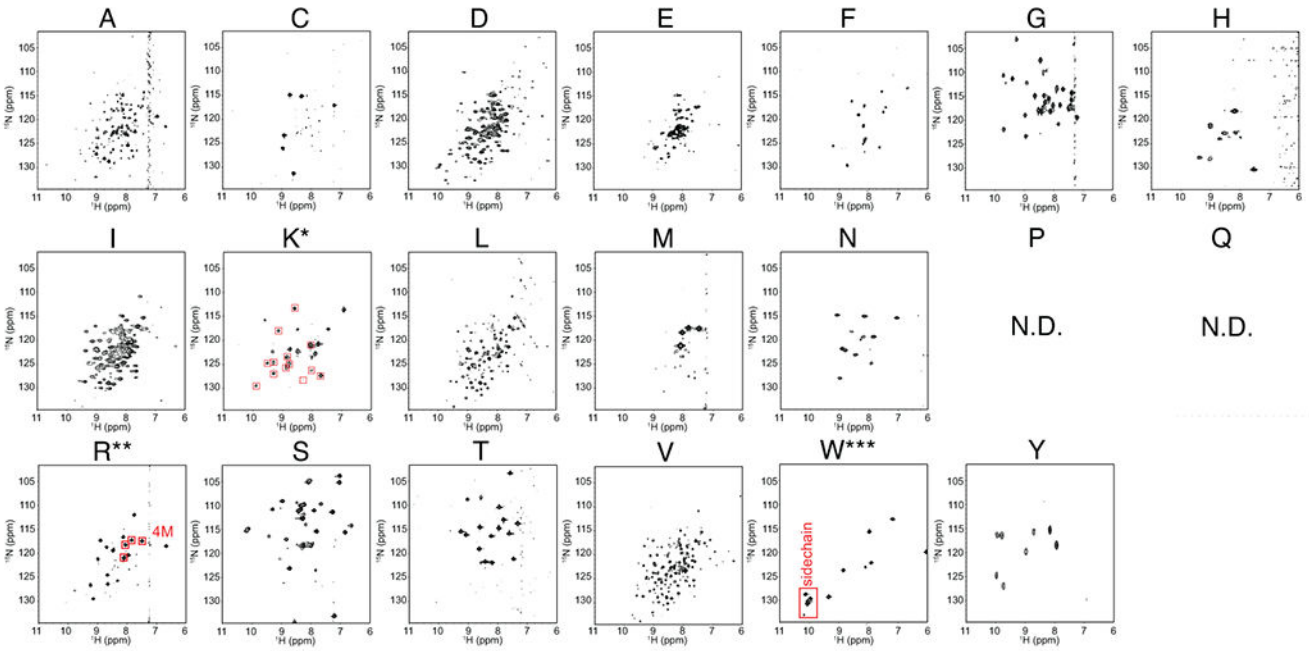
Funding

This work was supported by National Institutes of Health Award NIAID U01 AI148114 to AWB, NIH S10 OD025118 to IJA, and by funds from the Biochemistry and Molecular Biology Department at the University of Georgia and the Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Works Cited

1. Wuthrich K The way to NMR structures of proteins. Nat Struct Biol 8, 923–5 (2001). [PubMed: 11685234]
2. Callaway E Revolutionary cryo-EM is taking over structural biology. Nature 578, 201 (2020). [PubMed: 32047310]
3. Jumper J et al. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589 (2021). [PubMed: 34265844]
4. Pervushin K, Riek R, Wider G & Wuthrich K Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci U S A 94, 12366–71 (1997). [PubMed: 9356455]
5. Riek R, Wider G, Pervushin K & Wuthrich K Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules. Proc Natl Acad Sci U S A 96, 4918–23 (1999). [PubMed: 10220394]
6. Ollerenshaw JE, Tugarinov V & Kay LE Methyl TROSY: explanation and experimental verification. Mag Res Chem 41, 843–852 (2003).
7. Venters RA, Farmer BT 2nd, Fierke CA & Spicer LD Characterizing the use of perdeuteration in NMR studies of large proteins: 13C, 15N and 1H assignments of human carbonic anhydrase II. J Mol Biol 264, 1101–16 (1996). [PubMed: 9000633]
8. Subedi GP & Barb AW The Structural Role of Antibody N-Glycosylation in Receptor Interactions. Structure 23, 1573–1583 (2015). [PubMed: 26211613]
9. Ereno-Orbea J et al. Molecular basis of human CD22 function and therapeutic targeting. Nat Commun 8, 764 (2017). [PubMed: 28970495]
10. Takeuchi H & Haltiwanger RS Significance of glycosylation in Notch signaling. Biochem Biophys Res Commun 453, 235–42 (2014). [PubMed: 24909690]

11. Barb AW, Falconer DJ & Subedi GP The Preparation and Solution NMR Spectroscopy of Human Glycoproteins Is Accessible and Rewarding. Methods Enzymol 614, 239–261 (2019). [PubMed: 30611426]
12. Moremen KW, Tiemeyer M & Nairn AV Vertebrate protein glycosylation: diversity, synthesis and function. Nat Rev Mol Cell Biol 13, 448–62 (2012). [PubMed: 22722607]
13. Sastry M, Bewley CA & Kwong PD Effective isotope labeling of proteins in a mammalian expression system. Methods Enzymol 565, 289–307 (2015). [PubMed: 26577736]
14. Yamaguchi Y et al. Glycoform-dependent conformational alteration of the Fc region of human immunoglobulin G1 as revealed by NMR spectroscopy. Biochim Biophys Acta 1760, 693–700 (2006). [PubMed: 16343775]
15. Luchinat E, Barbieri L & Banci L A molecular chaperone activity of CCS restores the maturation of SOD1 fALS mutants. Sci Rep 7, 17433 (2017). [PubMed: 29234142]
16. Barbieri L, Luchinat E & Banci L Characterization of proteins by in-cell NMR spectroscopy in cultured mammalian cells. Nat Protoc 11, 1101–11 (2016). [PubMed: 27196722]
17. Dutta A, Saxena K, Schwalbe H & Klein-Seetharaman J Isotope labeling in mammalian cells. Methods Mol Biol 831, 55–69 (2012). [PubMed: 22167668]
18. Chalmers GR et al. NMR Resonance Assignment Methodology: Characterizing Large Sparsely Labeled Glycoproteins. J Mol Biol 431, 2369–2382 (2019). [PubMed: 31034888]
19. Nkari WK & Prestegard JH NMR resonance assignments of sparsely labeled proteins: amide proton exchange correlations in native and denatured states. J Am Chem Soc 131, 5344–9 (2009). [PubMed: 19317468]
20. Feng L, Lee HS & Prestegard JH NMR resonance assignments for sparsely 15N labeled proteins. J Biomol NMR 38, 213–9 (2007). [PubMed: 17487550]
21. Williams RV et al. AssignSLP_GUI, a software tool exploiting AI for NMR resonance assignment of sparsely labeled proteins. J Magn Reson 345, 107336 (2022). [PubMed: 36442299]
22. Barb AW et al. NMR characterization of immunoglobulin G Fc glycan motion on enzymatic sialylation. Biochemistry 51, 4618–26 (2012). [PubMed: 22574931]
23. Subedi GP, Johnson RW, Moniz HA, Moremen KW & Barb A High Yield Expression of Recombinant Human Proteins with the Transient Transfection of HEK293 Cells in Suspension. J Vis Exp, e53568 (2015). [PubMed: 26779721]
24. Yanaka S, Yagi H, Yogo R, Onitsuka M & Kato K Glutamine-free mammalian expression of recombinant glycoproteins with uniform isotope labeling: an application for NMR analysis of pharmaceutically relevant Fc glycoforms of human immunoglobulin G1. J Biomol NMR 76, 17–22 (2022). [PubMed: 34978013]
25. Wu FJ, Kronenberg D, Hertel I & Grzesiek S The key role of glutamine for protein expression and isotopic labeling in insect cells. J Biol Chem 299, 105142 (2023). [PubMed: 37553040]
26. Matthews DE Review of Lysine Metabolism with a Focus on Humans. J Nutr 150, 2548S–2555S (2020). [PubMed: 33000162]
27. Delaglio F et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6, 277–93 (1995). [PubMed: 8520220]
28. Johnson BA Using NMRView to visualize and analyze the NMR spectra of macromolecules. Methods Mol Biol 278, 313–52 (2004). [PubMed: 15318002]



Amino Acid		Protein	Observed Peaks	Expected Peaks	Scrambling?
A	Alanine	GFP-CD16a	~100	17	Yes
C	Cysteine	GFP-CD16a	6 strong	6	Minimal
D	Aspartic Acid	GFP-CD16a	>100	28	Yes
E	Glutamic Acid	CD16a	~30	10	Yes
F	Phenylalanine	GFP	15	15	No
G	Glycine	CD16a	26	9 (27 G + S)	to Ser
H	Histidine	CD16a	11	7	No
I	Isoleucine	GFP-CD16a	>100	20	Yes
K	Lysine	CD16a	10	11	No
L	Leucine	GFP-CD16a	>100	37	Yes
M	Methionine	GFP	4 strong	5	Minimal
N	Asparagine	GFP	13	15	No
P	Proline	N.D.	N.D.	N.D.	N.D.
Q	Glutamine	N.D.	N.D.	N.D.	N.D.
R	Arginine	GFP-CD16a	14 strong	15	Minimal
S	Serine	GFP	~33	13 (36 G + S)	to Gly
T	Threonine	GFP	16	18	No
V	Valine	GFP-CD16a	>100	30	Yes
W	Tryptophan	GFP-CD16a	7	7	No
Y	Tyrosine	IgG1 Fc	8	9	No

Figure 1. Metabolic scrambling of ^{15}N amino acids in HEK293 cells.

Cells grown in suspension cultures were transiently transfected with an expression plasmid and grown in medium containing 100 mg/L of the ^{15}N-labeled amino acid. Remaining unlabeled amino acids were added at 100 mg/L, except glutamine (1 g/L). *The ^{15}N lysine sample was expressed along with ^{15}N tyrosine and ^{15}N phenylalanine which are indicated with *red* boxes. **The ^{15}N arginine sample was labeled simultaneously with ^{15}N methionine; the methionine peaks identified separately are indicated with four *red* boxes. ***The ^{15}N tryptophan sample also contains sidechain ^{15}N labeling, as indicated with a

red box. Spectra are scaled to the height of the noise level to show the weakest identifiable peaks. Protein identities as well as the expected and observed number of peaks are reported in the table at the bottom of the figure. N.D. – not determined.

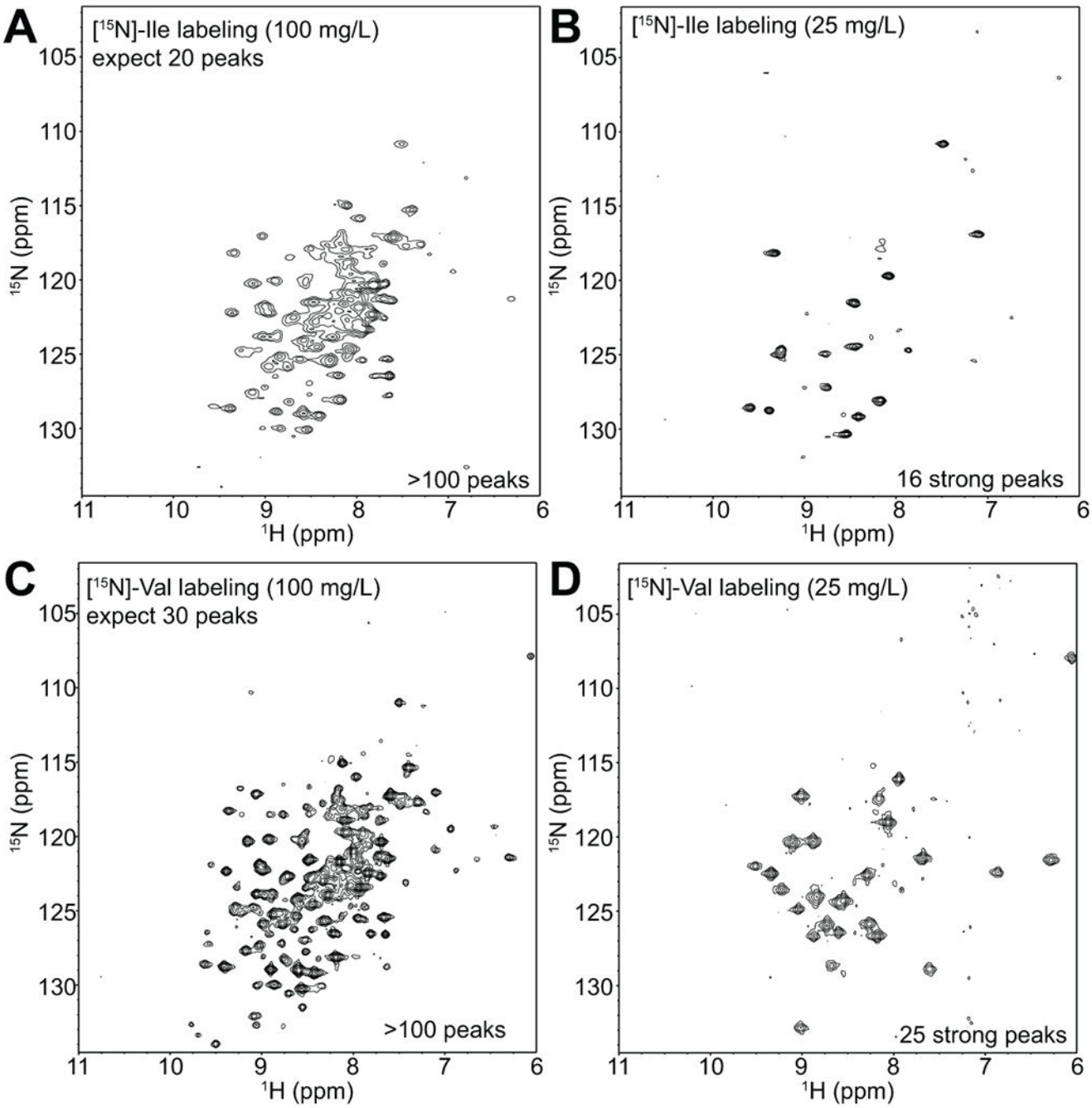


Figure 2. Scrambling of Isoleucine and Valine in GFP-CD16a is reduced by adjusting culture conditions.

Spectra on the left (A,C) are identical to those in Figure 1. Spectra on the right (B,D) were prepared by reducing the levels of the [^{15}N]-labeled amino acid to 25 mg/L. This fusion protein has 20 Ile and 30 Val residues.

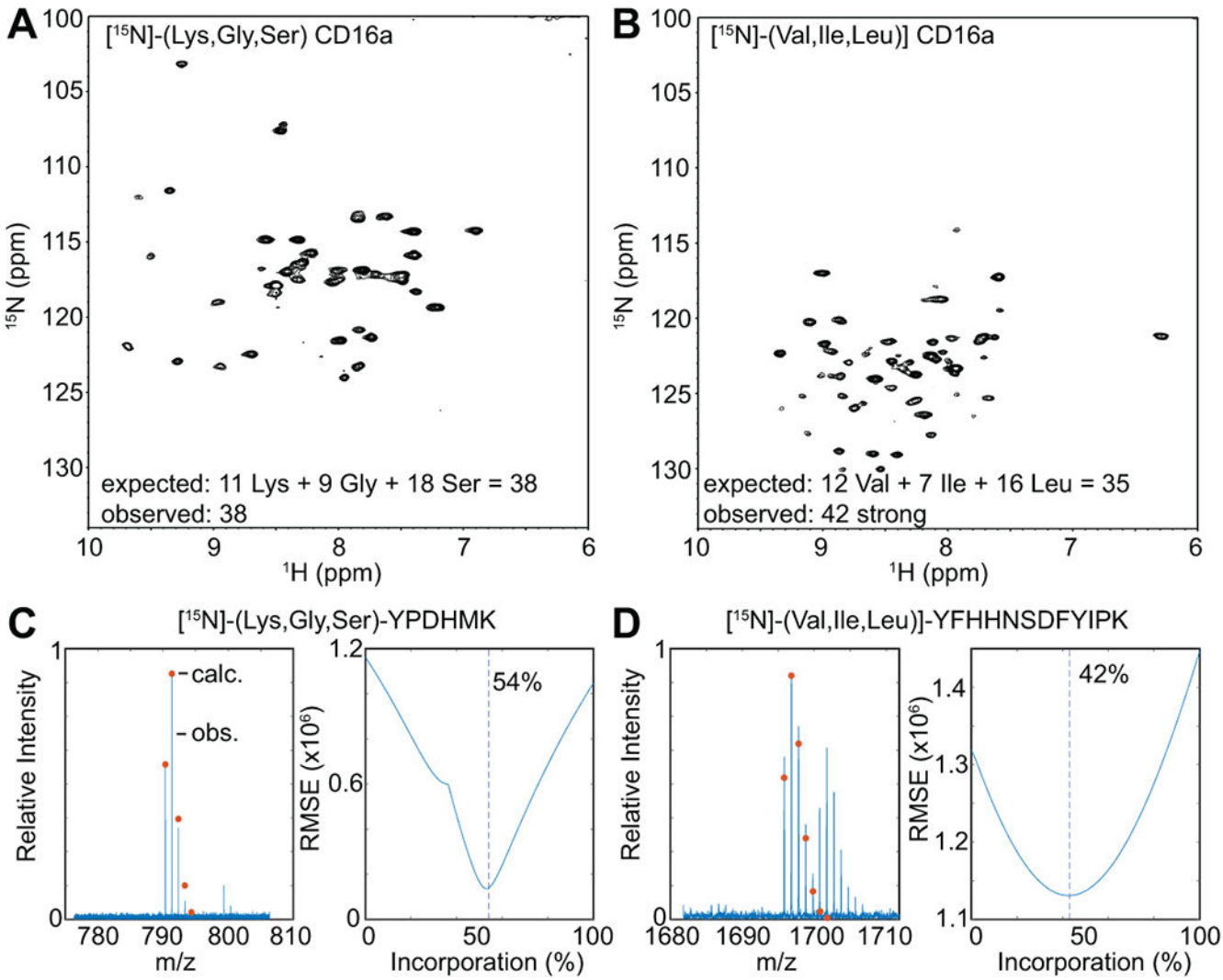


Figure 3. Combining selectively-labeled amino acids increases the number of peaks but retains high sensitivity and resolution with the 20 kDa CD16a.

A. ^{15}N-HSQC-TROSY spectrum collected at 800 MHz. **B.** Val, Ile and Leu interconverted but when simultaneously labeled provide a high quality ^{15}N-HSQC-TROSY spectrum. **C**. and **D**. Example MS spectra (left) with observed (*blue* line) and calculated (*red* dot) isotope distribution for best fit, and calculated root mean square error (right) between the observed signal and calculated isotope distribution as a function of ^{15}N incorporation, corresponding to the peptide mass including one ^{15}N atom per labeled amino acid.