



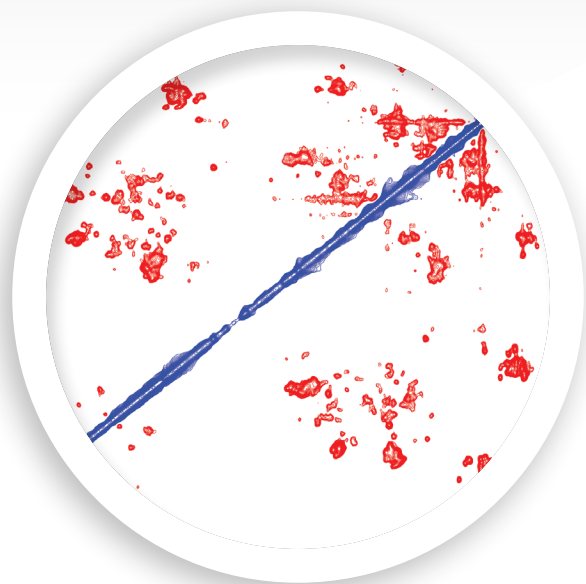
Uniform Isotope Labeling of Eukaryotic Proteins in Methylotrophic Yeast for High-Resolution NMR Studies – Extension to Membrane Proteins

Ying Fan, Lichi Shi, Vladimir Ladizhansky and Leonid S. Brown

Departments of Physics and Biophysics Interdepartmental Group

University of Guelph, 50 Stone Road East

Guelph, Ontario, N1G 2W1 Canada



Recent developments in multidimensional solid-state and solution NMR allow structural studies of new challenging protein targets, such as polytopic membrane proteins.¹⁻³ Especially attractive are medically relevant families of eukaryotic channels, transporters, and receptors, such as GPCRs.^{1,4-6} Unfortunately, uniform isotope labeling of any eukaryotic membrane protein for structural studies by high-resolution NMR can be a daunting task.⁷⁻¹⁰ While the mammalian and insect cell cultures provide native-like membrane environments and allow specific labeling of certain amino acids, relatively low protein yields can lead to extremely high costs for uniformly labeled samples.^{10,11} This is aggravated by the difficulty of protein deuteration, which is often required for structural studies of mid-size and large membrane proteins.¹⁰ The most common alternative, bacterial expression systems, such as *E. coli*, ensure cost-effective uniform labeling and deuteration, but often fail to produce the native fold of eukaryotic membrane proteins, and lack the capability for proper post-translational modifications.^{1,6,8,9} The third possibility, expression in eukaryotic microbes, such as methylotrophic yeast *Pichia pastoris*, has not been fully explored for isotopic labeling of membrane proteins. While the uniform isotopic labeling and deuteration of soluble secreted proteins in *Pichia* is well-established,¹²⁻¹⁶ and many eucaryotic membrane proteins have been expressed in yeast for crystallization, no high-resolution NMR spectra suitable for structural studies have been reported for yeast-expressed eukaryotic membrane proteins. In this application note, we report on the successful extension and modification of the cost-effective uniform double (¹³C, ¹⁵N) labeling protocol, previously employed for soluble secreted proteins,¹⁶ to a eukaryotic seven-transmembrane helical protein.¹⁷

The *Pichia pastoris* expression system is very attractive for isotopic labeling of eukaryotic proteins in general, as it combines native-like cellular environment and post-translational modifications with the ease of handling and cost-effectiveness similar to those of bacterial systems.^{10,14} Importantly, *Pichia* can be grown in fermenters, which can dramatically increase yields of isotopically labeled proteins compared to shake flasks.¹³ The availability of many different strains and vectors allow optimization of selection of best transformants, cellular targeting, and protein tagging for specific proteins.

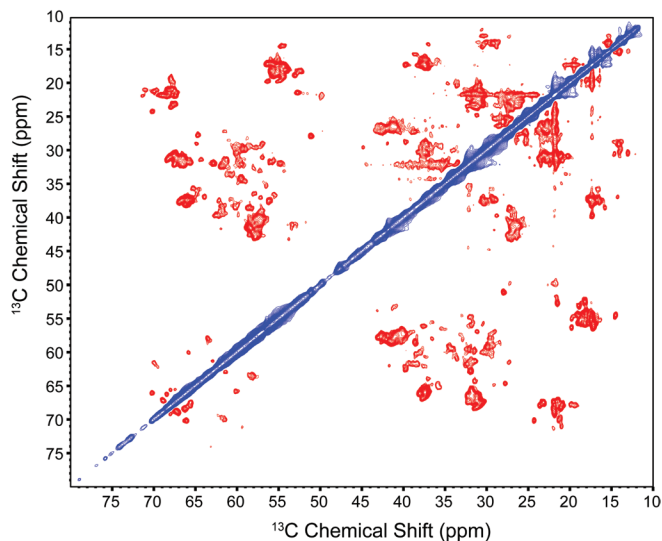


Figure 1. Two-dimensional carbon-carbon chemical shift correlation spectrum of the [U- ^{13}C , ^{15}N] labeled LR at 800 MHz proton frequency with 1.26 ms of SPC5₃ mixing.²⁶ Spinning frequency was 14.3 kHz. Acquisition lengths were 11 ms in t_1 and 20.48 ms in t_2 acquisition dimensions.

The double (^{13}C , ^{15}N) uniform isotope labeling of soluble secreted proteins in *Pichia* relies on a very straightforward protocol, in which the carbon source at the pre-induction growth phase (e.g. $^{13}\text{C}_6$ -glucose) is replaced by ^{13}C -methanol in the post-induction phase.¹⁶ Concentrations and timing of addition of isotope sources have been optimized to ensure the cost-effectiveness and completeness of labeling.¹⁴ In view of high-yield functional expression of many non-labeled mammalian membrane proteins, for example, aquaporins and GPCRs,¹⁸⁻²⁰ in *Pichia*, it is logical to adapt the existing isotope-labeling protocols for these attractive targets.

We have demonstrated the feasibility of such an approach¹⁷ by conducting cost-efficient uniform ^{13}C , ^{15}N -labeling of a ~31 kDa eukaryotic membrane protein, *Leptosphaeria* rhodopsin (LR),^{21,22} which shares seven-transmembrane architecture with GPCRs. The *lop* gene encoding N-terminally truncated, C-terminally 6-His-tagged LR was placed in pPICZ α A vector (Invitrogen), which allowed for selection of multiple integration events on zeocin plates and efficient targeting of the protein-to-plasma membrane using the secretion signal (α -factor). The expression was conducted using the protease-deficient strain SMD1168H (Invitrogen) in shake flasks at 30°C. The pre-induction growth was performed in 250 mL of ^{13}C , ^{15}N -BMD (buffered minimal dextrose) with 0.5% $^{13}\text{C}_6$ -glucose and 0.8% $^{15}\text{NH}_4\text{Cl}$, while the post-induction growth was done in 0.8 L of ^{13}C , ^{15}N -BMM (buffered minimal methanol) with 0.5% ^{13}C -methanol and 0.8% $^{15}\text{NH}_4\text{Cl}$. ^{13}C -methanol was added to the growth medium again after 24 hours of induction, to the final concentration of 0.5%, along with all-*trans*-retinal chromophore needed to regenerate the opsin. After the small-scale screening for the best-producing colonies and optimization of the harvesting time

(40 hours after the induction; shorter and longer times gave substantially lower yields), the yield of the purified protein exceeded 5 mg per liter of culture. As we used only 0.5% concentrations of ^{13}C -glucose and methanol, and the latter had to be replenished only once, the cost of this sample is close to that for similar bacterial proteins produced in *E. coli*.^{23,24}

After breaking the lyticase-treated cells with glass beads, and solubilizing the membranes with Triton X-100, the protein was purified using Ni²⁺-NTA 6-His-tag affinity resin (Qiagen). At this stage, the purity and homogeneity of solubilized LR were estimated spectrophotometrically, by SDS PAGE, and by MALDI TOF mass spectrometry. We found that LR was pure, lacked glycosylation, had the secretion signal cleaved off (with the exception of the last four residues), and had a high extent of isotopic labeling. The extent of isotopic labeling and functionality of the expressed protein were further explored by static and time-resolved Fourier-transform infrared (FTIR) spectroscopy,^{17,25} after reconstitution into the membrane-mimicking DMPC/DMPA liposomes. Static FTIR spectra of the dry proteoliposome films confirmed the high (>90%) overall extent of both ^{13}C and ^{15}N labeling, while time-resolved FTIR spectroscopy of hydrated liposomes showed fully functional photochemistry and proton transfers, along with complete isotope labeling of specific sidechains. The samples were stable for at least several weeks when kept at 4°C.

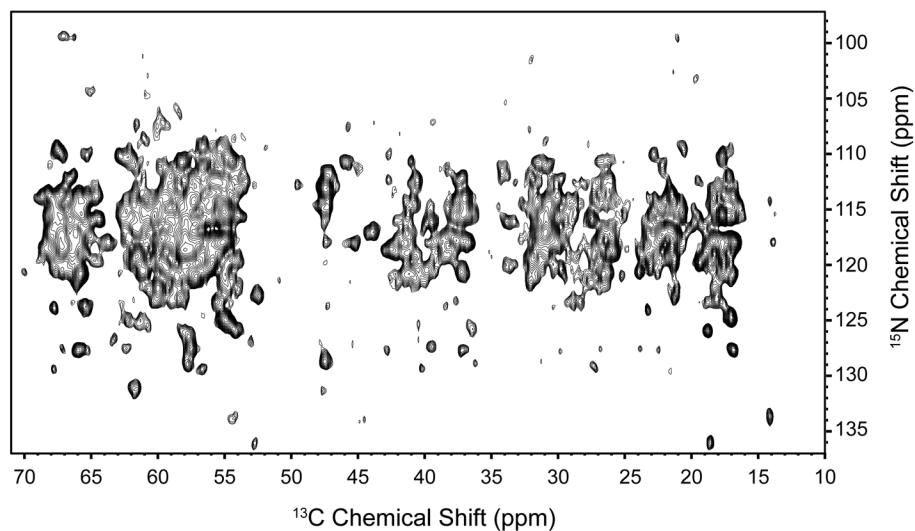


Figure 2. Aliphatic region of two-dimensional NCOCX chemical shift correlation experiment recorder at 600 MHz proton, 12.5 kHz spinning frequency, and with 30 ms of DARR carbon-carbon mixing. Acquisition lengths were 15 ms in t_1 and 22.3 ms in t_2 acquisition dimensions.

Magic angle spinning solid-state NMR measurements, conducted on hydrated LR proteoliposomes using 600 MHz and 800 MHz Bruker instruments, confirmed high structural homogeneity and purity of the sample, along with very low extent of glycosylation. Two-dimensional ^{13}C - ^{13}C SPC-5₃²⁶ and NCOCX correlation spectra (Figures 1 and 2) show many well-resolved resonances with estimated spectral linewidths of ~ 0.5 ppm for ^{13}C , 0.7 ppm for ^{15}N , on par with those observed in the best available samples of bacterial membrane proteins of similar size.^{23,24,27,28} Such spectral resolution allows identification of individual resonances, and is sufficient for conducting spectroscopic assignments by multi-dimensional spectroscopy. Many peaks can be identified by the amino acid type (e.g., prolines, alanines, threonines, serines). Resonances due to carboxylic acids could be resolved in previously published 2D ^{13}C - ^{13}C DARR correlation spectra.¹⁷ Furthermore, some resonances of prolines and protonated aspartates could be tentatively assigned based on the homology with bacteriorhodopsin.^{29,30} The peaks show high dispersion consistent with the co-existence of α -helical transmembrane regions and unstructured and β -stranded loops and tails. Integration of intensities of alanine and glycine peaks show that only a small portion of the resonances are not observed in the 2D ^{13}C - ^{13}C spectra, some of which may not be detectable because of unfavorable dynamics in the long unstructured loops and the C-tail.

In summary, we have demonstrated that cost-effective uniform ^{13}C , ^{15}N labeling of mid-size eukaryotic membrane proteins can be realized in *Pichia pastoris*, using ^{13}C -methanol as a main carbon source. The produced samples are homogeneous, functional, stable, and yield high-resolution solid-state NMR spectra suitable for structural studies. We hope that similar protocols can be adopted for challenging mammalian membrane proteins of medical relevance.

References

1. Kim, H.J.; Howell, S.C.; Van Horn, W.D.; Jeon, Y.H.; Sanders, C.R. **2009**. *Prog Nucl Magn Reson Spectrosc*, *55*, 335.
2. Gautier, A.; Mott, H.R.; Bostock, M.J.; Kirkpatrick, J.P.; Nietlispach, D. **2010**. *Nat Struct Mol Biol*, *17*, 768.
3. Renault, M.; Cukkemane, A.; Baldus, M. **2010**. *Angew Chem Int Ed Engl*, *49*, 8346.
4. Tikhonova, I.G.; Costanzi, S. **2009**. *Curr Pharm Des*, *15*, 4003.
5. Goncalves, J.A.; Ahuja, S.; Erfani, S.; Eilers, M.; Smith, S.O. **2010**. *Prog Nucl Magn Reson Spectrosc*, *57*, 159.
6. Tapaneyyakorn, S.; Goddard, A.D.; Oates, J.; Willis, C.L.; Watts, A. **2011**. *Biochim Biophys Acta*, in press.
7. Lundstrom, K.; Wagner, R.; Reinhart, C.; Desmyter, A.; Cherouati, N.; Magnin, T.; Zeder-Lutz, G.; Courtot, M.; Prual, C.; Andre, N.; Hassaine, G.; Michel, H.; Cambillau, C.; Pattus, F. **2006**. *Journal of structural and functional genomics*, *7*, 77.
8. Sarramegna, V.; Muller, I.; Milon, A.; Talmont, F. **2006**. *Cellular and Molecular Life Sciences*, *63*, 1149.
9. McCusker, E.C.; Bane, S.E.; O'Malley, M.A.; Robinson, A.S. **2007**. *Biotechnol Prog*, *23*, 540.
10. Takahashi, H.; Shimada, I. **2010**. *J Biomol Nmr*, *46*, 3.
11. Egorova-Zachernyuk, T.A.; Bosman, G.J.; Degrip, W.J. **2011**. *Appl Microbiol Biotechnol*, *89*, 397.
12. Laroche, Y.; Storme, V.; De Meutter, J.; Messens, J.; Lauwereys, M. **1994**. *Biotechnology (NY)*, *12*, 1119.
13. Wood, M.J.; Komives, E.A. **1999**. *J Biomol Nmr*, *13*, 149.
14. Rodriguez, E.; Krishna, N.R. **2001**. *J Biochem*, *130*, 19.
15. Morgan, W.D.; Kragt, A.; Feeney, J. **2000**. *J Biomol Nmr*, *17*, 337.
16. Pickford, A.R.; O'Leary, J.M. **2004**. *Methods in molecular biology* (Clifton, NJ), *278*, 17.
17. Fan, Y.; Shi, L.; Ladizhansky, V.; Brown, L.S. **2011**. *J Biomol Nmr*.
18. Oberg, F.; Ekvall, M.; Nyblom, M.; Backmark, A.; Neutze, R.; Hedfalk, K. **2009**. *Mol Membr Biol*, *26*, 215.
19. Singh, S.; Hedley, D.; Kara, E.; Gras, A.; Iwata, S.; Ruprecht, J.; Strange, P.G.; Byrne, B. **2010**. *Protein Expr Purif*, *74*, 80.
20. Andre, N.; Cherouati, N.; Prual, C.; Steffan, T.; Zeder-Lutz, G.; Magnin, T.; Pattus, F.; Michel, H.; Wagner, R.; Reinhart, C. **2006**. *Protein Sci*, *15*, 1115.
21. Idnurm, A.; Howlett, B. J. **2001**. *Genome*, *44*, 167.
22. Waschuk, S.A.; Bezerra, A.G.; Shi, L.; Brown, L.S. **2005**. *P Natl Acad Sci USA*, *102*, 6879.
23. Shi, L.; Ahmed, M.A.; Zhang, W.; Whited, G.; Brown, L.S.; Ladizhansky, V. **2009**. *J Mol Biol*, *386*, 1078.
24. Shi, L.; Kawamura, I.; Jung, K.H.; Brown, L.S.; Ladizhansky, V. **2011**. *Angew Chem Int Ed Engl*, *50*, 1302.
25. Egorova-Zachernyuk, T.A.; Bosman, G.J.; Pistorius, A.M.; DeGrip, W.J. **2009**. *Appl Microbiol Biotechnol*, *84*, 575.
26. Hohwy, M.; Rienstra, C.M.; Griffin, R.G. **2002**. *J Chem Phys*, *117*, 4973.
27. Etzkorn, M.; Martell, S.; Andronesi, O.C.; Seidel, K.; Engelhard, M.; Baldus, M. **2007**. *Angewandte Chemie-International Edition*, *46*, 459.
28. Varga, K.; Aslimovska, L.; Watts, A. **2008**. *J Biomol Nmr*, *41*, 1.
29. Lansing, J.C.; Hu, J.G.G.; Belenky, M.; Griffin, R.G.; Herzfeld, J. **2003**. *Biochemistry-Us*, *42*, 3586.
30. Metz, G.; Siebert, F.; Engelhard, M. **1992**. *Biochemistry-Us*, *31*, 455.

