J|A|C|S COMMUNICATIONS

Published on Web 02/24/2010

Challenging the Limit: NMR Assignment of a 31 kDa Helical Membrane Protein

Chengdong Huang and Smita Mohanty*

Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849

Received January 8, 2010; E-mail: mohansm@auburn.edu

Although predicted to constitute ~30% of the genome and 50% of current drug targets, membrane proteins remain one of the great challenges in structural biology.[1] In the Protein Data Bank (PDB), less than 1% of the atomic structures represent membrane proteins, and this percentage is actually decreasing as more structures of soluble proteins are deposited every day.[2] The primary reason for this disparity is the requirement of lipids (typically detergents) for the solublization of membrane proteins. For X-ray crystallography, detergents often interfere with crystallization,[3] while in NMR spectroscopy, membrane proteins embedded in detergent micelles tumble slowly, leading to rapid transverse relaxation rates, which broaden the resonance line widths and dramatically complicate the NMR spectra. As a result, membrane protein structure determination by NMR analysis has been possible only for some small α-helical proteins[4] and relatively larger outer-membrane bacterial porins having β-barrel folds that provide better amide proton chemical shift dispersion and ample long-range nuclear Overhauser effects (NOEs).[5] The NMR assignment of a 13 kDa α-helical membrane protein that forms a 40 kDa homotrimer has been reported.[6] Recently, a 241-residue helical membrane protein backbone assignment was reported.[7] Here we show that it is possible to use NMR spectroscopy to assign the backbone resonances for a 274 residue (253 residues plus 21 His-tagged residues) monomeric α-helical membrane protein, the C-terminal domain of Stt3p, in detergent micelles.

Oligosaccharyl transferase (OT) is a remarkably complex enzyme that catalyzes N-glycosylation, the most ubiquitous protein modification in eukaryotic cells. In the case of the yeast *Saccharomyces cerevisiae*, OT is composed of nine nonidentical membrane protein subunits, of which five subunits (including Stt3p) are essential for the viability of the cell.[8] Stt3p, the only conserved subunit in the three domains of life,[9] has been shown by several research groups[10] to be the catalytic subunit of the OT complex. To date, only the NMR structure of a 36-residue Ost4p,[11] a nonessential subunit of OT, and the crystal structure of the C-terminal domain of Stt3p from an archaeal source,[12] whose sequence similarity to eukaryotic Stt3p is very limited, have been reported. We have recently overexpressed the His-tagged C-terminal domain of Stt3p (residue 466−718) in *Escherichia coli* to produce very pure homogeneous protein at a very high level and demonstrated that sodium dodecyl sulfate is the optimal detergent for structure determination by high-resolution solution NMR analysis.[13] Here we report the NMR assignment and secondary structure of this 31.5 kDa membrane protein.

All of the NMR experiments were carried out at 55 °C using uniformly ^{2}H,^{13}C,^{15}N-labeled C-terminal domain of Stt3p. Except for HN(CO)CACB, which was acquired in a constant time (CT)-type experiment, all of the experiments, including HNCACB, HNCA, HN(CO)CA, HNCO, and HN(CA)CO, employed transverse-relaxation-optimized spectroscopy (TROSY). The TROSY−HNCACB experiment was recorded with the $^{13}C^{\alpha}-^{13}C^{\beta}$ transfer

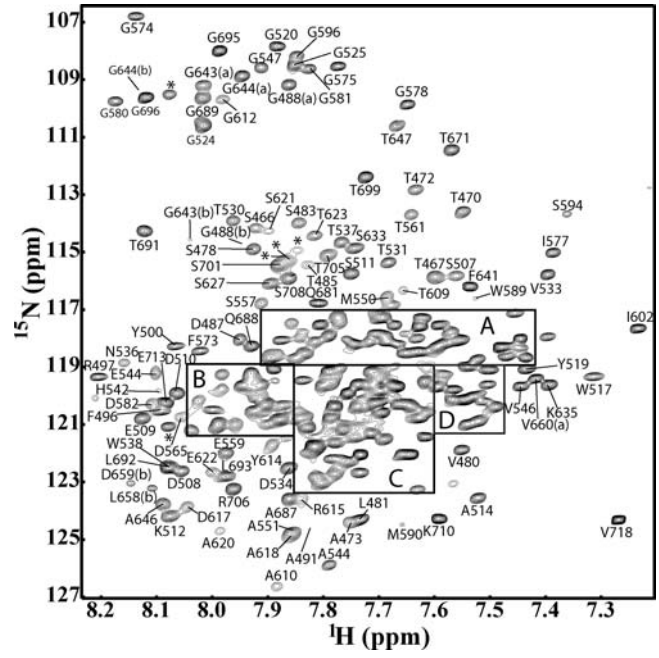

Figure 1. ^{15}N−^{1}HN-TROSY−HSQC spectrum of [U-^{15}N,^{13}C,^{2}H]-labeled C-terminal domain of Stt3p recorded at 55 °C. Four regions of the spectrum are enlarged (see Figure S1), and peaks are labeled with residue numbers. Because of isoasparaginyl and proline cis/trans isomerizational linkage, there are two sets of assignments for a few residues, which are labeled with (a) and (b), respectively. The His-tagged residues are labeled with *.

times optimized for maximum sensitivity of the $^{13}C^{\beta}$ peaks using delays that were less than $1/(2J_{C^{\alpha}C^{\beta}})$. This led to the appearance of typically weak $^{13}C^{\alpha}$ correlations in this spectrum but strong cross-peaks involving $^{13}C^{\beta}$, as expected. Protein stability at elevated temperatures was verified by circular dichroism (CD) and heteronuclear single-quantum correlation (HSQC) NMR experiments. The protein sample was found to be stable at 55 °C for at least 1 month. It is noteworthy that 100% perdeuteration was essential for assigning many of the resonances. It was also found that the TROSY-based experiments offered significant improvements in both resolution and sensitivity in these ^{1}HN−^{15}N correlation-based experiments.

The high content of α-helical secondary structure in this protein (see below), combined with the relatively large number of cross-peaks (263 nonproline residues), resulted in severe overlap in the central part of the HSQC spectrum (Figure 1), making the NMR assignment extremely challenging. The sequential NMR spin-system connectivities were established using [^{15}N−^{1}H]-TROSY−HNCACB and [^{15}N−^{1}H]-TROSY−HNCA (Figure S2 in the Supporting Information), which provided both intra- and inter-residue (sequential) cross-peaks of C^{β} and C^{α}, respectively. Ambiguities were resolved by [^{15}N−^{1}H]-TROSY−HN(CO)CA and CT−HN(CO)-CACB, which provide only sequential cross-peaks. All of the

10.1021/ja100078z

assignments were also confirmed by another complementary pair of experiments: [^{15}N−^{1}H]-TROSY−HNCO and [^{15}N−^{1}H]-TROSY−HN(CA)CO. It is noteworthy that the use of CT−HN-(CO)CACB was found not only to be extremely useful for improving the resolution in the carbon dimension but also, and more importantly, to provide the phase information that was used in the assignment process. In the CT−HN(CO)CACB experiment, the C$^{\beta}$ resonances of residues with an odd number of aliphatic carbons attached have sign opposite to that of the C$^{\beta}$ resonances of residues with an even number of attached aliphatic carbons.[14] On the basis of the fact that the number of amino acids containing odd numbers of aliphatic carbons attached to C$^{\beta}$ is approximately the same as that of those amino acids with even numbers of aliphatic carbons attached to C$^{\beta}$, the sign of the cross peaks of C$^{\beta}$ significantly facilitates the resolution of ambiguities during assignments. Approximately 93% of the backbone resonance assignments (255 out of 274 residues) were completed for the C-terminal domain of Stt3p.

Interestingly, during the course of the assignment, an isoaspartyl linkage in the protein sequence IsoAsp^{642}−Gly^{643}, which is an isomerized form of the deamidated Asn^{642}−Gly^{643} connection (β-linked peptide), was unambiguously identified on the basis of the fact that in the CT−HN(CO)CACB spectrum, the cross-peaks involving C$^{\alpha}$ and C$^{\beta}$ have opposite signs (Figure S3). Extensive studies of asparaginyl deamidation in proteins have shown that this nonenzymatic post-translational modification may play an important role in protein stability and have a significant impact on protein structure and/or function.[15] Another interesting finding is the presence of a proline cis/trans isomerizational linkage in this protein. In fact, there were two sets of cross-peaks assigned for residues between Leu^{658} and Val^{660} (Figure S4), suggesting that the peptide bond Val^{660}−Pro^{661} adopts both cis and trans conformations. The question as to whether this proline cis/trans isomerization plays an important role can be addressed only by further studies.

The secondary structure of the C-terminal domain of Stt3p was determined on the basis of the chemical shift index (CSI), which is the deviation of the C$^{\alpha}$ and C$^{\beta}$ chemical shifts from mean random coil values that have been corrected for deuterium isotope effects.[16] In Figure 2, the parameter ($\Delta C^{\alpha} - \Delta C^{\beta}$) is plotted versus the residue number. The NMR-based secondary structure results are consistent with the far-UV CD spectroscopy data, which also indicate that the C-terminal domain of Stt3p is highly helical.[13]

In summary, we have presented nearly complete backbone (^{1}HN, ^{15}N, ^{13}CO, and ^{13}C$^{\alpha}$) and side-chain ^{13}C$^{\beta}$ chemical shift assignments and the secondary structure for the His-tagged C-terminal domain of Stt3p, a 31.5 kDa, 274-residue helical membrane protein. The completion of the majority of the NMR resonance assignments demonstrates the feasibility of 3D structure determination for this 31.5 kDa monomeric helical membrane protein by solution NMR methods. Complete side-chain assignments and determination of the 3D structure of the C-terminal domain of Stt3p are progressing very well with the use of selectively protonated samples, such as methyl-protonated [Ile(δ1 only), Leu($^{13}CH_3$, $^{12}CD_3$), Val($^{13}CH_3$, $^{12}CD_3$)] [U-^{15}N,^{13}C,^{2}H]-labeled samples generated using biosynthetic precursors.[17]

Acknowledgment. This research was financially supported by NIH Grant DK082397 and NSF Grant IBN-0628064 (to S.M). C.H. was financially supported by an Alabama ACHE-GRSP Scholarship.

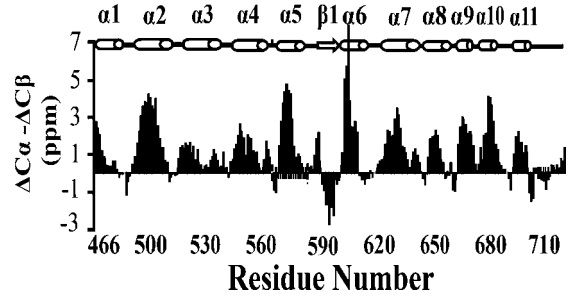

Figure 2. $\Delta C^{\alpha} - \Delta C^{\beta}$ as a function of residue number. ΔC^{α} and ΔC^{β} are the C$^{\alpha}$ and C$^{\beta}$ chemical shifts observed for the protein subtracted from the C$^{\alpha}$ and C$^{\beta}$ random-coil values, respectively. The His-tagged residues were excluded from the analysis.

The 900 MHz NMR data were collected at the NMR facility of the University of Georgia, which is funded by NIH Grant GM66340.

Note Added after ASAP Publication. In the original version of the article published on the Web on February 24, 2010, two of the residue labels in Figure 1 were incorrect. The corrected version was reposted on February 25, 2010.

Supporting Information Available: Figures S1−S4 and experimental details for sample preparation and NMR data collection. This material is available free of charge via the Internet at http://pubs.acs.org. The assigned backbone chemical shifts have been deposited in the BioMagResBank (BMRB accession number 16701).

References

(1) Korepanova, A.; Gao, F. P.; Hua, Y.; Nakamoto, R. K.; Cross, T. A. *Protein Sci.* **2005**, *14*, 148.
(2) Gao, F. P.; Cross, T. A. *Genome Biol.* **2005**, *6*, 244.
(3) (a) Sowadski, J. M. *J. Bioenerg. Biomembr.* **1996**, *28*, 3. (b) Ostermeier, C.; Michel, H. *Curr. Opin. Struct. Biol.* **1997**, *7*, 697. (c) Roosild, T. P.; Greenwald, J.; Vega, M.; Castronovo, S.; Riek, R.; Choe, S. *Science* **2005**, *307*, 1317.
(4) (a) Rastogi, V. K.; Girvin, M. E. *Nature* **1999**, *402*, 263. (b) MacKenzie, K. R.; Prestegard, J. H.; Engelman, D. M. *Science* **1997**, *276*, 131.
(5) (a) Fernandez, C.; Hilty, C.; Wider, G.; Guntert, P.; Wuthrich, K. *J. Mol. Biol.* **2004**, *336*, 1211. (b) Hwang, P. W.; Choy, W.; Lo, E. I.; Chen, L.; Forman-Kay, J. D.; Raetz, C. R.; Prive, G. G.; Bishop, R. E.; Kay, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 13560. (c) Arora, A.; Abildgaard, F.; Bushweller, J. H.; Tamm, L. K. *Nat. Struct. Biol.* **2001**, *8*, 334.
(6) Oxenoid, K.; Kim, H. J.; Jacob, J.; Sönnichsen, F. D.; Sanders, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 5048.
(7) Gautier, A.; Kirkpatrick, J. P.; Nietlispach, D. *Angew. Chem., Int. Ed.* **2008**, *47*, 7297.
(8) Yan, A.; Lennarz, W. J. *J. Biol. Chem.* **2005**, *280*, 3121.
(9) Zufferey, R.; Knauer, R.; Burda, P.; Stagljar, I.; te Heesen, S.; Lehle, L.; Aebi, M. *EMBO J.* **1995**, *14*, 4949.
(10) (a) Yan, Q.; Lennarz, W. J. *J. Biol. Chem.* **2002**, *277*, 47692. (b) Nilsson, I.; Kelleher, D. J.; Miao, Y.; Shao, Y.; Kreibich, G.; Gilmore, R.; von Heijne, G.; Johnson, A. E. *J. Cell Biol.* **2003**, *161*, 715. (c) Glover, K. J.; Weerapana, E.; Numao, S.; Imperiali, B. *Chem. Biol.* **2005**, *12*, 1311. (d) Feldman, M. F.; Wacker, M.; Hernandez, M.; Hitchen, P. G.; Marolda, C. L.; Kowarik, M.; Morris, H. R.; Dell, A.; Valvano, M. A.; Aebi, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 3016. (e) Nasab, F. P.; Schulz, B. L.; Gamarro, F.; Parodi, A. J.; Aebi, M. *Mol. Biol. Cell* **2008**, *19*, 3758.
(11) Zubkov, S.; Lennarz, W. J.; Mohanty, S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 3821.
(12) Igura, M.; Maita, N.; Kamishikiryo, J.; Yamada, M.; Obital, T.; Maenaka, K.; Kohda, D. *EMBO J.* **2008**, *27*, 234.
(13) Huang, C.; Mohanty, S.; Banerjee, M. *Biochemistry* **2010**, *49*, 1115.
(14) Shan, X.; Gardner, K. H.; Muhandiram, D. R.; Rao, N. S.; Arrowsmith, C. H.; Kay, L. E. *J. Am. Chem. Soc.* **1996**, *118*, 6570.
(15) (a) Manning, M. C.; Patel, K.; Borchard, R. T. *Pharm. Res.* **1989**, *6*, 903. (b) Aswad, D. W.; Paranandi, M. V.; Schuter, B. T. *J. Pharm. Biomed. Anal.* **2000**, *21*, 129.
(16) (a) Venters, R. A.; Farmer, B. T.; Fierke, C. A.; Spicer, L. D. *J. Mol. Biol.* **1996**, *264*, 1101. (b) Wishart, D. S.; Sykes, B. D. *J. Biomol. NMR* **1994**, *4*, 171.
(17) Tugarinov, V.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 13868.

JA100078Z