



Published in final edited form as:

*J Am Chem Soc.* 2010 March 24; 132(11): 3662–3663. doi:10.1021/ja100078z.

## 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

Smita Mohanty: monhansm@auburn.edu

### Abstract

Structural determination of membrane proteins by NMR, especially for helical membrane proteins remains a challenge. Here we report the NMR assignment and secondary structure of a 31 kDa helical membrane protein, the C-terminal domain of Stt3p. The C-terminal domain of Stt3p has been proposed to be the catalytic domain of the yeast Oligosaccharyl Transferase (OT), a multi-subunit membrane-associated enzyme complex catalyzing N-glycosylation, an essential and highly conserved protein modification. NMR assignment is the first critical step in the determination of high-resolution solution structure and for further structure-function studies.

Although predicted to constitute about 30% of the genome and 50% of current drug targets, membrane proteins remain one of the great challenges in structural biology.<sup>1</sup> In the Protein Data Bank (PDB), less than 1% of the atomic structures represent that of membrane proteins, and this percentage is actually decreasing since more structures of soluble proteins are being deposited everyday.<sup>2</sup> The primary reason for this disparity is the requirement of lipids (typically detergents) for the solubilization of membrane proteins. For X-ray crystallography, detergents often interfere with crystallization,<sup>3</sup> while for NMR, 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 have only been possible for some small  $\alpha$ -helical proteins,<sup>4</sup> and relatively larger outer membrane bacterial porins having  $\beta$ -barrel folds that provide better amide proton chemical shift dispersion and ample long-range nuclear Overhauser effects (NOEs).<sup>5</sup> The NMR assignment of a 13 kDa  $\alpha$ -helical membrane protein that forms 40 kDa homotrimer has been reported.<sup>6</sup> Recently, a 241-residue helical membrane protein backbone assignment was reported.<sup>7</sup> Here, we show that it is possible to use NMR to assign the backbone resonances for a 274-residue (253 residues plus 21 His-tagged residues) monomeric  $\alpha$ -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 *S. cerevisiae*, OT is composed of nine nonidentical membrane protein subunits, of which, five subunits including Stt3p are essential for the viability of the cell.<sup>8</sup> Stt3p, the only conserved subunit in the three domains of life,<sup>9</sup> has been shown to be the catalytic subunit of the OT complex by several research groups.<sup>10</sup> To date, only the NMR structure of a 36-residue Ost4p,<sup>11</sup> a non-essential subunit of OT, and the crystal structure of the C-terminal domain of Stt3p from an archaeal source,<sup>12</sup> that has very limited sequence similarities to the eukaryotic Stt3p,

Correspondence to: Smita Mohanty, monhansm@auburn.edu.

**Supporting Information Available:** Figures 1S-4S, 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>.

have been reported. We have recently over-expressed the His-tagged C-terminal domain of Stt3p (residue 466-718) in *E. coli* and produced very pure homogeneous protein at very high level and have demonstrated that SDS is the optimal detergent for structure determination by high-resolution solution NMR.<sup>13</sup> Here we report the NMR assignment and secondary structure of this 31.5 kDa membrane protein.

All NMR experiments were carried out at 55 °C using uniformly <sup>2</sup>H, <sup>13</sup>C, <sup>15</sup>N-labeled C-terminal domain of Stt3p. Except the HN(CO)CACB which was acquired as constant time (CT) type experiment, all experiments, including HNCACB, HNCA, HN(CO)CA, HNCO, HN(CA)CO, were collected as TROSY-based. TROSY-HNCACB experiment was recorded with the <sup>13</sup>C<sup>α</sup>-<sup>13</sup>C<sup>β</sup> transfer times optimized for maximum sensitivity of <sup>13</sup>C<sup>β</sup> peaks using delays which were less than 1/(2JC<sup>α</sup>C<sup>β</sup>). This led to the appearance of typically weak <sup>13</sup>C<sup>α</sup> correlations in these spectra, but strong cross-peaks involving the <sup>13</sup>C<sup>β</sup> as expected. Protein stability at elevated temperatures was verified by Circular Dichroism (CD) and HSQC NMR experiments. Protein sample was found to be stable at 55 °C for at least one month. It is noteworthy that 100% perdeuteration was essential for assigning many resonances. It was also found that the TROSY-based experiments offer significant improvements in both resolution and sensitivity in these <sup>1</sup>HN-<sup>15</sup>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 non-proline residues), resulted in severe overlap in the central part of the HSQC spectra (Figure 1), making NMR assignment extremely challenging. The sequential NMR spin system connectivities were established using [<sup>15</sup>N-<sup>1</sup>H]-TROSY-HNCACB and [<sup>15</sup>N-<sup>1</sup>H]-TROSY-HNCA (Figure 2S), which provided both intra- and inter-residue (sequential) cross-peaks of C<sup>β</sup> and C<sup>α</sup> respectively. Ambiguities were resolved by [<sup>15</sup>N-<sup>1</sup>H]-TROSY-HN(CO)CA and CT-HN(CO)CACB, which provides only sequential cross-peaks. All the assignments were also confirmed by another complementary pair experiments: [<sup>15</sup>N-<sup>1</sup>H]-TROSY-HNCO and [<sup>15</sup>N-<sup>1</sup>H]-TROSY-HN(CA)CO. It is noteworthy that, the use of CT-HN(CO)CACB was found to be extremely useful not only for the improved resolution in the carbon dimension, but more importantly, it provided the phase information, which was used in the assignment process. In the CT-HN(CO)CACB experiment, the C<sup>β</sup> resonance of residues with an odd number of aliphatic carbons attached has opposite sign to that of the C<sup>β</sup> resonance of residues with an even number of attached aliphatic carbons.<sup>14</sup> Based on the fact that the number of amino acids containing odd sets of aliphatic carbons attached to the C<sup>β</sup> is approximately same as that of those amino acids with even sets of aliphatic carbons attached to the C<sup>β</sup>, the sign of the cross peaks of C<sup>β</sup> significantly facilitates in resolving ambiguities during assignments. Approximately 93% (255 out of 274 residues) backbone resonance assignments were completed for the C-terminal domain of Stt3p.

Interestingly, during the course of the assignment, an isoaspartyl linkage in the protein sequence, -IsoAsp<sup>642</sup>-Gly<sup>643</sup>, which is an isomerized form of deamidated Asn<sup>642</sup>-Gly<sup>643</sup> connection (β-linked peptide), was unambiguously identified based on the fact that in the CT-HN(CO)CACB spectra, the cross peaks involving C<sup>α</sup> and C<sup>β</sup> are of opposite signs (Figure 3S). Asparaginyl deamidation in proteins has been studied extensively, showing that this nonenzymatic post-translational modification may play an important role in protein stability and have significant impact on protein structure and/or function.<sup>15</sup> Another interesting finding is the presence of proline *cis/trans* isomerizational linkage in this protein. In fact, there were two sets of cross peaks that were assigned for residues between Leu<sup>658</sup> and Pro<sup>661</sup> (Figure 4S); suggesting the peptide bond Val<sup>660</sup>-Pro<sup>661</sup> adopts both *cis* and *trans* conformations. The question as to whether this proline *cis/trans* isomerization plays an important role, can only be addressed 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 deviations of  $C^\alpha$  and  $C^\beta$  chemical shifts from mean random coil values that have been corrected for deuterium isotope effects.<sup>16</sup> In Figure 2, the parameter ( $\Delta C^\alpha - \Delta C^\beta$ ) is plotted versus the protein sequence. The NMR-based secondary structure result is consistent with the far-UV CD spectroscopy data, which also indicates that the C-terminal domain of Stt3p is highly helical.<sup>13</sup>

In summary, we have presented nearly complete backbone ( $^1\text{HN}$ ,  $^{15}\text{N}$ ,  $^{13}\text{CO}$  and  $^{13}\text{C}^\alpha$ ) and side-chain  $C^\beta$  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( $\delta 1$  only), Leu( $^{13}\text{CH}_3$ ,  $^{12}\text{CD}_3$ ), Val( $^{13}\text{CH}_3$ ,  $^{12}\text{CD}_3$ )} U- $\{^{15}\text{N}$ ,  $^{13}\text{C}$ ,  $^2\text{H}\}$  that were generated using biosynthetic precursors.<sup>17</sup>

The assigned backbone chemical shifts have been deposited in the BioMagResBank (BMRB accession number of 16701).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

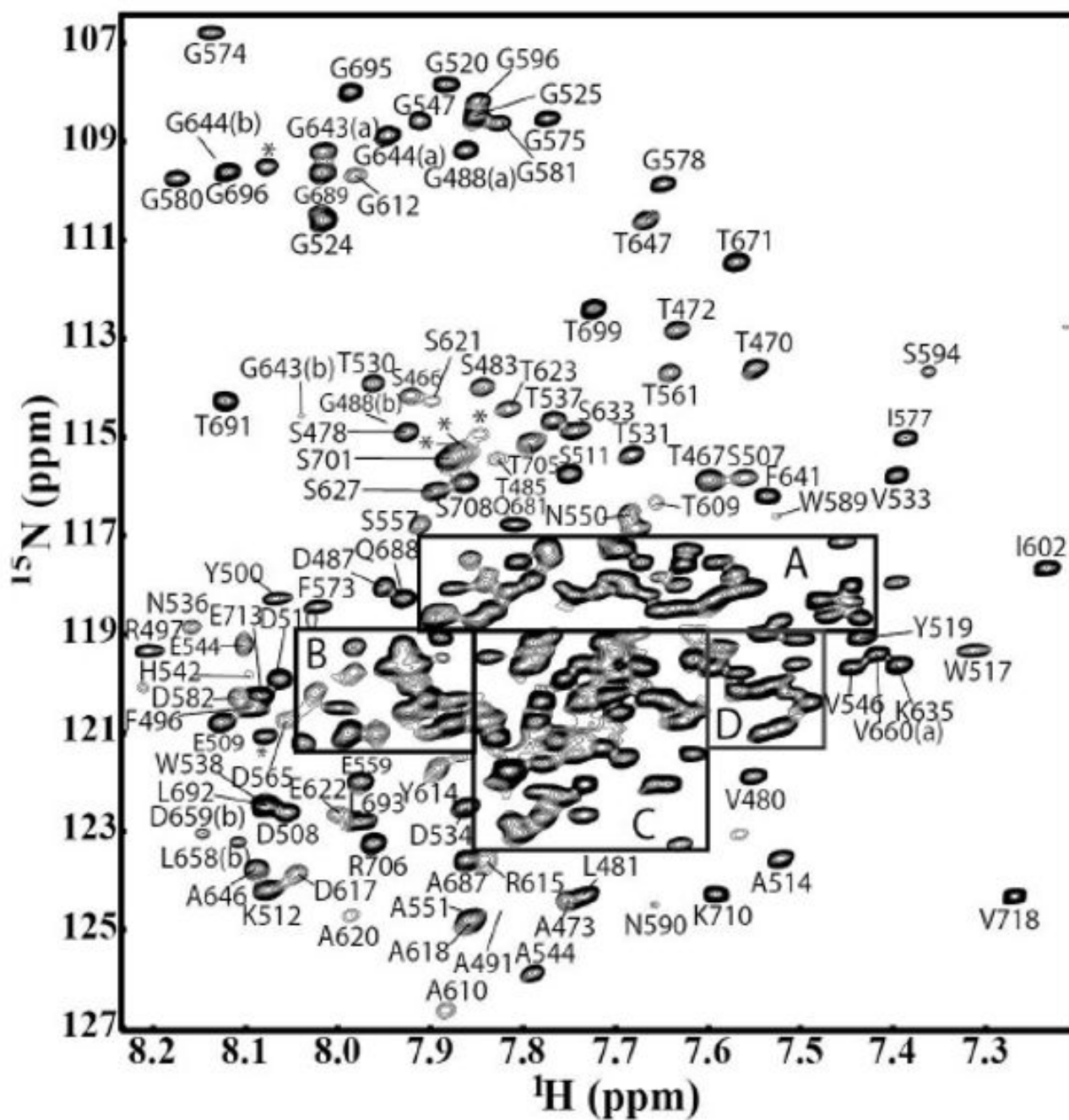
## Acknowledgments

This research was financially supported by NIH grant DK082397 and NSF grant IBN-0628064 (to S. M). C. H. was financially supported by Alabama ACHE-GRSP Scholarship. The 900 MHz NMR data were collected at the NMR facility of the University of Georgia, funded by NIH grant GM66340.

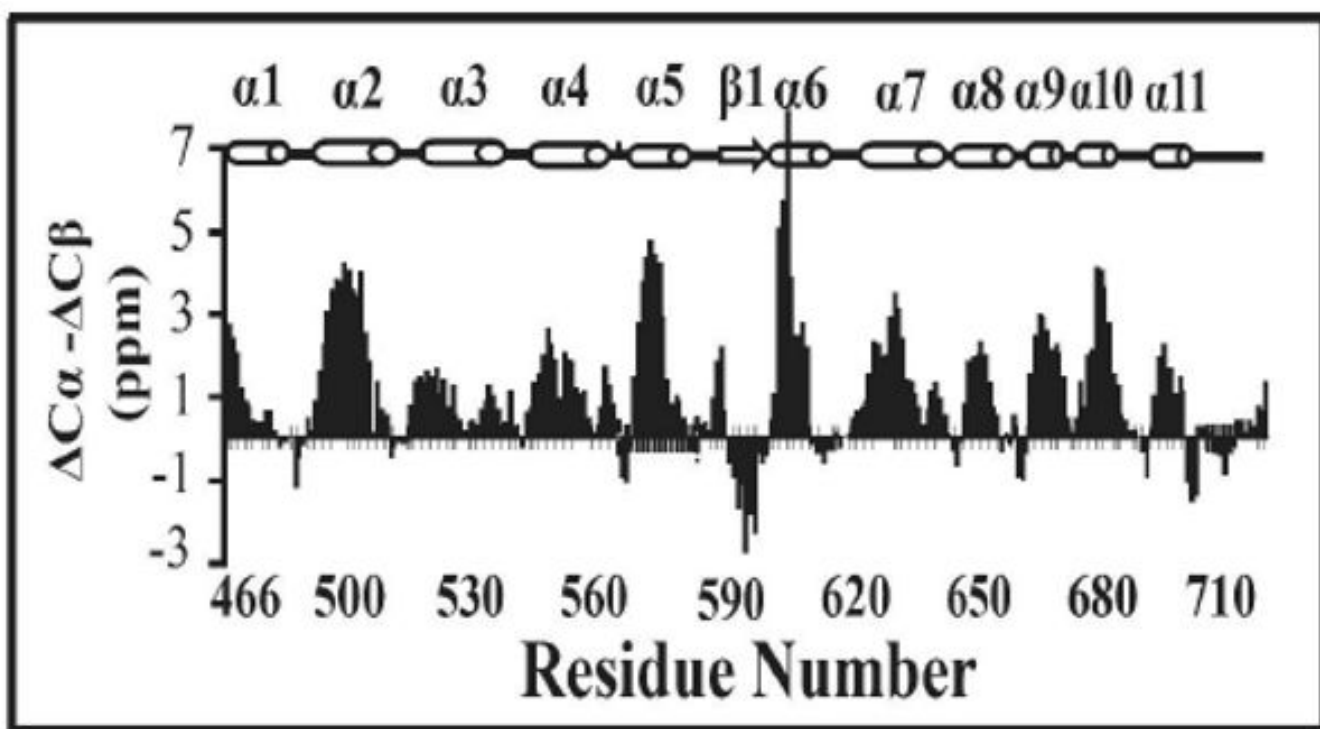
## References

1. Korepanova A, Gao FP, Hua Y, Nakamoto RK, Cross TA. *Protein Sci* 2005;14:148–158. [PubMed: 15608119]
2. Gao FP, Cross TA. *Genome Biology* 2005;6:244. [PubMed: 16420681]
3. (a) Sowadski MJ. *Bioenerg Biomembr* 1996;28:3–5. (b) Ostermeier C, Michel H. *Curr Opin Struct Biol* 1997;7:697–701. [PubMed: 9345629] (c) Roosild TP, Greenwald J, Vega M, Castronovo S, Riek R, Choe S. *Science* 2005;307:1317–1321. [PubMed: 15731457]
4. (a) Rastogi VK, Girvin ME. *Nature* 1999;402:263–268. [PubMed: 10580496] (b) MacKenzie KR, Prestegard JH, Engelman DM. *Science* 1997;276:131–133. [PubMed: 9082985]
5. (a) Fernandez C, Hilty C, Wider G, Guntert P, Wuthrich K. *J Mol Biol* 2004;336:1211–1221. [PubMed: 15037080] (b) Hwang PW, Choy W, Lo EI, Chen L, Forman-Kay JD, Raetz CR, Prive GG, Bishop RE, Kay LE. *Proc Natl Acad Sci* 2002;99:13560–13565. [PubMed: 12357033] (c) Arora A, Abildgaard F, Bushweller JH, Tamm LK. *Nature Struct Biol* 2001;8:334. [PubMed: 11276254]
6. Oxenoid K, Kim HJ, Jacob J, Snnichsen FD, Sanders CR. *J Am Chem Soc* 2004;126:5048–5049. [PubMed: 15099070]
7. Gautier A, Kirkpatrick JP, Nietlispach D. *Angew Chem Int Ed* 2008;47:7297–7300.
8. Yan A, Lennarz WJ. *J Biol Chem* 2005;280:3121–3124. [PubMed: 15590627]
9. Zufferey R, Knauer R, Burda P, Stagljar I, te Heesen S, Lehle L, Aebi M. *EMBO J* 1995;14:4949–4960. [PubMed: 7588624]
10. (a) Yan Q, Lennarz WJ. *J Biol Chem* 2002;277:47692–47700. [PubMed: 12359722] (b) Nilsson I, Kelleher DJ, Miao Y, Shao Y, Kreibich G, Gilmore R, von Heijne G, Johnson AE. *J Cell Biol* 2003;161:715–725. [PubMed: 12756234] (c) Glover KJ, Weerapana E, Numao S, Imperiali B. *Chem Biol* 2005;12:1311–1315. [PubMed: 16356848] (d) Feldman MF, Wacker M, Hernandez M, Hitchen

- PG, Marolda CL, Kowarik M, Morris HR, Dell A, Valvano MA, Aebi M. *Proc Natl Acad Sci* 2005;102:3016–3021. [PubMed: 15703289] (e) Nasab FP, Schulz BL, Gamarro F, Parodi AJ, Aebi M. *Mol Biol Cell* 2008;19:3758–3768. [PubMed: 18596231]
11. Zubkov S, Lennarz WJ, Mohanty S. *Proc Natl Acad Sci* 2004;101:3821–3826. [PubMed: 15001703]
  12. Igura M, Maita N, Kamishikiryo J, Yamada M, Obital T, Maenaka K, Kohda D. *EMBO J* 2008;27:234–243. [PubMed: 18046457]
  13. Huang C, Mohanty S, Banerjee M. *Biochemistry*. 2010 in press.
  14. Shan X, Gardner KH, Muhandiram DR, Rao NS, Arrowsmith CH, Kay LE. *J Am Chem Soc* 1996;118:6570–6579.
  15. (a) Manning MC, Patel K, Borchard RT. *Pharm Res* 1989;6:903–918. [PubMed: 2687836] (b) Aswad DW, Paranandi MV, Schuter BT. *J Pharm Biomed Anal* 2000;21:129–136.
  16. (a) Venters RA, Farmer BT, Fierke CA, Spicer LD. *J Mol Biol* 1996;264:1101–1116. [PubMed: 9000633] (b) Wishart DS, Sykes BD. *J Biomol NMR* 1994;4:171–180. [PubMed: 8019132]
  17. Tugarinov V, Kay LE. *J Am Chem Soc* 2003;125:13868–13878. [PubMed: 14599227]



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



**Figure 2.**  
(A)  $\Delta C\alpha$  and  $\Delta C\beta$  are chemical shifts observed for the protein subtracted from the  $C\alpha$  and  $C\beta$  random-coil values respectively. The His-tagged residues were excluded for analysis.