Eukaryotic N-glycosylation Occurs Via Membrane-Anchored C-terminal domain of Stt3p subunit of Oligosaccharyl Transferase*

Chengdong Huang¹, Rajagopalan Bhaskaran¹, and Smita Mohanty¹

¹From the Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849, USA

* Running title: NMR structure of the C-terminal domain of yeast Stt3p

To whom correspondence should be addressed: Smita Mohanty, Department of Chemistry and Biochemistry, Auburn University, 179 Chemistry Building, Mell Street, Auburn, AL 36849. Tel.: (334) 826-7980; Fax: (334) 844-6959; E-mail: mohansm@auburn.edu.

Keywords: NMR, Structure, Stt3p, Oligosaccharyl Transferase, N-glycosylation

Background: Stt3p is the catalytic subunit of Oligosaccharyl Transferase (OT) that catalyses protein N-glycosylation.
Results: We report the first high-resolution NMR structure of the acceptor binding domain of yeast OT.
Conclusion: This work provides structural basis for the function of the C-terminal Domain of Stt3p subunit.
Significance: Structure determination of this critical domain is an important step towards understanding the mechanisms of the eukaryotic N-glycosylation process.

SUMMARY

N-glycosylation is an essential and highly conserved protein modification. In eukaryotes, it is catalyzed by a multisubunit membrane associated enzyme, Oligosaccharyl Transferase (OT). We report the high-resolution structure of the C-terminal domain of eukaryotic Stt3p. Unlike its soluble β-sheet rich prokaryotic counterparts, our model reveals that the C-terminal domain of yeast Stt3p is highly helical and has an overall oblate spheroid shaped structure containing a membrane-embedded region. Anchoring of this protein segment to the ER membrane is likely to bring the membrane embedded donor substrate closer, thus facilitating the glycosylation efficiency. Structural comparison of the region near the WWDYG signature motif reveals that the acceptor substrate binding site of yeast OT strikingly resembles to its prokaryotic counterparts, suggesting a conserved mechanism of N-glycosylation from prokaryote to eukaryote. Furthermore, comparison of the NMR and cryo-EM structures of the yeast OT reveals that the molecular architecture of this acceptor substrate recognizing domain has interesting spatial specificity for interactions with other essential OT subunits.

Oligosaccharyl transferase (OT) catalyzes N-glycosylation, the most ubiquitous co-translational protein modification in all eukaryotes. In the central reaction, a preassembled oligosaccharide moiety is transferred from a dolichol carrier (Dol-PP-oligosaccharide) to an asparagine residue in the nascent polypeptide chain defined by the consensus sequon N-X-T/S (X
≠ P) (1,2). The transferred N-glycans play essential roles in modulating protein function in many critical processes including development, inflammation, cancer, and the immune response (3,4). For many higher eukaryotes, protein N-glycosylation is a highly coordinated and complex process. In the most extensively studied eukaryotic system, *Saccharomyces cerevisiae*, OT contains nine non-identical integral membrane protein (IMP) subunits: Ost1p-Ost6p, Stt3p, Wbp1p and Swp1p (2). Among these, five subunits (Stt3p, Wbp1p, Swp1p, Ost1p, and Ost2p) are essential for the viability of the cell, whereas Ost3p and Ost6p are homologous, interchangeable subunits (5). In contrast, a single protein chain defines the OT activity in prokaryotes emphasizing the complexity of the glycosylation process in eukaryotes.

Although the detailed enzymatic reaction mechanism and the roles of the other subunits are not yet fully understood, a multitude of experimental evidence has suggested that Stt3p contains the catalytic site (6-8). As shown in Figure S-1, eukaryotic Stt3p is highly conserved. In fact, it is the most conserved subunit in the OT complex (6). The most convincing evidence favoring this conclusion has come from the composition of the OT of some lower eukaryotes such as trypanosomatids (9), as well as a few prokaryotes in which N-glycosylation occurs. In these organisms, OT is a single integral membrane protein homologous to the Stt3p subunit, such as PglB proteins in the bacterium *Campylobacter jejuni* (10) and *Campylobacter lari* (11), or AglB protein in the archaeon *Pyrococcus furiosus* (12). Further studies showed that when transferred into *Escherichia coli* or *stt3*-deficient yeast cells, these single-polypeptide OTs can either enable or complement N-glycosylation in these host cells (13,14). Moreover, it has also been demonstrated that the introduction of *Leishmania major* Stt3p into yeast can replace the entire OT complex (15).

Despite the biological significance of N-glycosylation, very limited success has been achieved with regard to the structural investigation of eukaryotic OT. The scarcity of structural knowledge significantly hampers our understanding of the enzymatic mechanism of OT complex. To date, there are only three structural reports at atomic resolution in terms of eukaryotic OT subunits: NMR structures of the minimembrane protein Ost4p from yeast (16) and human (17), and the crystal structure of the N-terminal luminal domain of Ost3p/Ost6p (18). The low-resolution structure (at 12 Å) of the yeast OT complex was also determined by electron microscopy (EM) techniques (19). In comparison to the eukaryotic OT, the structures of prokaryotic OT are relatively well understood. Two crystal structures of the soluble C-terminal domain of the prokaryotic homologs of Stt3p have been determined in succession: the AglB protein of archaeon *P. furiosus* (20), and the PglB protein of bacterium *C. jejuni* (12). Recently, the crystal structure of a full length bacterial Stt3p homolog, the PglB protein of *Campylobacter lari*, was determined in complex with an acceptor peptide (21). However, sequence alignment reveals that the bacterial and archaean Stt3p homologs show very limited sequence similarities to the eukaryotic Stt3p.

We have previously reported the preparation and nuclear magnetic resonance (NMR) assignments of the 274 residues C-terminal domain of Stt3p (residue 466-718 plus 21 His-tagged residues) in detergent micelles (22,23). Our results revealed that the C-terminal domain of yeast Stt3p is highly helical. Here, we present the NMR structure of this 31.5 kDa helical membrane protein fragment in detergent micelles. Our results represent the feasibility of *de novo* structure determination of a medium to high molecular weight helical membrane protein fragment.
using solution NMR methodologies. Structural comparison of the acceptor substrate binding site of yeast OT to that of its prokaryotic homologs suggest an evolutionary conserved N-glycosylation mechanism. Furthermore, fitting the NMR structure of the C-terminal domain of Stt3p to the EM model provides insight into the interaction between Stt3p and other OT subunits.

EXPERIMENTAL PROCEDURES

Protein expression and purification—Protein samples prepared for side-chain and NOE assignments include $[^{13}C, ^{15}N]$-double labeled, partially deuterated (50%) $[^2H, ^{13}C, ^{15}N]$-triple labeled, and perdeuterated $[^2H, ^{13}C, ^{15}N]$-triple labeled protein samples. The labeled C-terminal domain of Stt3p samples were prepared as described previously (22,23).

NMR experiments and resonances Assignments—All spectra were acquired at 55 °C using either a Bruker 600 MHz Avance spectrometer equipped with a triple resonance cryoprobe in the Department of Chemistry and Biochemistry, Auburn University or a Varian Inova 900 MHz NMR spectrometer equipped with a cold probe in the Complex Carbohydrate Research Center (CCRC), University of Georgia. NMR data collected were subsequently processed by NMRPipe program and analyzed by NMRView software (42).

NMR data collected for protein side-chain assignments are: HBHA(CO)NH, HCH-TOSY, HNHA, HCC(CO)NH, (H)CC(CO)NH, and TOCSY-HSQC, using either $[^{13}C, ^{15}N]$-double labeled or highly deuterated $[^2H, ^{13}C, ^{15}N]$-triple labeled protein as NMR samples. For NOE experiments, $^{15}N$-edited 3D NOESY-HSQC, $^{13}C$-edited 3D NOESY-HSQC, and $[^{13}C, ^{15}N]$- edited 4D HSQC-NOESY-HSQC data were recorded. The mixing times for $^{15}N$- edited 3D NOESY-HSQC, $^{12}C$-edited 3D NOESY-HSQC and $[^{15}N, ^{13}C]$-edited 4D HSQC-NOESY-HSQC were set as 150 ms, 110 ms and 150 ms, respectively.

ILV methyl protonated sample preparation and assignment—The protocol of expression of methyl protonated [I(δ1 only), L($^{13}C,H_{12},CD_{3}$),V($^{13}C,H_{12},CD_{3}$)] U-[$^{15}N, ^{13}C, ^2H]$ sample of C-terminal domain of Stt3p was same as the expression of U-[$^{15}N, ^{13}C, ^2H]$ sample except for one hour prior to induction, 2-keto-3,3-d$_2$-1,2,3,4-$^{13}C$-butyrate sodium salt (Sigma Aldrich) and 2-keto-3-methyl-d$_3$-3-d$_1$-1,2,3,4-$^{13}C$-butyrate sodium salt (Sigma Aldrich) were added to the media. Pulse programs for NMR experiments for site-specific methyl assignment were written in our lab, mainly based on the pulse schemes reported by Kay’s group (43). Both 3D $^{13}C$-edited NOE-SY-HSQC and 4D $[^{13}C, ^{13}C]$-edited HSQC-NOESY-HSQC spectra were recorded for the ILV methyl protonated sample.

Topology Determination—16-Doxyl-stearic acid (16-DSA) was used as the hydrophobic paramagnetic spin probe to determine the trans-membrane domain of the C-terminal domain of Stt3p in SDS micelles. Titrations were performed by stepwise addition of 16-DSA over a concentration range of 0-2 mM to a constant amount of U-$^{15}N$ protein sample (0.1 mM). $[^1H, ^{15}N]$-HSQC experiments were carried out using the same parameters except P$_1$ (the 90 degree hard pulse) and shimming values. The peak intensities were measured at each titration point to assess the amount of paramagnetic induced line broadening.

HSQC titrations—HSQC titrations were carried out as previously described (22).

Structure Calculation and Refinement—Structure calculation was done with CYANA 3.0. NOESY spectra were peak picked and integrated interactively. Dihedral angle constraints were determined by TALOS (+) (31). Structure calculation of the C-terminal domain of Stt3p was done in two steps. In the first step, the structure of each secondary
structural element was determined to high resolution. The structures of these elements were then fixed by medium-range NOE restraints and relatively tight backbone dihedral angle restraints. In a second step, these segments were folded together with long-range NOE restraints. 100 structures were calculated and 10 structures of lowest total energy were used to represent their ensemble conformation. In the final calculated structures, the percentages of residues that reside in the most favored, additionally allowed and generously allowed regions of the Ramachandran diagram are 60.5%, 28.6% and 8.1%, respectively.

RESULTS AND DISCUSSION

Structure determination of the C-terminal domain of Stt3p- De novo structure determination of membrane proteins remains a challenge. For X-ray crystallography, the presence of detergents often hinders sample crystallization (24); whereas in NMR spectroscopy, the slow tumbling of membrane proteins embedded in detergent micelles often dramatically broadens the resonance line widths due to their rapid transverse relaxation rates. The difficulty of structure determination for α-helical membrane proteins by NMR is even more striking in comparison to their β-barrel counterparts. This is mainly because unlike the β-barrel membrane proteins, where the long-range interstrand NOEs are abundant and readily provide adequate restraints to define the global folds, there are typically very few long-range NH-NH NOEs for α-helical membrane proteins (25). Additionally, the NMR spectral dispersion for the amide region is very narrow for α-helical membrane proteins which results in severe overlapping of resonances, making unambiguous assignments much harder. Consequently, to date, successful cases of α-helical membrane protein structure determination using NMR methods have been largely limited to proteins of low molecular weights (25-30).

To achieve unambiguous assignments, we have used a combination of NOEs from [13C, 15N]-double-labeled, partially deuterated (50%) triple-labeled, and uniformly [2H, 13C, 15N]-triple-labeled samples, as well as an ILV methyl protonated otherwise uniformly [2H, 13C, 15N]-triple-labeled sample, together with backbone dihedral angles from chemical shift analysis (TALOS+) (31). It is noteworthy that 4D NOESY data proved to be very helpful in unambiguous identification of many medium-range NOEs, a characteristic for α-helices, together with many long-range NOEs. The detailed statistics of restraints used for structure determination are provided in Table-1 and in methods and materials section.

Overall Structure and Topology of the C-terminal Domain of Stt3p- As presented in Figure 1A, the C-terminal domain of Stt3p reveals a globally compact oblate spheroid shape structure, with a major axis of ~680 Å and minor axis of ~400 Å. This domain of Stt3p is primarily helical, containing eleven helices (α1-α11) and a disordered C-terminus (Figure 1B). These eleven helices are connected by short or medium-length loops. This high helicity is consistent with our previous experimental data from far-UV CD and CSI (Chemical Shift Index) analysis (22,23). Although both TALOS+ program and CSI predict the formation of a β-strand for residues R592-W598, we failed to unambiguously assign supportive NOEs to confirm this prediction, presumably due to peak overlapping. In fact, the absence of β-strand in the C-terminal domain of eukaryotic Stt3p has been predicted by some structure predictions programs (20).

The topology of the full length yeast Stt3p has been proposed to be N cyt-C-lum with 11 transmembrane helices and a soluble C-terminal domain (32). However, some prediction programs predict that there is a transmembrane domain near the middle of the C-terminal domain (22,33), while some other programs predict it only contains a
hydrophobic region at this position (32). Therefore, the detailed topology of the C-terminal domain of Stt3p remains controversial. In this study, we used a nonpolar paramagnetic probe, 16-doxylstearate (16-DSA), to investigate the topology of the C-terminal domain of Stt3p with respect to the membrane. This nonpolar probe accelerates the relaxation of neighboring nuclear spins via dipole-dipole interactions of the residues located in the transmembrane (TM) domain, thus reducing the intensities of HSQC peak (34). The HSQC spectra were acquired both in the absence and presence of 2 mM 16-DSA (Fig. S-2). Further analysis showed that the addition of 16-DSA results in significant reduction in peak intensities for the following segments: residues 488-504, 511-526, 539-551, and 566-582 (Fig. 2). From the structure calculated in this study, we observed that except segment 566-582 which is partially protruding out, all other segments could form a hydrophobic pocket. This explained the reason for the resonance intensity reduction in those segments, since the hydrophobic probe 16-DSA can fit into a hydrophobic pocket. Taken together with the results obtained from transmembrane (TM) prediction programs (22), it is likely that the C-terminal segment encompassing residues 566-582 (α5) along with the adjoining loop residues are membrane embedded. Because it has been established that the C-terminal domain of Stt3p is located in the luminal side of ER (32), we concluded that this membrane embedded segment only penetrates the lipid bilayer but does not span it completely, and therefore is not a transmembrane domain. Taken together, we propose that the protein domain under study is monotopic. This is similar to the case of bacterial peptidoglycan glycosyltransferases, which was shown to be a monotopic membrane protein (35). As shown in Fig. 3, this topology model is consistent with the structure presented in this study, in which the highly hydrophobic C-terminus of α5 and the adjoining loop protrudes into the lipid bilayer. Accordingly, combining with the existing result, we propose that the full length Stt3p has an Ncyt-C-lum topology with 11 transmembrane domains in its N-terminal domain and a luminal monotopic C-terminal domain (Fig. 3B).

We have previously reported that different concentrations of detergent led to changes in chemical shift positions of some residues in the [1H, 15N]-HSQC spectra of the C-terminal domain of Stt3p (22). To confirm the topology of the C-terminal domain of Stt3p, we further analyzed the above data. Resonances in the HSQC spectra collected at different detergent concentrations were assigned (23). The effect of increasing detergent concentration (50 mM to 200 mM) on chemical shift positions of each residue in the protein is shown in Fig. S-3A. Residues that underwent significant chemical shift perturbations (≥ 0.05 ppm) when mapped into the protein structure (Fig. S-3B) revealed that the most of the perturbed residues are located at the proposed interface region between the protein and the ER membrane (shown in blue color) except a few that are located in a peripheral helix (shown in red color). These results support our topology model of the C-terminal domain of Stt3p. Additionally, the last transmembrane segment of yeast Stt3p is predicted to encompass residues 442-464 (32), which suggests that the N-terminal residue of the protein under study (residue 466) should be located close to the ER membrane. Our model is in agreement with the above studies, showing close proximity between residue 466 and the lipid bilayer (Fig. 3A).

Comparison to the crystal structures of prokaryotic Stt3p homologs-Two crystal structures of the C-terminal soluble domain of prokaryotic Stt3p homologs have been previously reported, AglB protein of P. furiosus (12) and PglB protein of C. jejuni.
Recently, the crystal structure of full length PgIB of C. lari was also determined (21). The C-terminal domain of these structures comprises mainly an α-helical “central core” domain and some β-sheet-rich domains either encircling or inserting into the “central core” domain. The “central core” domain contains the well-conserved WWDYG motif and therefore was previously proposed as catalytic domains (12,20).

However, recent crystal structural studies on the full-length Stt3p of C. lari demonstrates that the WWDYG motif is indeed involved in binding of the signature sequon N-X-T/S (X ≠ P) on the acceptor peptide, but not directly in the catalysis process (21).

Comparison of the solution structure of the C-terminal domain of yeast Stt3p with the structures of its prokaryotic homologs reveals two major differences between them. First, the C-terminal domains of prokaryotic Stt3p homologs are water-soluble, lacking membrane-embedded segment; while our data suggest that the C-terminal domain of yeast Stt3p contains an ER membrane-embedded region, the C-terminus of α5 and residues located in the adjoining loop. We postulate that the anchoring of the C-terminal domain of Stt3p which contains the acceptor binding site to the ER membrane makes it closer to its donor substrate, dolichol-linked oligosaccharide, which is also embedded in ER membrane. This might potentially increase the effective local concentration of the donor substrate, and hence facilitates the N-glycosylation process. Another striking difference is that the counterparts of the β-sheet-rich domains in AgIB and PgIB proteins are, however, missing in the C-terminal domain of yeast Stt3p. It is thus reasonable to postulate that the C-terminal domain of Stt3p as a whole is corresponding to the “central core domain” in the prokaryotic homologs, although it has a larger size and contains more helix elements. The function of those β-sheet-rich domains might be fulfilled by the other subunit(s) in the case of yeast OT.

**Acceptor substrate binding studies of the C-terminal domain of Stt3p:** As mentioned above, the WWDYG motif, which is conserved across the three domains of life, has been shown to be the acceptor substrate binding site of OT (21). To investigate the interaction between the C-terminal domain of Stt3p and the acceptor substrate, NMR titration studies were carried out using Asn-Asp-Thr-NH2 acceptor peptide, which contains the consensus N-linked glycosylation sequon. Binding results were evaluated by monitoring the changes in chemical shift positions in the [1H, 15N]-HSQC spectra of the C-terminal domain of Stt3p, as described earlier (22). As shown in Fig. 4A and supplemental Fig. S-4, the resonance positions of some residues were perturbed upon addition of the acceptor substrate peptide, confirming the interaction between the protein and peptide. Mapping of these residues within the structure of C-terminal domain of Stt3p reveals they can be divided into two groups: either located in the peripheral loop regions (shown as blue balls) or clustered around the WWDYG motif area (WWDYG motif and two previous Ala residues are shown as red balls while the adjacent residues are shown as yellow balls) as shown in Fig. 4B. Except the residues in the loop regions, other perturbed residues are found to be part of a pocket that is formed around the WWDYG motif. Based on the NMR mapping studies of the C-terminal domain of Stt3p with the acceptor peptide, we propose this pocket as the acceptor substrate binding site located in the luminal C-terminal domain of Stt3p. As reported earlier, the apparent Kₐ₀.₅ of the interaction between the substrate peptide and the C-terminal domain of Stt3p is about 10 mM (22). The weak interactions could possibly due to the fact that we are not working with the whole OT complex. Moreover, the in vitro experimental
conditions can never be same as within a cell. However, if the in vivo interaction is actually weak, it may provide the enzyme with a fast association and disassociation with the substrate, thus facilitating an efficient screening and glycosylation process of a nascent polypeptide chain.

Comparison of the acceptor binding site of Stt3p with its prokaryotic homologs-

Limited sequence similarity between the eukaryotic Stt3p and its prokaryotic homologs does not allow a meaningful sequence alignment. Since the architecture surrounding the WWDYG motif is highly relevant to understand the mechanisms of N-glycosylation, structural comparison of these regions of Stt3p and its prokaryotic homologs should provide information with respect to how well the mechanism of N-glycosylation is conserved throughout the three domains of life.

As shown in Fig. 5A-D, despite very low overall amino acid sequence similarities, the structural architecture of the region surrounding WWDYG motif of Stt3p is strikingly similar to that of its prokaryotic homologs, AglB protein from the archaeon P. furiosus, and PglB proteins from the bacterium C. jejuni, and C. lari. This observation suggests that OT shares a common mechanism of acceptor substrate binding in all three domains of life. In the binding site of both eukaryotic and prokaryotic OT (Fig. 5), the local helical structural feature of the WWDYG motif is strictly conserved, although the N-terminal part of the WWDYG motif in AglB adopts a rare left-handed helical conformation while a typical right-handed helix in the other three structures. Structural studies of PglB from the bacterium C. lari in complex with an acceptor peptide has revealed that the side chains of the two tryptophan in this motif directly interact with the β-hydroxyl group of the +2 position Thr of the acceptor sequon via hydrogen bonds (21). Additionally, an Ile residue, which is located adjacent to the WWDYG motif, was proposed to stabilize the interaction between the enzyme and the glycosylable acceptor sequon through van der Waals contact (21). In our model, we also observe an Ile residue, Ile572 in close proximity to the WWD residues (Fig. 5A). This observation appears to be reminiscent of PglB of C. lari, suggesting that residue Ile572 might play a similar role in binding of acceptor substrate. In addition, sequence alignment shows the residue Ile 572 is absolutely conserved from yeast to human (Fig. S-1).

Close comparison of the acceptor binding sites of these different models reveals that the aspartate residue side chain orientation within the WWDYG motif is different. In both yeast Stt3p and the archaeon P. furiosus AglB, the side chain of this aspartate residue orients away from the two Trp residues (Fig. 5A-B). In contrast, in the PglB of the bacterium C. jejuni, and C. lari, the side chains of the aspartate residue in WWDYG motif are pointing towards the two Trp residues (Fig. 5C-D). This observation implies that this Asp518 of Stt3p and AglB might play a different role than that of its bacterial homolog PglB, in which Asp518 together with the Trp516 and Trp517 directly interact with the acceptor sequon via hydrogen bonds (21). Indeed, if Asp518 in yeast OT also forms a hydrogen bond with the acceptor substrate, the OT activity would be expected to be largely unaffected upon its mutation to glutamate residue, because both aspartate and glutamate have comparable ability to form hydrogen bonds with the acceptor substrate due to their similar side chains. However, it has been shown that D518E mutation in yeast Stt3p results in a lethal phenotype rendering the enzyme completely nonfunctional (6). We have proposed a structural role for Asp518 in our earlier studies (22). Further structural and functional studies are needed to clarify the
exact role of Asp518 in the glycosylation process of higher organisms.

Based on phylogenetic tree analysis and structural studies of AglB protein from the archaeon \textit{P. furiosus}, it was postulated that the acceptor substrate binding site (which was tentatively proposed as the catalytic site of OT in this study) of eukaryotic OT was formed by the WWDYG motif and the so-called DK motif (20). Our model reveals the DK motif of Stt3p (residues D583-K586) is indeed located in close proximity to the WWDYG motif (Fig. 5A), in agreement with the above proposal. However, the spatial orientation of the side chains of the residues in DK motif can not be precisely positioned in our model due to the missing or weak NMR resonances for these residues. For the same reason, we are uncertain about the local secondary structure of the DK motif. However, CSI analysis and TALOS+ prediction based on the neighboring residues shows that it is very likely that residues D582 to E587 form a small helix.

Structural insights into the interaction between the C-terminal domain of Stt3p and other OT subunits—Arguably, one of the most striking features of eukaryotic OT probably lies in its structural complexity. Yeast OT has two isoforms. Each isoform is composed of eight membrane protein subunits containing either Ost3p or Ost6p (36-38). In the absence of high-resolution structure of the whole OT complex, any information on the subunit-subunit interactions might help to understand the functional role of each subunit as well as the mechanism of enzyme catalysis. The EM structure of the whole OT complex has been previously reported (19). Comparison of the NMR structure of the C-terminal domain of Stt3p to the corresponding region in the EM structure demonstrates a very good spatial fit despite the fact that protein expression and reconstitution methods are different (Fig. 6A-B). Because the OT samples used for EM studies are enzymatically active, we conclude that the solution structure presented here represents the native fold of the C-terminal domain of Stt3p.

Besides the catalytic Stt3p subunit, the EM model also demonstrates the spatial orientations of Ost1p, the proposed acceptor substrate-recognizing subunit (39) and Wbp1p, the proposed donor substrate-recognizing subunit (40). Based on the above information, our structure reveals that \(\alpha_4\), which is aligned approximately parallel to the interface region of Ost1p, is responsible for the primary contact between the C-terminal domain of Stt3p and the luminal domain of Ost1p (Fig. 6C). More specifically, residues D534, N536, T537, N540, T541, A544 and K548 that are oriented towards Ost1p are very likely responsible for interactions between these two subunits. Sequence alignment of the C-terminal domain of Stt3p from yeast to human shows that these residues are all highly conserved (Fig. S-1), suggesting their biological importance. With the exception of residue A544, other polar residues can have electrostatic interactions or hydrogen bonding with the residues of Ost1p at the interface, thus providing hydrophilic environments to the groove formed between Stt3p and Ost1p. Further investigation will be needed to confirm the roles of those amino acid residues, and to address the physiological significance for the hydrophilic properties of this proposed protein-conducting groove. Similarly, structural fitting of the C-terminal domain of Stt3p to the EM model allows us to identify the C-terminus of \(\alpha_2\) and the loop between \(\alpha_2-\alpha_3\) (L23) as the putative site for interactions with Wbp1p. Sequence alignment demonstrates that the C-terminus of \(\alpha_2\) helix is highly conserved while L23 is only partially conserved (Fig. S-1). The presence of a disordered loop in the interface region between Stt3p and Wbp1p is in agreement with the dynamic feature of Wbp1p (19). This might provide a structural prerequisite for efficient N-glycosylation since it is shown...
that the substrate of eukaryotic N-glycosylation reaction is in a flexible form (41).

In summary, we present here the high-resolution structure of the C-terminal domain of Stt3p, the proposed acceptor binding domain of OT. Furthermore, considering the high sequence homology of eukaryotic Stt3p, determination of the first high-resolution structure of the acceptor binding domain of yeast OT is an important step towards understanding the mechanisms of the eukaryotic N-glycosylation process.

**Acknowledgements** This research was financially supported by USDA PECASE award 2003-35302-12930, NSF IBN-0628064, NIH DK082397, and USDA 2011-65503-20030 grants to S. M. The 900 MHz NMR data were collected at the NMR facility of the University of Georgia, funded by NIH grant GM66340. We thank Dr. Uma V. Katre for critical reading of the manuscript.

**Author Contributions** S.M. conceived and designed the research plan; C. H. prepared samples, performed all NMR experiments, processed data, and completed resonance assignments; C.H. R.B. and S.M. analyzed the data; C.H. and S.M. wrote the paper.

**Accession codes**: The atomic coordinates (2LGZ) have been deposited in the Protein Data Bank.

**References**

Figure Legends

**Figure 1.** Solution structure of the C-terminal domain of Stt3p: A. superposition of 10 conformers representing the final NMR structure. B. Ribbon structure of the lowest energy conformer.

**Figure 2.** Effects of the spin-labeled detergent 16-DSA on $[^{1}H,^{15}N]$-HSQC peak intensities for the U-$^{15}$N-labeled C-terminal domain of Stt3p the in SDS micelles at pH 6.5 and 55 °C. Resonance intensity change of each residue upon addition of 2mM 16-DSA to the C-terminal domain of Stt3p in SDS micelles was plotted against the amino acid sequence. The chemical shift changes was calculated as combined $^1$H- and $^{15}$N-chemical shift ($\Delta\delta(HN) = (\Delta\delta(^1H))^2 + (0.2\cdot\Delta\delta(^{15}N))^2)^{0.5}$).

**Figure 3.** Proposed topology of the C-terminal domain of Stt3p. A. The C-terminus of $\alpha_5$ and its adjoining loop (depicted as mesh) is proposed as the ER membrane embedded region. The residue 466, which is known to be adjacent to the ER membrane, is depicted as blue sticks. B. Topology of the yeast Stt3p. The last transmembrane domain has been shown to be from residue 442-464 (32), suggesting the residue 466 is adjacent to the ER membrane.

**Figure 4.** Interactions of the C-terminal domain of Stt3p with the acceptor peptide, Asn-Asp-Thr-NH$_2$. A. Chemical shift plot of combined $^1$H- and $^{15}$N-chemical shift changes ($\Delta\delta(HN) = (\Delta\delta(^1H))^2 + (0.2\cdot\Delta\delta(^{15}N))^2)^{0.5}$) of the amide moieties of the C-terminal domain of Stt3p in SDS micelles induced by the addition of 50-fold molar excess of ligand peptide. B. Upon addition of peptide, the residues undergoing significant NMR resonance chemical shift perturbation are shown as balls. The unspecific interactions in the peripheral regions are shown as blue balls. The proposed acceptor substrate binding pocket is formed by the WWDYG motif (red balls) and its adjacent residues (yellow balls).

**Figure 5.** Comparison of the acceptor substrate binding sites of the C-terminal domain of yeast Stt3p (A) with its prokaryotic homologs: AglB protein from the archaeon *P. furiosus* (B), PglB proteins from the bacterium *C. jejuni* (C) and *C. lari* (D). Figures B, C and D were prepared using the pdb files of 2ZAI, 3AAG and 3RCE, respectively.

**Figure 6.** Comparison of the NMR structure of the C-terminal domain of Stt3p to the corresponding region in the EM model of OT complex (19). A. NMR model of the C-terminal domain of Stt3p. B. EM model of the OT complex. The corresponding region of the C-terminal domain of Stt3p is depicted in yellow color in the EM model. Both the NMR and EM structures of the corresponding region present a very good spatial fit to each other. C. Proposed interactions between the C-terminal domain of Stt3p and the other OT subunits. To clearly show the interacting surface, the protein orientation in C is different from that shown in A. The proposed surfaces that interact with Ost1p and Wbp1p are colored in blue and yellow, respectively, while the proposed active site is circled in dashed red line.
Table 1. NMR and refinement statistics for the NMR ensemble of the C-terminal domain of Stt3p.

<table>
<thead>
<tr>
<th>NMR distance and dihedral constraints</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance constraints</td>
<td></td>
</tr>
<tr>
<td>Total NOE</td>
<td>1445</td>
</tr>
<tr>
<td>Intra-residue</td>
<td>0</td>
</tr>
<tr>
<td>Inter-residue</td>
<td>1445</td>
</tr>
<tr>
<td>Sequential ((</td>
<td>i - j</td>
</tr>
<tr>
<td>Medium-range ((</td>
<td>i - j</td>
</tr>
<tr>
<td>Long-range ((</td>
<td>i - j</td>
</tr>
<tr>
<td>Total dihedral angle restraints</td>
<td>272</td>
</tr>
<tr>
<td>(\phi)</td>
<td>136</td>
</tr>
<tr>
<td>(\psi)</td>
<td>136</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Structure statistics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Violations (mean and s.d.)</td>
<td></td>
</tr>
<tr>
<td>Distance constraints (Å)</td>
<td>0.27 +/- 0.15</td>
</tr>
<tr>
<td>Dihedral angle constraints (º)</td>
<td>4.24</td>
</tr>
<tr>
<td>Max. dihedral angle violation (º)</td>
<td>4.73</td>
</tr>
<tr>
<td>Max. distance constraint violation (Å)</td>
<td>0.52</td>
</tr>
<tr>
<td>Average pairwise r.m.s. deviation** (Å)</td>
<td></td>
</tr>
<tr>
<td>Heavy</td>
<td>3.07</td>
</tr>
<tr>
<td>Backbone</td>
<td>2.3</td>
</tr>
</tbody>
</table>

**[AU: Please indicate number of structures used in r.m.s. deviation calculations: “Pairwise r.m.s. deviation was calculated among 10 refined structures.”]
Figure 1.
Figure 2.
Figure 3.
Figure 4.
Figure 5.
Figure 6.
Eukaryotic N-glycosylation Occurs Via membrane-anchored C-terminal domain of Stt3p subunit of oligosaccharyl transferase
Chengdong Huang, Rajagopalan Bhaskaran and Smita Mohanty
J. Biol. Chem. published online August 3, 2012

Access the most updated version of this article at doi: 10.1074/jbc.M112.342253

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2012/08/03/M112.342253.DC1

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/early/2012/08/03/jbc.M112.342253.full.html#ref-list-1