Published in final edited form as:

Biopolymers. 2012 July ; 97(7): 499–507. doi:10.1002/bip.22028.

A Novel and Simple Method of Production and Biophysical Characterization of a Mini-Membrane protein, Ost4p: A Subunit of Yeast Oligosaccharyl Transferase

Amit Kumar, Pricilla Ward, Uma V. Katre, and Smita Mohanty*

Department of Chemistry and Biochemistry, Auburn University, Auburn, AL 36849, USA

Abstract

Asparagine-linked glycosylation is an essential and highly conserved protein modification reaction. In eukaryotes, oligosaccharyl transferase (OT), a multisubunit membrane-associated enzyme complex, catalyzes this reaction in newly synthesized proteins. In *Saccharomyces cerevisiae*, OT consists of nine non-identical membrane proteins. Ost4p, the smallest subunit, bridges the catalytic subunit Stt3p with Ost3p. Mutation of transmembrane residues 18–24 in Ost4p has negative effect on OT activity, disrupts the Stt3p-Ost4p-Ost3p complex, results in temperature-sensitive phenotype, and hypoglycosylation. Heterologous expression and purification of integral membrane proteins (IMPs) are the bottleneck in membrane protein research. We report the cloning, successful overexpression and purification of recombinant Ost4p with a novel but simple method producing milligram quantities of pure protein. GB1 protein was found to be the most suitable tag for the large scale production of Ost4p. The cleavage of Ost4p conveniently leaves GB1 protein in solution eliminating further purification. The precipitated pure Ost4p is reconstituted in appropriate membrane mimetic. The recombinant protein is highly helical as indicated by the far-UV CD spectrum. The well-dispersed HSQC spectrum indicates that this minimembrane protein is well folded. The successful production of pure recombinant Ost4p with a novel yet simple method may have important ramification for the production of other membrane proteins.

Keywords

membrane protein; oligosaccharyl transferase; Ost4p; glycosylation; protein expression; NMR

Introduction

Production and structural characterization of membrane proteins remains one of the greatest challenges in the era of structural genomics and proteomics. Although one third of the genome encodes for membrane proteins and over 60 % of drug targets belong to a class of membrane proteins called G protein-coupled receptors (GPCR), only 0.3 % of the PDB structures represent integral helical membrane proteins [1, 2]. This disparity is due to the difficulties associated with recombinant expression and purification of large quantities of integral membrane proteins that is required for structural and functional analysis. Much of

*Corresponding author. Department of Chemistry and Biochemistry, Auburn University, Auburn, AL 36849, USA. Tel: +1(334)826-7980 (S. Mohanty), bionmr61@yahoo.com.

Dedication:

The authors dedicate this article in the loving memory of Dr. Marie W. Wooten, Dean and Professor of the College of Sciences and Mathematics (COSAM) at Auburn University, who tragically passed away on November 5, 2010. Dean Wooten was our mentor, colleague, friend, and role model. She will be deeply and painfully missed, yet celebrated, honored, and never forgotten.

the efforts in the field have been devoted to the production of pure and homogeneous recombinant membrane protein and reconstitution in suitable membrane mimetic for biochemical and biophysical characterization.

Oligosaccharyl transferase (OT) is a multi-subunit enzyme complex that is located next to the ribosome in the rough endoplasmic reticulum (RER). This integral membrane protein complex catalyzes N-glycosylation of proteins in eukaryotes. N-Glycosylation of proteins is an essential, highly conserved protein modification reaction that occurs in all eukaryotic [3] and some prokaryotic organisms [4]. OT catalyzes the transfer of a preassembled high mannose oligosaccharide ($\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$) to the asparagine side chain of Asn-X-Ser/Thr acceptor sites on nascent polypeptides as they are translocated into the lumen of RER [5]. Genetic defects in human in the N-linked glycosylation pathway cause a class of disorders now known as congenital disorders of glycosylation (CDG) with severe clinical manifestations of mental retardation, developmental delay, hypoglycemia, liver dysfunction and gastrointestinal disorders [6–8]. Complete loss of N-glycosylation is lethal in all organisms. In yeast, *Saccharomyces cerevisiae*, OT consists of following nine nonidentical membrane protein subunits: Ost1p, Ost2p, Ost3p, Ost4p, Ost5p, Ost6p, Stt3p, Swp1p, Wbp1p [9]. Ost3p and Ost6p have 46% sequence homology and 21% sequence identity. All nine subunits have been cloned and sequenced and have been reported to exist in 3 subcomplexes through genetic and biochemical studies [10]. Ost4p has been proposed to bridge Ost3p and Stt3p in the Stt3p-Ost4p-Ost3p subcomplex [11]. This 3.9 kDa Ost4p protein is also reported to be required for the incorporation of either Ost3p or Ost6p resulting in two functionally distinct OT complexes [12]. It has been reported through genetic screens that both STT3 and OST4 are required for full OT activity *in vivo* [12]. OST4 encodes a 36 residues hydrophobic minimembrane protein and OST4-deleted cells show temperature-sensitive phenotype resulting in hypoglycosylation of both soluble and membrane N-glycoproteins [13]. There is a high sequence similarity between most of the subunits of OST identified in higher eukaryotes and those found in yeast. Ost4p homologs which are also minimembrane proteins have been found in *C. elegans*, *Drosophila*, *A. thaliana*, mouse, and human.

It has been shown through comprehensive mutagenesis studies that introduction of a single ionizable amino acid in positions 18 to 24 of Ost4p results in impaired cell growth and *in vitro* OT activity [11, 14]. These mutations were shown to cause the disruption of the interactions of Ost4p with Ost3p and Stt3p through coimmunoprecipitation approach [14]. Interactions of Ost4p with Stt3p were disrupted with Met18Lys, Leu21Lys or Ile24Lys mutation, while mutation of any residue from Met 18 through Ile 24 disrupted the Ost4p-Ost3p complex [14]. In contrast, no effect was observed when similar single mutations were introduced in residues 2–17 of Ost4p [14]. The high-resolution structure of Ost4p [15] in mixed polarity solvent (chloroform-methanol-water) that was determined in our laboratory did explain the results of the comprehensive mutagenesis studies on Ost4p [11, 14]. This minimembrane protein folded into a kinked helix in the mixed polarity solvent. The residues in positions 18–24 formed a well-defined i+4 ridges as seen in the 3-dimensional structure [15]. These residues form the central part of the α 2-helix (residues 15–28) for the formation of a functional complex and for proper OT activity. On the basis of the 3-dimensional structure and results of mutagenesis study, it was proposed that in the OT complex, Stt3p is packed against the α 2-helix of Ost4p using “ridges-into-grooves” interactions with Met18, Leu21, and Ile24 as the packing interface on one face, whereas Ost3p is involved in interactions with Met19, Thr20, Ile22, and Val23 on the other [14, 15]. Recently atomic resolution structure was determined for the human Ost4p [16] in the same mixed polarity solvent system. The human Ost4p also folds into a kinked helix in a similar manner as the yeast Ost4p.

Although the atomic resolution structures have been determined for two eukaryotic Ost4p using chemically synthesized proteins, many questions regarding impact of mutation on the 3-dimensional structure remain unanswered, such as what happens to the α 2-helix when there is a single point mutation of the residues 18–24 to an ionizable amino acid? Do the “ridges-into-grooves” interactions between the Ost4p α 2-helix and Stt3p get disrupted due to point mutation in the α 2-helix of Ost4p? What is the structural effect of a point mutation on the 3-dimensional structure of Ost4p? Does this point mutation on α 2-helix of Ost4p disrupt the kinked helix of Ost4p? Would this structure determined in mixed-organic solvent be same as in micelles or bicelles? In addition to these questions, it would be very interesting to study direct protein-protein interactions between Ost4p and the transmembrane domain of the interacting subunits, Stt3p and Ost3p.

The above high-resolution structures for both yeast and human Ost4p were determined with chemically synthesized Ost4p. Being very hydrophobic, chemical synthesis of these proteins are expensive and purification to homogeneity offers many challenges. Additionally, chemical synthesis of NMR sensitive isotope (^{15}N or ^{13}C) enriched form of Ost4p or other manipulation has become a practical problem and has limited the use of very sensitive NMR experiments for protein-protein interaction studies. This is illustrated by the fact that there has been no research on yeast Ost4p structure and function since the report of its structure in 2004 except the publication of human Ost4p structure in 2011. To date, there are only three structural reports at atomic resolution for eukaryotic OT subunits: NMR structures of the minimembrane protein Ost4p from yeast [15] and human [16], and the crystal structure of the N-terminal luminal domain of Ost3p/Ost6p [17]. Crystal structures of the C-terminal domain and that of the full-length prokaryotic homolog of yeast Stt3p have been solved recently (18–20). There is a report of a low resolution cryo-EM structure of yeast OT complex however it is difficult to know which one is Ost4p in this structure (21). The lack of high-resolution structures for different eukaryotic OT subunits clearly demonstrates that chemically synthesized protein is neither a practical option nor can provide the ease or the easy manipulation of a gene that is possible through genetic engineering. Thus, heterologous expression primarily in bacteria for the production of recombinant proteins will remain the workhorse for scientists in academics and in pharmaceutical companies for biochemical or biophysical or structure based drug design studies.

Despite the wealth of well established methods for the water soluble recombinant protein expression and purification, production of eukaryotic proteins, especially, integral membrane proteins (IMPs), remains a challenge. There is a bottleneck at every stage of the production of an IMP: starting from cloning to biophysical characterization. The biggest obstacle is in obtaining large amounts of a given IMP for atomic resolution structural studies which requires usually between 5 to 50 mg of pure protein for either high resolution NMR or X-ray crystallographic studies. This is the primary reason why IMPs represent only 0.3 % of the structures deposited in the Protein Data Bank (PDB) [22].

For past several years, we have unsuccessfully tried the production of recombinant Ost4p with various large tags such as GST. Here, we report the production of pure and homogeneous recombinant Ost4p using a novel but simple method of purification using a small protein GB1 (the beta 1 domain of *Streptococcus* protein G) as an N-terminal tag and with a C-terminal His-tag. Most importantly, efficient cleavage of the GB1 tag with thrombin yields pure Ost4p that precipitates out of the solution leaving GB1 protein behind requiring no further purification. This cleavage reaction is elegantly monitored through mass spectrometry. We have successfully optimized the cleavage reaction generating 30–40 mg of pure Ost4p per liter of bacterial culture.

Results

Cloning, Overexpression and Purification of GB1-Ost4p

After trying several different large and small protein tags, we were able to overexpress Ost4p with N-terminal GB1 tag and C-terminal His6 tag in the GEV2 vector. In order to clone the OST4 gene into the GEV2 vector, three different strategies were employed. Each PCR reaction was set up with the same forward primer but with a different reverse primer, where cleavage sites were engineered for cleavage at Tryptophan or Thrombin or Arginine site. The strategy that worked best for us was with Thrombin (Fig. 1). The PCR amplified and double digested gene product of OST4 and that of the vector GEV2 were ligated to create recombinant recombinant pGEV2/Ost4p plasmids. Ost4p was overexpressed in BL21(DE3)pLysS by inserting the recombinant pGEV2/Ost4p plasmids into the cells. Protein expression was optimized using different *E. coli* strains, temperatures, duration, and IPTG concentrations (see Materials and Methods for details). The target protein was expressed as soluble protein. The GB1-Ost4p was purified to homogeneity by taking advantage of the histidine tag at the C-terminus of the GB1-Ost4p. The crude lysate containing the GB1-Ost4p was loaded to with Ni-NTA column. The pure GB1-Ost4p was eluted with 500 mM imidazole (Fig. 2A).

As mentioned earlier, the recombinant GB1-Ost4p was expressed as soluble protein and remained in solution during Ni-NTA purification. However, cleavage of GB1 tag allowed C-terminal His-tagged Ost4p to precipitate out of the solution, thus eliminating the need for the removal of contaminating GB1 protein that remained in the solution. This simple method of cleavage of the GB1 tag in the GB1-Ost4p fusion protein with thrombin was optimized after many trials under various conditions: different buffer conditions, pH, temperature, denaturant, detergent and duration (Fig. 2B). Addition of 0.01% SDS was essential for the cleavage to occur in this case. Cleavage of the GB1 tag was highly successful after the optimization of the cleavage conditions as the GB1-Ost4 fusion protein was not detectable by SDS PAGE and mass spectrometry analysis.

Determination of Mass and Monitoring of Cleavage Reaction Simultaneously via Mass Spectrometry

The cleavage of GB1 tag in the GB1-Ost4p protein was very conveniently monitored by electrospray ionization-mass spectrometry (Figs. 3A, 3B, 3C & 3D). As the cleavage reaction proceeded, the intensity for M⁺ peak for GB1-Ost4p at 11,981.001 Da (Fig. 3A) started to reduce while the intensity for M⁺ peak for unlabeled GB1 around 6646.29 Da and M⁺ for unlabeled C-terminal his tagged-Ost4p around 5349.07 Da (Fig. 3B) appeared and increased in intensity. The cleavage of ¹⁵N-labeled GB1-Ost4p occurred in a similar manner as that of the unlabeled GB1-Ost4p. Briefly, the intensity of the M⁺ peak for ¹⁵N-GB1-Ost4p around 12,113.7 Da (Fig. 3C) started to reduce while the intensity for M⁺ peak for ¹⁵N-labeled GB1 around 6715.7 Da and M⁺ peak for ¹⁵N-labeled C-terminal his-tagged Ost4p around 5413.8 Da appeared and increased in intensity (Fig. 3D). The Ost4p precipitated out of the solution slowly with time. Thus, the solution was clear initially but white precipitate appeared and increased with time.

Characterization of the Ost4p by Far-UV CD Spectroscopy

Far-UV CD spectroscopy was employed to probe the secondary structure of the Ost4p reconstituted in 100 mM SDS detergent micelle system. The far-UV CD spectra (Fig. 4) in the SDS detergent micelle system had the characteristics of a typical α -helical protein with CD minima at 208 and 222 nm. This observation is consistent with the 3-dimensional structure that was determined in our laboratory for chemically synthesized Ost4p [14].

Characterization of Ost4p by NMR Spectroscopy

The 2D $\{^1\text{H}, ^{15}\text{N}\}$ HSQC spectrum provides both qualitative and quantitative information for the evaluation of whether a protein is well-folded and exists in a single conformation [23]. The quality and number of peaks present in a 2D $\{^1\text{H}, ^{15}\text{N}\}$ HSQC NMR spectrum reveal whether a protein is monomeric or if it exists in oligomeric forms [23–24]. 2D $\{^1\text{H}, ^{15}\text{N}\}$ HSQC NMR spectrum of a protein called the fingerprint region of the protein [25–29] since this region is very sensitive to change in protein structure or conformation. Any change in the protein structure/conformation due to change in pH or temperature also results in changes in the 2D $\{^1\text{H}, ^{15}\text{N}\}$ HSQC spectrum. Well dispersed 2D $\{^1\text{H}, ^{15}\text{N}\}$ HSQC spectra of both GB1-Ost4p and pure Ost4p (Figs. 5A & 5B) indicate that both proteins are folded properly.

Discussion

Expression and purification of a recombinant integral membrane protein (IMP) are arduous, time consuming and very challenging tasks despite the fact that protein chemistry for production and purification of globular proteins is well established. There is a bottleneck at every stage of the production of an IMP: starting from cloning to biophysical characterization. Production of pure recombinant of a 4 kDa integral membrane protein, Ost4p, has been very difficult and challenging. This small protein containing only 36 residues was previously cloned into pGEX3X vector with GST (glutathione transferase) as N-terminal tag; however, our efforts to produce pure Ost4p were unsuccessful. While searching for a suitable protein tag, GB1 protein with a mass of about 6 kDa was found to be very attractive alternative for the heterologous expression of Ost4p. Thus the OST4 gene was subcloned into GEV-2 vector with N-terminal GB1 tag and C-terminal His-tag with a thrombin cleavage site engineered between GB1 and Ost4p since neither GB1 nor Ost4p contain intrinsic thrombin cleavage site. After optimization of expression using several different cell lines, the GB1-Ost4p was successfully overexpressed in *E. coli* BL21 DE3pLysS cell line. Very interestingly, 100 % of the recombinant protein was expressed as soluble protein. The GB1-Ost4p was purified to homogeneity in a single step using Ni-NTA column. The molecular mass of the GB1-Ost4p was confirmed by mass spectrometry analysis. The NMR analysis of the pure protein suggested that the recombinant GB1-Ost4p fusion protein was folded into a stable structure. Two-dimensional $\{^1\text{H}-^{15}\text{N}\}$ - HSQC NMR is a very robust method to find out whether a protein is folded properly or not (23–24). In the 2D $\{^1\text{H}-^{15}\text{N}\}$ - HSQC experiment, the amide proton of each amino acid residue in the protein correlates to the attached nitrogen atom. Thus except for proline, the amide proton and nitrogen atom of each amino acid residue in a protein, would contribute a resonance in this region. Thus this region is called the fingerprint region of a protein. For a well folded protein, the resonances in the fingerprint region are well dispersed [23–27]. However, for a denatured or unfolded protein, the resonances collapse to a very narrow region [23, 27]. It is clear from the fingerprint region in the 2D HSQC spectrum that the GB1-Ost4p protein is well folded and has a stable tertiary structure. The $\{^1\text{H}-^{15}\text{N}\}$ - HSQC spectrum of GB1-Ost4p contains peaks that correspond to those observed in the $\{^1\text{H}-^{15}\text{N}\}$ - HSQC spectrum of pure GB1 protein reported previously [28–29]. These peaks have been labeled with one letter amino acid code and residue number in the GB1 protein primary sequence.

The cleavage of GB1 tag in the GB1-Ost4p was optimized under various conditions of buffer, detergents, denaturants, temperature and duration of the reaction. It was found that 0.01 % SDS was necessary for the cleavage to occur. The mass of pure GB1 and pure C-terminal his-tagged Ost4p after cleavage were confirmed by mass spectrometry. The $\{^1\text{H}-^{15}\text{N}\}$ - HSQC NMR spectrum of his-tagged Ost4p in mixed organic solvents was found to be well dispersed as well. The narrow amide proton chemical shift dispersion is typical for helical proteins and is consistent with the CD data. It has been shown previously that

His-tagged membrane protein can be reconstituted in detergent micelles without any problems [23]. The C-terminal His tagged Ost4p was successfully reconstituted in SDS micelles or mixed organic solvents.

Conclusion

We have reported a novel and very simple method for the overexpression of a minimembrane protein, the Ost4p, a subunit of *S. cerevisiae* OT. After many unsuccessful trials with large protein tags, Ost4p was successfully overexpressed as a completely soluble protein when fused to the 6 kDa GB1 protein to its N-terminus. The recombinant protein was purified in a single step with Ni-NTA column. Most importantly Ost4p was efficiently cleaved with thrombin from the N-terminal GB1 tag leaving GB1 in the solution, thus eliminating any additional step for purification. The precipitated Ost4p was reconstituted either in mixed organic solvent or detergent micelles. Far-UV CD data show Ost4p is highly helical consistent with our previous structural work. The peak dispersion in the HSQC data demonstrates that the protein is well-folded in the mixed organic solvent system. It is likely that this novel method of expression and production of this minimembrane protein may have important ramification for the production of other smaller integral membrane proteins.

Materials and Methods

Subcloning

The OST4 gene was cloned into GEV 2 vector using the following primers: forward, 5' - GGCCGCGGATCCATGATCTCTGATGAACAGCTGAACTC - 3' and reverse, 5' - TCGCGG CTCTGAGAGAACCACGAGGCACCAGGTTCTTAGGAGACATGGTGGAGTC-3'. Yeast plasmid pHP84 containing OST4 gene was used as the template. The PCR amplified product were purified, digested with *Bam*HI and *Xho*I, and cloned into the *Bam*HI and *Xho*I restriction sites of GEV 2 vector with a N-terminal GB1 protein and a C-terminal His tag. A thrombin cleavage site was introduced between the C-terminus of GB1 protein and the N-terminus of Ost4p for the cleavage of GB1 tag. The orientation and sequence of the Ost4p cDNA in the resulting pGEV2-Ost4p vector were confirmed by restriction mapping analysis and DNA sequencing.

Overexpression of GB1-Ost4p

Unlabeled and uniformly ^{15}N labeled GB1-Ost4p were produced in *E. coli* BL21(DE3) pLysS cells (Stratagene). Expression of the N-terminal GB1 and C-terminal His-tagged Ost4p in the GEV2 vector were under the control of an IPTG (isopropyl- β -thiogalactopyranoside) inducible promoter. The expression was carried out as described here. The overnight starter culture was diluted to an OD₆₀₀ of 0.08–0.1 in a fresh LB medium containing 100 $\mu\text{g}/\text{mL}$ of ampicillin. The culture was incubated at 37 °C in a shaker until the OD₆₀₀ reached to 0.4–0.5, after which it was induced with 1 mM IPTG. The culture was further incubated at 37 °C for another 4 h and then cooled on ice for 30 min. The cells were harvested by centrifugation at 8000 rpm for 20 min at 4 °C and frozen at –80 °C until needed. For NMR studies, uniformly ^{15}N labeled GB1-Ost4p was expressed in M9 minimal medium with [^{15}N] ammonium chloride (Cambridge Isotope Laboratories, MA). After induction, the culture was further incubated at 30 °C for 15 h and the cells were harvested. Tris-Tricine gel electrophoresis was carried out to verify the protein expression.

Purification of GB1-OST4p

The cell pellet was freeze-thawed 5–6 times, by freezing in liquid nitrogen for 5 min followed by thawing on ice for 1.5 h. After freeze-thawing, the cells were suspended in the

lysis buffer (50 mM sodium phosphate buffer, pH 6.5, 200 mM NaCl, 0.01 % NaN₃) and lysed by sonication. DNase or lysozyme was not used for the bacterial cell lysis. The insoluble fraction was removed from the cell lysate by centrifugation at 12,000 rpm for 20 min at 4 °C.

Unlabeled as well as ¹⁵N labeled GB1-Ost4p proteins were purified using affinity chromatography on a Ni-NTA column. To start with, 3 mL of Ni-NTA resin (Qiagen) was packed in a column and washed extensively with deionized water. The column was then equilibrated with the binding buffer (20 mM sodium phosphate buffer, pH 7.4, 500 mM NaCl, and 10 mM imidazole). For equilibration, 3–4 ml buffer was added to the column, incubated at room temperature with continuous shaking and drained. This step was repeated for three times. To 3 mL of the cell lysate, NaCl, and imidazole were added to get final concentrations of 350 mM and 10 mM respectively. This cell lysate was added to the column and incubated at room temperature for ~16 h on a shaker, to allow the binding of the protein to Ni-NTA resin.

The unbound proteins were drained from the column and the impurities were removed using a washing buffer in 20 mM sodium phosphate buffer, pH 7.4 containing 500 mM NaCl, 25 mM imidazole several times with shaking. The bound GB1-Ost4p was eluted using the elution buffer (20 mM sodium phosphate buffer, pH 6.5, 500 mM NaCl, 500 mM imidazole) with shaking. Each time, 500 µL –1 ml of elution buffer was added to the column, shaken for 1 h at room temperature and the elution was collected. This step was repeated 6–7 times.

Production of Pure Ost4p in a Single Step with the Cleavage of GB1 tag

The GB1 tag was cleaved from Ost4p using thrombin. The pure protein was first concentrated and exchanged into the cleavage buffer (20 mM Tris-HCl pH 8.4, 150 mM NaCl, 2.5 mM CaCl₂). Cleavage of the GB1 tag in the GB1-Ost4p fusion protein with thrombin (VWR) was optimized after many trials under various conditions, such as: buffer, pH, temperature, denaturant, detergent and duration. The GB1 tag was cleaved very successfully at room temperature in 20 mM Tris-HCl, pH 8.4, 150 mM NaCl, 2.5 mM CaCl₂, and 0.01 % SDS allowing the precipitation of pure Ost4p from cleavage solution requiring no further purification to remove GB1 tag. The concentration of SDS used in the cleavage reaction was very low (< 5 mM) to keep the cleaved Ost4p in solution. The reaction was conveniently monitored by mass spectrometry. The precipitated C-terminal his-tagged Ost4p was separated by centrifugation at 14,000 rpm for 15 minutes at 25 °C leaving GB1 protein in solution. The pure Ost4p solid was air dried and reconstituted in the appropriate membrane-mimic environment (mixed organic solvent or in detergent micelles) through sonication for ~ 30 minutes in a water bath at room temperature.

Mass Spectrometry

Mass spectrometry was carried out at the Department of Chemistry and Biochemistry, Auburn University Mass Spectrometry Facility. The cleaved GB1-Ost4p was analyzed by electrospray ionization-mass spectrometry on a Q-ToF Premier™ (Waters) mass spectrometer.

Circular Dichroism (CD) Spectropolarimetry

All CD experiments were performed on a JASCO J-810 automatic recording spectropolarimeter using a 0.05 cm path length quartz cell at room temperature. The wild-type Ost4p was assessed in 20 mM phosphate buffer, pH 6.5 containing 100 mM SDS. The protein concentration was 300 µM for far-UV CD measurement. Data were averaged over 100 scans with a response time of 1 s and a scan speed of 100 nm/min. CD data were converted to mean residual ellipticity (θ) by standard procedures.

NMR Measurements

All NMR data were collected at 25 °C on a Bruker Avance 600-MHz spectrometer equipped with a triple resonance $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ TCI cryoprobe. NMR samples consisted of 300 μM of uniformly ^{15}N -labeled GB1-Ost4p or Os4p in 50 mM phosphate buffer containing 5 % D_2O , pH 6.5, 1 mM EDTA and 0.01 % sodium azide (for GB1-Ost4p) or 4:4:1 of $\text{CDCl}_3:\text{CD}_3\text{OD}:\text{H}_2\text{O}$ (4:4:1) for Ost4p in Shigemi tubes. Two-dimensional $\{^1\text{H}, ^{15}\text{N}\}$ heteronuclear single quantum coherence (HSQC) spectra were collected using suitable Bruker pulse program. The raw NMR data were processed offline on a SGI (Silicon Graphics Inc.) computer using NMRPipe (30) and visualized using NMRView (31) software.

Acknowledgments

This research was financially supported by USDA Presidential Early Career Award for Scientists and Engineers (PECASE) award 2003-35302-12930, USDA grant 2011-65503-20030, NSF grant IBN-0628064, and NIH grant DK082397 to Smita Mohanty. We thank Mr. Suman Mazumder and Mr. Mohiudeen Ovee for their help.

References

1. Vitkup D, Melamud E, Moulton J, Sander C. *Nat Struct Biol.* 2001; 8:559–566. [PubMed: 11373627]
2. Wallin E, von Heijne G. *Protein Sci.* 1998; 7:1029–1038. [PubMed: 9568909]
3. Kornfeld R, Kornfeld S. *Annu Rev Biochem.* 1985; 54:631–664. [PubMed: 3896128]
4. Wacker M, Linton D, Hitchen PG, Nita-Lazar M, Haslam SM, North SJ, Panico M, Morris HR, Dell A, Wren BW. *Science.* 2002; 298:1790–1793. [PubMed: 12459590]
5. Marshall RD. *Annu Rev Biochem.* 1972; 41:673–702. [PubMed: 4563441]
6. Westphal V, Xiao M, Kwok PY, Freeze HH. *Hum Mutat.* 2003; 22:420–421. [PubMed: 14517965]
7. Martin PT, Freeze HH. *Glycobiology.* 2003; 13:67R–75R.
8. Freeze HH. *J Pediatr.* 1998; 133:593–600. [PubMed: 9821413]
9. Knauer R, Lehle L. *Biochim Biophys Acta.* 1999; 1426:256–273.
10. Karaoglu D, Kelleher DJ, Gilmore R. *J Biol Chem.* 1997; 272:32513–32520. [PubMed: 9405463]
11. Kim H, Yan G, Heijne G, Caputo GA, Lennarz WJ. *Proc Natl Acad Sci USA.* 2003; 100:7460–7464. [PubMed: 12810948]
12. Zufferey R, Knauer R, Burda P, Stagljar I, te Heesen S, Lehle L, Aebersold M. *EMBO J.* 1995; 14:4949–4960. [PubMed: 7588624]
13. Chi JH, Roos J, Dean N. *J Biol Chem.* 1996; 271:3132–3140. [PubMed: 8621712]
14. Kim H, Park H, Montalvo L, Lennarz WJ. *Proc Natl Acad Sci USA.* 2000; 97:1516–1520. [PubMed: 10677492]
15. Zubkov S, Lennarz WJ, Mohanty S. *Proc Natl Acad Sci USA.* 2004; 101:3821–3826. [PubMed: 15001703]
16. Gayen S, Kang CB. *Biochem Biophys Res Commun.* 2011; 409:572–576. [PubMed: 21609714]
17. Schulz BL, Stirnimann CU, Grimshaw JP, Brozzo MS, Fritsch F, Mohorko E, Capitani G, Glockshuber R, Grütter MG, Aebersold M. *Proc Natl Acad Sci USA.* 2009; 106:11061–11066. [PubMed: 19549845]
18. Maita N, Nyirenda J, Igura M, Kamishikiryo J, Kohda D. *J Biol Chem.* 2010; 285:4941–4950. [PubMed: 20007322]
19. Igura M, Maita N, Kamishikiryo J, Yamada M, Obita T, Maenaka K, Kohda D. *EMBO J.* 2008; 27:234–243. [PubMed: 18046457]
20. Lizak C, Gerber S, Numao S, Aebersold M, Locher KP. *Nature.* 2011; 474:350–355. [PubMed: 21677752]
21. Li H, Chavan M, Schindelin H, Lennarz WJ, Li H. *Structure.* 2008; 16:432–440. [PubMed: 18334218]

22. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. J Mol Biol. 2001; 305:567–580. [PubMed: 11152613]
23. Huang C, Mohanty S, Banerjee M. Biochemistry. 2010; 49:1115–1126. [PubMed: 20047336]
24. Gronenborn AM, Clore GM, Louis JM, Wingfield PT. Protein Sci. 1999; 8:426–429. [PubMed: 10048336]
25. Kwan AH, Mobli M, Gooley PR, King GF, Mackay JP. FEBS J. 2011; 278:687–703. [PubMed: 21214860]
26. Cavanagh J, Fairbrother W, Palmer AG III, Sklepton NJ. 1996
27. Gronenborn AM, Frank KM, Clore GM. FEBS Lett. 1996; 398:312–316. [PubMed: 8977129]
28. Selenko P, Serber Z, Gadea B, Ruderman J, Wagner G. Proc Natl Acad Sci USA. 2006; 103:11904–11909. [PubMed: 16873549]
29. Gronenborn AM, Clore GM. Protein Sci. 1996; 5:174–177. [PubMed: 8771212]
30. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. J. Biomol. NMR. 1995; 6:277–293. [PubMed: 8520220]
31. Johnson BA, Blevins RA. J. Biomol. NMR. 1994; 4:603–614.

ATG	CAG	TAC	AAG	CTT	GCT	CTG	AAC	GGT	AAA	ACC	CTG	AAA	GGT	GAA	ACC	ACC	ACC	GAA	GCT	GTT	GAC
M	Q	Y	K	L	A	L	N	G	K	T	L	K	G	E	T	T	T	E	A	V	D
GCT	GCT	ACC	GCG	GAA	AAA	GTT	TTC	AAA	CAG	TAC	GCT	AAC	GAC	AAC	GGT	GTT	GAC	GGT	GAA	TGG	ACC
A	A	T	A	E	K	V	F	K	Q	Y	A	N	D	N	G	V	D	G	E	W	T
TAC	GAC	GAC	GCT	ACC	AAA	ACC	TTC	ACG	GTA		ACC	GAA	CTG	GTT	CCG	CGT	GGA	TCC	ATG	ATC	TCT
Y	D	D	A	T	K	T	F	T	V		T	E	L	V	P	R	G	S	M	I	S
GAT	GAA	CAG	CTG	AAC	TCC	TTG	GCC	ATC	ACC	TTC	GGT	ATT	GTG	ATG	ATG	ACT	TTA	ATT	GTC	ATT	TAC
D	E	Q	L	N	S	L	A	I	T	F	G	I	V	M	M	T	L	I	V	I	Y
CAT	GCT	GTT	GAC	TCC	ACC	ATG	TCT	CCT	AAG	AAC	CGC	CTC	GAG	CAC	CAC	CAC	CAC	CAC	CAC	CAC	TGA
H	A	V	D	S	T	M	S	P	K	N	R	L	E	H	H	H	H	H	H	H	Stop

Fig. 1.

Nucleotide (top) and corresponding amino acid (bottom) sequences of GB1-Ost4p. Amino acid sequence in **bold** represents GB1 sequence. Amino acids in **red** are the first amino acids in the GB1 or Ost4p sequence. Amino acids in **cyan background** represent thrombin cleavage site. Amino acids in **yellow background** are those which are not present originally in OSt4 sequence. Amino acids in **gray background** are C-terminal hexa-Histidine tag that was used to purify GB1-Ost4p using Ni-NTA immobilized metal ion affinity chromatography (IMAC).

Fig. 2.

Expression, purification and cleavage of GB1-Ost4p monitoring by Tris-Tricine SDS-PAGE. (A) Expression of GB1-Ost4p in BL21(DE3)pLysS using the Ni^{2+} -NTA column. Lane 1: Protein marker, Lane 2+3: Cells before induction with 1mM IPTG, Lane 4: Cells after induction with IPTG, Lane 5: Lysate before passing through Ni-NTA column, Lane 6: Lysate after passing through Ni-NTA column after binding solution. Lane 7: Washing. Lanes 8–11: Elutions of GB1-Ost4p. Lane 12: Protein Marker. (B) Optimization of Cleavage Reaction of GB1-Ost4p. Lane -1 is marker, lanes 2 and 3 are pure Ost4p (purified on Ni-NTA column), and lanes 4–9 are cleavage products of GB1-Ost4p. Although about 5 % of the GB1-Ost4p is not cleaved in this gel, we have optimized the cleavage reaction successfully such that GB1-Ost4p fusion protein is not detectable by SDS PAGE and mass spectrometry analysis.

Fig. 3.

Characterization of unlabeled and ^{15}N labeled GB1-Ost4p protein by mass spectrometry. (A) Mass spectrometry analysis of unlabeled GB1-Ost4p. (B) Mass spectrometry analysis of fully cleaved unlabeled GB1-Ost4p. (C) Mass spectrometry analysis of partially cleaved ^{15}N labeled GB1-Ost4p. (D) Mass spectrometry analysis of fully cleaved ^{15}N labeled GB1-Ost4p.

OST4p (wild type)

Fig. 4. Circular Dichroism (CD) spectrum of pure Ost4p in 20 mM phosphate buffer, pH 6.5 containing 100 mM SDS.

Fig. 5. 2D $\{^1\text{H}, ^{15}\text{N}\}$ - heteronuclear single quantum coherence (HSQC) spectrum of ^{15}N labeled GB1-Ost4p and pure Ost4p. (A) HSQC spectrum of ^{15}N labeled GB1-Ost4p in 50 mM phosphate buffer containing 5 % D_2O at pH 6.5. (B) HSQC spectrum of ^{15}N labeled Ost4p in $\text{CDCl}_3:\text{CD}_3\text{OH}:\text{H}_2\text{O}$ (4:4:1). 2D $\{^1\text{H}, ^{15}\text{N}\}$ HSQC NMR is a very useful technique for investigating protein folding and conformation. Each peak in the spectrum represents an amino acid residue in the protein under investigation except Proline. The peaks are well dispersed for both ^{15}N -GB1-Ost4p (A) and ^{15}N -Ost4p (B) indicating that both the proteins are well folded. Peaks corresponding to the amino acid residues of the GB1 protein have been labeled in the spectrum of GB1-Ost4p fusion protein (A). The spectrum of Ost4p (B) indicates a helical protein due to narrow dispersion in the amide proton region which is typical for helical proteins.