EXPRESSION, PURIFICATION, NMR STRUCTURE, AND MOLECULAR DYNAMICS STUDIES OF Ost4 AND Ost4V23D: A CRITICAL SUBUNIT OF YEAST OLIGOSACCHARYLTRANSFERASE

By

BHARAT PRASAD CHAUDHARY

Bachelor of Science in Chemistry
Butwal Multiple Campus
Butwal, Nepal
2005

Master of Science in Chemistry
Tribhuwan University
Kirtipur, Nepal
2007

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfilment of the
requirements for
the degree of
DOCTOR OF PHILOSOPHY
May, 2021

EXPRESSION, PURIFICATION, NMR STRUCTURE, AND MOLECULAR DYNAMICS STUDIES OF Ost4 AND Ost4V23D: A CRITICAL SUBUNIT OF YEAST OLIGOSACCHARYLTRANSFERASE

Dissertation Approved:

Dr. Smita Mohanty

Dissertation Adviser

Dr. Christopher Fennell

Committee Member

Dr. Richard Bunce

Committee Member

Dr. Donghua Zhou

Outside Committee Member

Acknowledgement:

I would like to express my deepest gratitude to my thesis advisor, Dr. Smita Mohanty for providing me with an opportunity to conduct cutting edge research on a life-related project in her laboratory. My research would not be possible without her training, guidance, support, encouragement throughout my graduate study. I thank her for being not only my research advisor but also my parent who has been guiding me to become successful in the future. I would like to express heartfelt gratitude to committee members in my thesis, Dr. Richard Bunce, Dr. Christopher Fennell, and Dr. Donghua Zhou for their help, support, and encouragement in completing my dissertation. I would like to thank Dr. Suman Mazumder, a postdoc with Dr. Smita Mohanty, with whom I had a chance to work on various projects and learned a lot about protein expression, purification and characterization. I would like to thank Dr. David Zoetewey for helping me with the resonance assignments and structure calculations of my protein by solution NMR. I would like to thank Dr. Mohiuddin Ovee with whom I had a chance to learn about NMR data collection and processing. I would like to thank Dr. Donghua Zhou and his student Hem Moktan for helping to reconstitute a membrane protein into the lipid bilayer. I would like to thank Dr. Jochem Struppe, a senior application scientist at Bruker Biospin corporation, Billerica, MA, USA, for collecting all the ssNMR data required for the ssNMR structure in the lipid bilayer. I would like to thank Dr. Martin McCullagh, assistant professor at OSU, for helping by performing and interpreting a portion of MD simulation results. I would like

to thank my colleagues Salik Ram Dahal, Omar Al-Danoon, my fellow graduate students Vishwanath Nukala and Arth Patel, Dr. Mohanty's undergraduate students, Jacob Lewellen, and Ishani Ray, for being supportive and helpful during the research study. I would like to thank former undergraduate students with Dr. Mohanty, Benton Bishop, Gage Calhoun, Courtney Dunn, Keely Ratcliff, and Kayley Powell for helping me in the lab. I would like to thank Dr. Thomas Webb, an emeritus professor at Auburn university, for helping with the proofreading of our research articles to increase the importance of articles. I would like to thank the former chair of the Department of Chemistry Dr. Frank Blum for his support and encouragement towards the successful completion of my Ph. D. dissertation. I would like to thank all the professors, administrative staff at the Department of Chemistry, and all my friends at Oklahoma State University for helping me out directly or indirectly to complete my Ph. D. successfully. I would like to thank U.S. Department of Agriculture NIFA Award number 2011-65503-23501, National Science Foundation Grant ISO-0628064, and National Institutes of Health Grant DK082397, Oklahoma State University startup fund AA-1-52223, National Science Foundation Award CHE-1807722 and DBI-1726397 to Dr. Smita Mohanty. I would like to thank National High Magnetic Field Laboratory (NHMFL), Tallahassee, FL, which is supported by NSF Cooperative Agreement No. DMR-1644779 and the State of Florida. The majority of the NMR data reported in this dissertation were acquired at the NHMFL and my dissertation would not be possible without their contribution. I would like to thank my parents, my wife's parents, all my family members, and relatives for their continuous support and help. At last, I would like to thank my wife, Mina K Chaudhary, for her continuous love, encouragement, being on my side in any situation, and support.

Name: BHARAT PRASAD CHAUDHARY

Date of Degree: MAY, 2021

Title of Study: EXPRESSION, PURIFICATION, NMR STRUCTURE, AND MOLECULAR DYNAMICS STUDIES OF Ost4 AND Ost4V23D: A CRITICAL SUBUNIT OF YEAST OLIGOSACCHARYLTRANSFERASE

Major Field: CHEMISTRY

Abstract: *N*-linked glycosylation is an essential and highly conserved protein modification reaction that takes place in all eukaryotes and some prokaryotes. This reaction is catalyzed by the enzyme complex, oligosaccharyltransferase (OST). In the central step of the *N*-linked glycosylation reaction, the pre-assembled high mannose oligosaccharide moieties are transferred from dolichol linked donors to the side chain of a specific asparagine residue in the Asn-X-Thr/Ser (where X ≠ proline) sequence of the nascent protein. Genetic defects in the *N*-linked glycosylation in humans results a group of disorders known as congenital disorders of glycosylation (CDG) that include but are not limited to mental retardation, developmental delay, hypoglycemia etc. The complete loss of *N*-linked glycosylation is lethal to all organisms.

In *Saccharomyces cerevisiae*, the functional OST is composed of eight of the nine non-identical integral membrane protein subunits. The subunits Wbp1, Swp1, Ost1, Ost2, and Stt3p are essential for the viability of cells. The subunits Ost3 / Ost6 and Ost5 are non-essential but are required for optimal OST function. The subunit Ost4 is essential for the growth of cells at 37^{o} C, but not at 25^{o} C. Ost4 is the smallest subunit critical for the OST activity and the stability of the Stt3-Ost4-Ost3 sub-complex. Any mutation of the residues from 18 to 24 to a charged residue results in the destabilization of the sub-complex and impairs cell growth and *in vitro* OST activity. Mutation of valine (V) at position 23 in Ost4 to aspartate (D) causes defects in the *N*-linked glycosylation process. To understand the structure, function and role of Ost4 in *N*-linked glycosylation, characterizations of Ost4 and its functionally important mutant/s are critical.

My doctoral dissertation is focused on the following three parts: (1) production and biophysical characterization of Ost4V23D, (2) 3D structure determination of the Ost4 and Ost4V23D in DPC micelles by solution-state NMR and molecular dynamics (MD) simulation, and (3) 3D structure determination of Ost4 and Ost4V23D in bilayer by solid state NMR followed by MD simulation. A comparison of the structure of the V23D mutant protein to its wild-type to reveal how the mutation affects the overall structure and function of the enzyme. Additionally, we have shed light on the molecular basis of why a point mutation of certain hydrophobic residues to charged residues destabilizes the catalytic sub-complex rendering the OST enzyme dysfunctional.

TABLE OF CONTENTS

List of Table

List of Figures

List of Abbreviations

NMR: Nuclear magnetic resonance

FT: Fourier transform

TMS: Tetramethylsilane

DSS: 2,2-dimethylsilapentane-5-sulfonic acid

NOE: Nuclear Overhauser effect

RF: Radio frequency

HSQC: Heteronuclear single quantum coherence

COSY: Correlation spectroscopy

TOCSY: Total correlation spectroscopy

NOESY: Nuclear Overhauser effect spectroscopy

cryo-EM: Cryo-electron microscopy

MAS: Magic angle spinning

ssNMR: Solid-state NMR

CP: Cross polarization

DARR: Dipolar assisted rotational resonance

PDSD: Proton driven spin diffusion

OST: Oligosaccharyltransferase

ER: Endoplasmic reticulum

CDG: Congenital disorders of glycosylation

LLO: Lipid-linked oligosaccharide

AglB: Archaeal glycosylation B

PglB: Protein glycosylation B

DPC: Dodecyl phosphocholine

PDB: Protein data bank

NTA: Nitrilotriacetic acid

SDS PAGE: Sodium dodecyl sulfate polyacrylamide gel electrophoresis

CD: Circular dichroism

WT: Wild type

IPTG: Isopropyl-β-D-thiogalactopyranoside

BMRB: Biological magnetic resonance data bank

SSP: Secondary structure propensity

TALOS: Torsion angle likelihood obtained from shift

CSI: Chemical shift index

MD: Molecular dynamics

RMSD: Root mean squared deviation

NPT: Constant number of particles, pressure, and temperature

NVT: Constant number of particles, volume, and temperature

DPPC: Dipalmitoylphosphatidylcholine

SASA: Solvent accessible surface area

CSP: Chemical shift perturbations

RCI: Random coil index

NESG: Northeast structural genomics consortium

TM: Transmembrane

TMH: Transmembrane helix

POPC: 1-palmitoyl-2-oleoyl-snglycero-3-phosphocholine

POPE: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine

DMPC: 1,2-dimyristoyl-sn-glycero-3-phosphocholine

DMPG: 1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol)

POPG: 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)

DOPC: 1,2-dioleoyl-sn-glycero-3-phosphocholine

DOPE: 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine

LUV: Large unilamellar vesicle

CHAPS: 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate

RB: Round bottomed

DANGLE: Dihedral angles from global likelihood estimates

CHAPTER 1

Introduction

1.1. NMR

1.1.1 History and principles of NMR

Nuclear Magnetic Resonance (NMR) spectroscopy was first introduced by Isidor Rabi in 1938. He was awarded the Nobel Prize in 1944. Two physicists, Felix Bloch and Edward Mills Purcell, extended the NMR technique for its use in solution and solid state in 1946. They were awarded the Nobel Prize in 1952. Since its discovery, NMR spectroscopy has been through some major theoretical and technical advancements. Richard R. Ernst introduced the Fourier transform (FT) to NMR in 1964. He was awarded a Nobel Prize in 1992 for this discovery. In this method, a single short pulse or a series of pulses in the form of radiofrequency is applied to the sample. The signal emitted by the nuclear spins in the response to the pulses is Fourier transformed to get an NMR spectrum. This yields a much better signal to noise (S/N) ratio. In the 1980s and 1990s, homo-and hetero-nuclear multidimensional NMR were developed, in which signals were acquired as a function of multiple frequency variables. The determination of high-resolution structures of macromolecules and their complexes was possible due to the hetero-nuclear multidimensional NMR. In multidimensional NMR, the resonances are spread into more than one dimension resulting in enhanced resolution.

In the early days, NMR was used as an analytical tool to determine the purity, quality, and content of the samples. In the modern era, NMR has become a versatile method to determine the molecular structures of chemical and biomolecules. NMR is based on the principle that all nuclei of a spinning atom have an intrinsic angular momentum of a spinning sphere. Based on the theory of quantum mechanics, for a nucleus containing spin, the optimum observable component of the angular momentum is a half-integral or integral multiple of $h/2\pi$, where h is Plank's constant. For an atomic nucleus possessing angular momentum P, and spin quantum number I. The optimum observable component of the angular momentum can be represented by:

$$P = \frac{Ih}{2\pi} \qquad (1.1)$$

For an atom having non-zero spin ($I \neq 0$), the magnetic moment, μ, can be defined by the equation:

$$\mu = \frac{\gamma h[I(I+1)]1/2}{2\pi} \qquad (1.2)$$

Where γ is a constant, the gyromagnetic ratio of the nucleus. Depending on the atomic number and mass number of different atomic nuclei, spin quantum numbers, I, can be different. For example, a nucleus of an element containing an even atomic number and an even mass number will have zero spin quantum number, I, and angular momentum of such nuclei will be zero. The spin quantum number can be 1, 2, 3, 4… for a nucleus of an element having an even atomic number and odd mass number. Similarly, a nucleus of an element with an even or odd atomic number and odd mass number can have $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots\ldots$ spin quantum numbers (I).

In the presence of an external magnetic field B_0, a nucleus having a spin quantum number, *I,* can have $2I+1$ possible states or orientations, and each state is related to a different potential energy – the Zeeman splitting. The potential energy of the nucleus in each state is $\mu B_0 \cos\theta$, where θ is the angle between the axis of the spin and the direction of the magnetic field. A nucleus such as ^{1}H, ^{15}N, ^{31}P, ^{19}F having a spin of one-half, has two linearly independent energy states, α and β spin states, in presence of an external magnetic field. In the absence of a magnet field, these states have the same energy, i.e. the states are degenerate.

The lower energy state corresponds to a lower potential energy and a higher energy state corresponds to a higher potential energy of the nucleus. The difference in energy ΔE between the two states is directly proportional to the applied uniform external magnetic field, B_0.

$$\Delta E = \gamma h B_0 \qquad (1.3)$$

And this results in a small population bias favoring the lower energy state in thermal equilibrium. With more spins pointing up than down, a net spin magnetization along the magnetic field $\mathbf{B}_0$ results. The energy orientation states a nucleus having one-half spin can be represented by Figure 1.1.

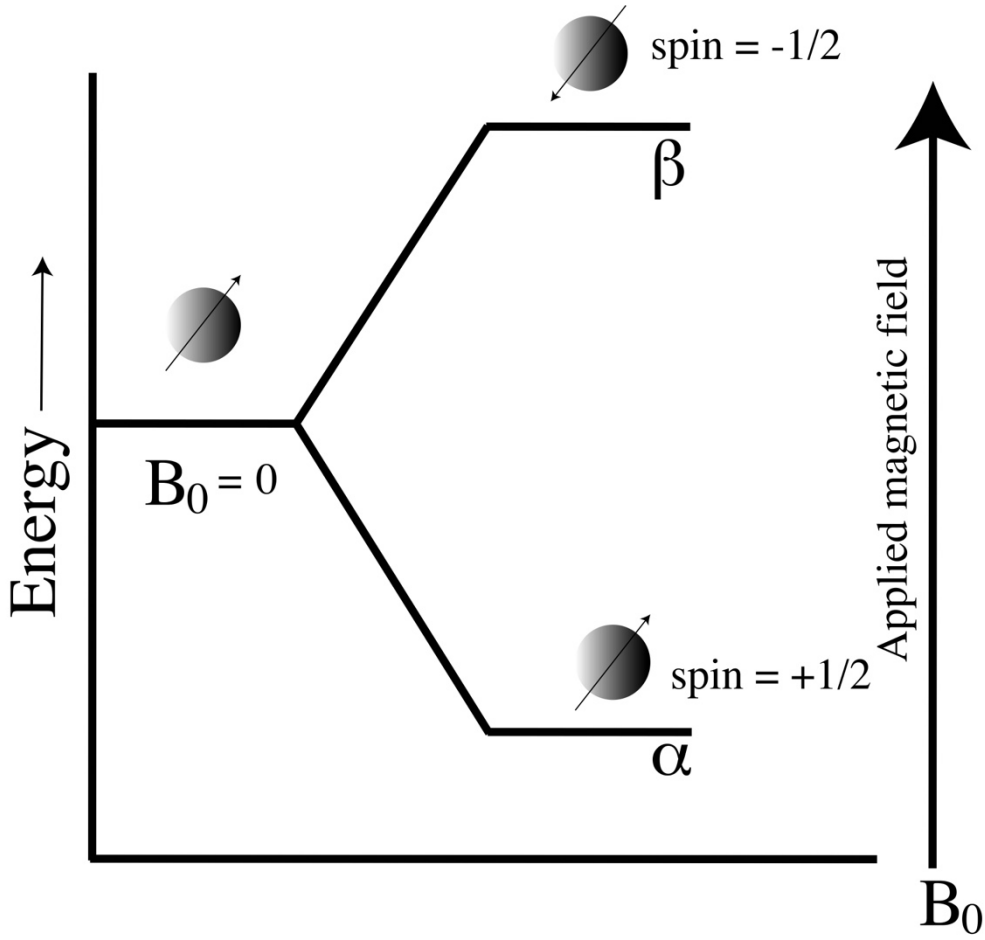


Figure 1.1: Energy states splitting as a function of applied external magnetic field strength. The figure was adapted from reference *(1)*.

For NMR spectroscopy, the frequency of the electromagnetic radiation can be calculated using the following equation.

$$\Delta E = h\nu = \mu B_0/I \qquad \text{(Eq. 1.4)}$$

In the presence of the external magnetic field, the spinning nucleus precesses with a certain frequency called the precession frequency or Larmor frequency. Upon applying a radio frequency wave exactly equal to precession frequency perpendicular to the external magnetic field, the nucleus transition from the lower energy state to the upper energy state takes place. The gyromagnetic ratio, γ, of a nucleus and its radiation frequency, ν are related to each other under the influence of the external magnetic field as:

$$\nu = \gamma B_0/2\pi \qquad \text{(Eq. 1.5)}$$

i.e. $$\nu \propto B_0 \qquad \text{(Eq. 1.6)}$$

The precession frequency or Larmor frequency of a nucleus depends on the gyromagnetic ratio of the nucleus and the applied magnetic field.

1.1.2 Chemical shifts

The chemical shift of a nucleus is the difference between the resonance frequency of the nucleus and the standard. It is expressed in parts per million (ppm). The chemical shifts of different nuclei in a molecule are slightly different depending upon the local chemical environment. The chemical shift of a particular nucleus can be defined by the following equation:

$$\delta = (\nu_{peak} - \nu_{Standard})/\text{frequency of instrument in MHz} \qquad (1.7)$$

Where,

ν_{peak} = resonance frequency of the nucleus in Hz.

$\nu_{Standard}$ = resonance frequency, in Hz, of an internal standard

For ^{1}H and ^{13}C nuclei, tetramethylsilane (TMS) is the most commonly used reference. Nuclei with different molecular environments show different chemical shifts. This is very useful for structure determination by NMR. However, because of its low solubility in water, it is not used as an internal reference for biomolecules in aqueous solvents. Instead, a low concentration of 2,2-dimethylsilapentane-5-sulfonic acid (DSS) is used as the reference for biomolecules in aqueous solutions.

When the external magnetic field is applied to a nucleus, the applied magnetic field is not always equal to the magnetic field experienced by the nucleus. The electrons present around the nucleus create a secondary magnetic field that either shield or deshield it from the applied magnetic field. Therefore, a difference in the field is observed between the applied magnetic field and the actual field experienced by the nucleus. This difference in the field experienced by the nucleus is termed *nuclear shielding*. The s-electrons have spherical symmetry and circulate in the presence of a magnetic field opposing the applied field. This means that the nucleus needs a higher magnetic field to be applied for it to come to the resonance. As a result, an *upfield shift* occurs. This upfield shift is also known as a *diamagnetic shift*. In contrast, the electrons in p-orbitals do not have spherical symmetry. Instead of opposing, they augment the applied field resulting in comparatively large magnetic fields at the nucleus producing a *low field shift*. This *low field shift* is also known as nuclear deshielding or a *paramagnetic shift*.

1.1.3 Relaxation in NMR

For the proper interpretation of NMR spectra, understanding the relaxation is very important. At the equilibrium state of molecules, the population difference between spins at a higher energy level and spins at a lower energy level is zero. This condition is called a degenerate state. The population of the spins at an energy state is defined by Boltzmann distribution. When a molecule is subjected to an external energy source, some of the spins at a lower energy level are excited to the higher energy level. Upon removal of the external energy source, the spins at higher energy states return to the equilibrium state by a process called relaxation. In solution, which is also called an isotropic system, the nuclei relax to equilibrium by two processes: longitudinal or spin-lattice relaxation (T1), and transverse or spin-spin relaxation (T2). In longitudinal relaxation, the excited magnetization vector returns to its thermal equilibrium state. This occurs in the z-direction. The gain or loss of magnetization follows an exponential curve as given in Eq. 1.5.

$$M_z = M_0 [1-e^{-t/T_1}] \qquad \text{(Eq. 1.8)}$$

An example of T1 relaxations is shown in Figure 1.2a

Longitudinal relaxation is due to energy exchange between the spins and anything around it. Since the magnetization vector returns to thermal equilibrium in T1, the spin-lattice relaxation is an enthalpy-driven phenomenon. T1 relaxation is inversely proportional to the magnetic field. The higher the magnetic field, the slower the longitudinal relaxation (*2*).

Transverse or spin-spin relaxation (T2) is the mechanism by which the excited magnetization returns to equilibrium in the x-y direction and this occurs by a decay process. The T2 relaxation process can be described by an exponential curve as given by Eq. 1.9.

$$M_{xy} = M_0 e^{-t/T_2} \qquad \text{(Eq. 1.9)}$$

Figure 1.2b provides an example of the T2 relaxation process.

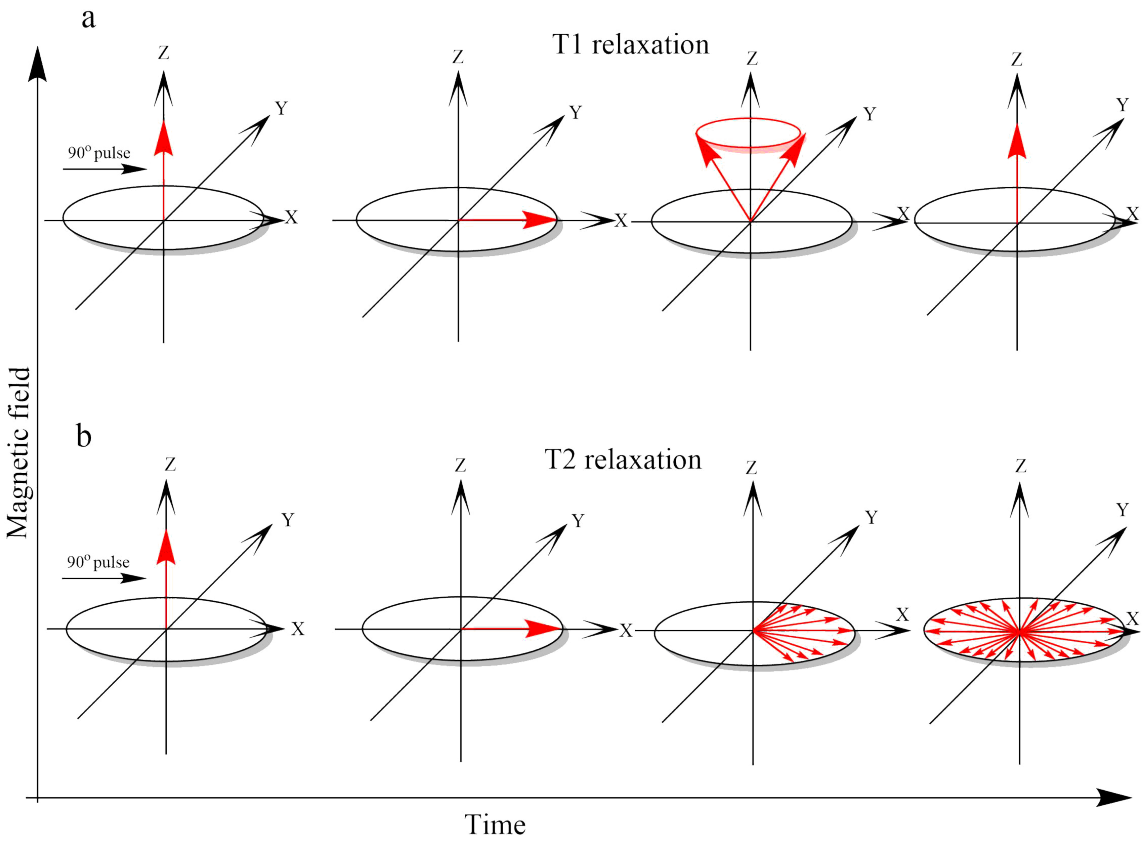


Figure 1.2: Process of relaxation in NMR. (a) Longitudinal relaxation and (b) transverse relaxation.

Transverse relaxation is caused by transient magnetic fields at any frequency and spin-spin interaction. Therefore, slower molecular tumbling results in a shorter T2. In contrast to T1 relaxation, which is an enthalpy-driven process, T2 relaxation results in the loss of phase coherence, and therefore, is an entropy-driven process. Unlike T1, T2 is independent of the magnetic field. Both T1 and T2 can be determined by NMR experiments and T2 is always shorter than T1 (*3*).

1.1.4 Spin-spin coupling

The magnetic interaction between nuclei with non-zero spin is termed spin-spin coupling. Two important types of spin-spin couplings are observed in NMR.

1.1.4.1 Scalar coupling (or J-coupling)

The scalar coupling or J-coupling is the interaction between nuclei through three chemical bonds. Splitting of NMR signals occurs due to scalar coupling. The splitting of NMR peaks occurs due to two spin states, alpha and beta, of a nucleus. The two nuclei connected by three chemical bonds have different spin states and influence the magnetic field of each other. A nucleus, which is bonded to another nucleus augmenting the applied magnetic field, comes to resonance at a slightly lower magnetic field. In contrast, a nucleus bonded to another nucleus that opposes the applied magnetic field come to resonance at a higher magnetic field. In scalar coupling, the spin of one nucleus perturbs the spins of related electrons which in turn perturbs the energy levels of neighboring nuclei. This changes the resonance frequency of the neighboring nucleus resulting in the splitting of signal lines. The frequency difference between the split signal lines is termed as a J-coupling constant. Scalar coupling provides much information about a molecule. The three-bond scalar coupling provides information about the dihedral angle. The coupling pattern gives insight into the connectivity of atoms in a molecule. The coherence transfer in multidimensional NMR is possible due to scalar coupling. Since protein NMR is dependent on multidimensional NMR, J-coupling constants are crucial for protein NMR experiments. Moreover, J-coupling constants are very important for the investigation of drugs and a variety of small molecules.

1.1.4.2 Dipolar coupling

The second type of coupling is dipolar coupling. In dipolar coupling, the spin-spin interaction takes place through space. Dipolar coupling results in some essential phenomenon in NMR spectroscopy, such as spin-spin relaxation and Nuclear Overhauser Effect (NOE). Spin-spin relaxation essentially determines the peak linewidth. NOE is responsible for the change in the intensity of the peak. The closer peak will have a larger peak intensity compared to the farther ones due to the NOE effect. This is the key to determine the structure of molecules and interaction processes between molecules as well.

1.1.5 Nuclear Overhauser effect (NOE)

When a nucleus is perturbed by applying a radio frequency (RF), it attempts to relax back to its equilibrium state by scalar coupling relaxation mechanism. However, the scalar coupling relaxation is not sufficient to relax the nucleus back to the equilibrium state. In such a case, the nucleus reaches its equilibrium state with an additional dipolar relaxation process. The irradiated nucleus transfers some of its energy, during relaxation by dipolar coupling method, to any nucleus nearby in space. Then, the second nucleus acts as if it is perturbed with the RF. Therefore, this second nucleus relaxes back to equilibrium increasing its intensity. This process is called the Nuclear Overhauser Effect (or NOE).

If two nuclei X and Y having two spin states, alpha and beta respectively, are coupled through the dipolar coupling, the two nuclei contain four energy states: $\alpha\alpha$ (representing both spins at lower energy states), $\alpha\beta$ (X at lower and Y at higher energy states), $\beta\alpha$ (X at higher and Y at lower energy states), and $\beta\beta$ (both spins at higher energy states). There are

two steps of transitions for each spin: αα–>βα and αβ–>ββ for X and αα–>αβ and βα–>ββ for Y (W1). Upon irradiation of one of the nuclei, the population αα becomes equal to βα and αβ are equal to ββ. But there is a still population difference between αα and ββ states. Here, the population states are minimized via dipolar coupling relaxation (W2) resulting in the enhancement of the NMR signal intensity of the Y called a positive NOE. This occurs in the case of small molecules which tumble fast in solution. However, for larger molecules like proteins which tumble slowly in solution, the population difference is restored via another type of relaxation process that occurs between βα and αβ (W0). This results in the decrease of NMR signal intensity of other nuclei known as negative NOE. Figure 1.3 shows all the transitions occurring between X and Y spin systems.

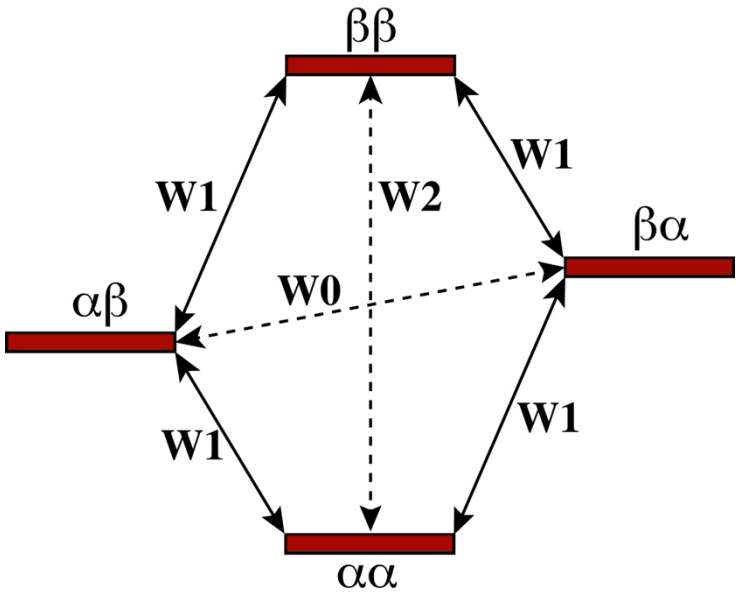


Figure 1.3: Energy diagram for a dipolar-coupled two-spin system. The four states are αα, αβ, βα, and ββ; the zero- single- and double-quantum transitions are represented by W0, W1 and W2, respectively. This figure is drawn according to reference *(2)*.

For medium-sized molecules ranging from 1 KDa to 3 KDa, the positive and negative NOEs compete and sometimes the NOES for such molecules is negligible or zero. NOE between two spins can be detected only if they are close enough (~6Å). The intensity of the NOE between two nuclei separated by distance r is inversely proportional to r^6 (i.e. NOE $\propto r^6$) (*4*).

1.1.6 Multidimensional NMR

Recent advancements in the NMR field have made the use of multidimensional NMR the most effective method for the elucidation of high-resolution structures of macromolecules. 2D NMR provides the data in which the resonances are plotted in space defined by two axes. Jean Jeener was the first scientist who proposed the first 2D NMR in 1971, which was later implemented by Walter P. Aue, Richard R. Ernst, and Enrico Bartholdi (*5, 6*). Unlike 1D NMR, two additional factors, evolution period and mixing time, are included in 2D NMR. Therefore, a 2D NMR experiment contains the preparation period, the evolution time, mixing, and the acquisition as shown in Figure 1.4.

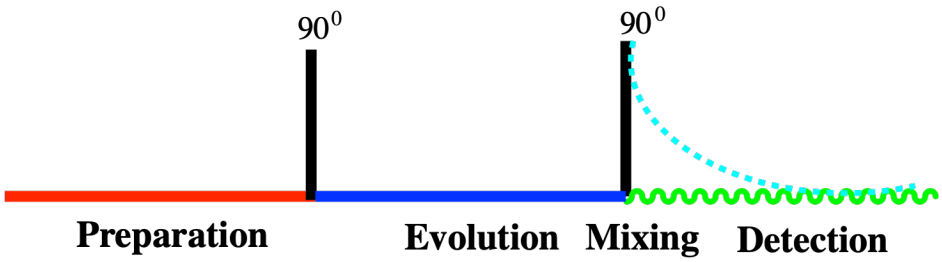


Figure 1.4: A schematic representation of the 2D pulse program. The schematic was adapted from reference (*4*).

Some of the important 2D NMR experiments include Heteronuclear Single Quantum Coherence (HSQC), Correlation Spectroscopy (COSY), Total Correlation Spectroscopy (TOCSY), and Nuclear Overhauser Effect Spectroscopy (NOESY).

1.1.6.1 HSQC

HSQC stands for heteronuclear single quantum coherence. The HSQC generates an NMR signal for two different types of magnetically active nuclei that are bonded together. In general, HSQC contains two axes: the proton chemical shift and a heteronuclear chemical shift. The ^{1}H nucleus has a larger Boltzmann population difference, the highest gyromagnetic ratio, and the higher detection sensitivity. Because of these reasons, it favors the magnetization transfer to and from the heteronucleus. And the heteronuclear experiments in which ^{1}H is utilized as a passage by the heteronucleus have a much better sensitivity (signal to noise, S/N). Therefore, in the HSQC experiment, the magnetization is transferred from ^{1}H to a directly connected heteronucleus and then transferred back to the proton for detection. Generally, two types of HSQCs are used in protein NMR: [^{1}H, ^{15}N]-HSQC and [^{1}H, ^{13}C]-HSQC. In [^{1}H, ^{15}N]-HSQC spectrum, each cross peak is due to the amide proton and nitrogen (amide) in the peptide bond of the protein. Therefore, the number of peaks in this experiment is equal to the number of non-proline residues present in the protein. This experiment is very sensitive to changes in the chemical environment, pH, temperature etc. Due to this reason, this experiment is also known as the fingerprint of the protein.

1.1.6.2 COSY

COSY is an example of the most popular multidimensional NMR spectroscopy which is used to identify the spin systems that are connected by a 3-bond distance. This allows us to identify the resonances that are connected by J-coupling.

Figure 1.5 represents an example of the COSY spectrum.

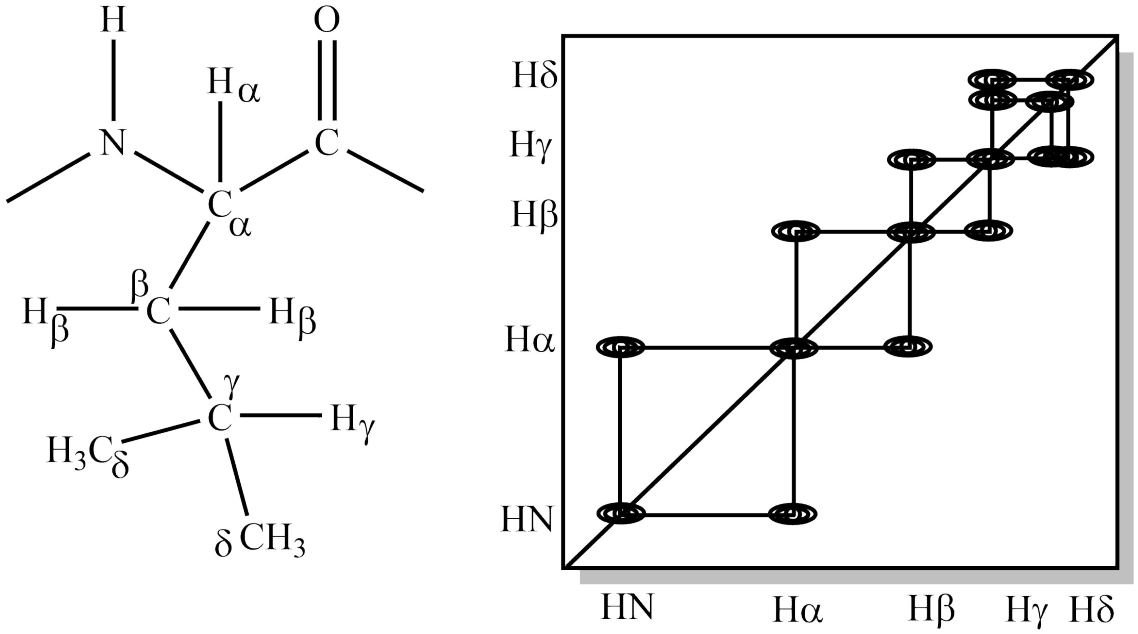


Figure 1.5: Example of 2D COSY spectrum for leucine. COSY provides a diagonally symmetric spectrum.

1.1.6.3 TOCSY

TOCSY stands for Total Correlation Spectroscopy. In TOCSY experiments, the magnetization is kept constant (spin-locked) by applying a certain amount of RF frequency. This time is called mixing time. During this mixing time, the coherence transfer takes place through the scalar coupling. As a result, resonances of all the protons attached to a particular spin system are observed. Figure 1.6 provides a general idea about the TOCSY experiment. In protein NMR, this experiment is used to identify the sidechain protons of

residues.

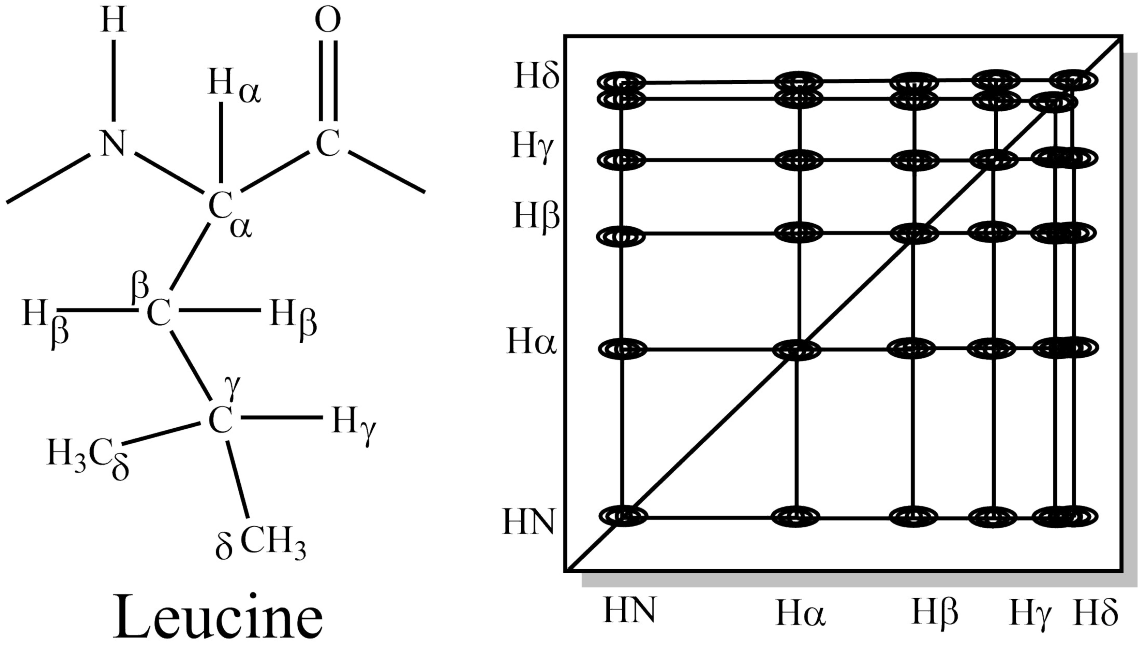


Figure 1.6: Schematic representation of the spectrum of a 2D TOCSY for the leucine residue. This experiment shows all sidechain hydrogen atoms of an amino acid residue. The spectrum is diagonally symmetric.

1.1.6.4 NOESY

NOESY is an abbreviated form of Nuclear Overhauser Effect Spectroscopy. The NOESY is based on the NOE theory developed by American physicist Albert Overhauser who proposed that *nuclear* spin polarization could be augmented by the microwave irradiation of the conduction *electrons* in certain metals (*7*). NOESY utilizes the dipolar coupling to transfer the magnetization. It provides information about the distances between nuclei in space (within 5 Å). NOESY, in protein NMR, is an essential experiment for the determination of protein structure. The intensity of cross-peaks provides the distance between two nuclei in NOESY.

1.1.6.5 3D NMR

For small proteins having a molecular weight less than 10 kDa, 2D proton NMR can be used to elucidate the structure. But as the molecular size of the molecule increases, the 2D proton NMR can not be used because of several reasons. The main reason is that the 2D proton NMR spectra become very complex for analysis and hinder the unambiguous assignment of the resonances of larger biomolecules in solution. In addition, large biomolecules tumble slowly in solution. This results in the line broadening in the 2D spectrum of the molecules. As a consequence, the spin systems of the molecule are severely overlapped for unambiguous resonance assignment. Thus, 2D proton NMR fails for the structure determination of larger biomolecules in solution.

These limitations can be overcome by both increasing the dimension of the 2D NMR to 3D and if needed to 4D NMR along with the utilization of other nuclei such as ^{15}N and ^{13}C. Thus, heteronuclear 3D NMR experiments solve the resonance overlap problem due to the line broadening of macromolecules. Since heteronuclear coupling is used in the case of 3D NMR, a ^{13}C, ^{15}N (double) labeled protein sample is required to perform 3D NMR experiments for larger proteins. Two 2D NMR pulse sequences can be combined to generate a 3D pulse sequence. During the construction of a 3D pulse sequence, the acquisition time of the first 2D pulse sequence and the preparation time from second 2D pulse sequence are removed and then the rest from each 2D pulse sequence are combined to yield a 3D pulse sequence. The sensitivity and digital resolution decrease with the increase in dimensionality in NMR. This limits the increase in dimensions in NMR experiments to a maximum of 4 dimensions or 4D NMR.

1.1.6.6 Protein NMR

NMR spectroscopy, X-ray crystallography, and cryo-electron microscopy (cryo-EM) are the three main methods for protein 3D structure determination at atomic resolution. Although X-ray crystallography is the predominant method for protein structure determination, it has certain limitations; for example, many proteins especially membrane proteins are difficult to crystallize. Additionally, if some of these difficult proteins can be crystallized, they may not diffract well to generate the structure. Furthermore, crystallography can not provide information on the dynamic behavior of proteins in solution. Although the cryo-EM method has made significant advancements recently for the determination of atomic resolution structures of large biomolecules and their complexes, it also cannot provide any information on the dynamics of the biological molecule under study (*8*). On the other hand, NMR spectroscopy is routinely used for structure, dynamics, and interactions studies for small to medium size biological molecules in solution (*9-11*). Solution-state NMR provides information on protein dynamics on a picosecond to second time scale. However, NMR is an intrinsically an insensitive technique and requires large amounts (both concentration and volume) of isotope-labeled samples to get meaningful data. Recent advancements in NMR techniques such as NMR instrumentation, magnetic field strengths, microcoils, cryogenically cooled probes, dynamic nuclear polarization, and sensitivity-enhanced pulse sequences have made it possible to acquire good quality NMR data on a protein sample at submicromolar concentrations.

Atomic resolution structure determination of proteins by NMR spectroscopy include sample preparation, NMR data collection, data processing, data analysis, protein backbone assignment, sidechain assignment, NOE assignment, structure calculation, and refinement.

The refinement of the calculated NMR structure is an iterative process that takes several rounds of iteration. Therefore, it can take a tremendous amount of time before the final refined structure is obtained.

1.1.6.6.1 Backbone chemical shift assignments

The complete protein backbone assignment can be achieved by three pairs of 3D NMR experiments: (i) CBCA(CO)NH and HNCACB (*12*), (ii) HNCA and HN(CO)CA (*13-15*), and (iii) HNCO and HN(CA)CO (*14-16*). Each of these NMR spectrums has a [^{1}H, ^{15}N] HSQC plane in the X-Y dimension and a ^{13}C plane in the Z-dimension.

The 3D CBCA(CO)NH and 3D HNCACB Pair:

The 3D CBCA(CO)NH and 3D HNCACB NMR experiments are the main pair for backbone assignments of the protein residues. In this pair, the magnetization is transferred from Hβ to Cβ and Hα to Cα, and again from Cβ to Cα. Finally, to N and HN. In CBCA(CO)NH, the chemical shift is evolved on N and HN of its own (i) residue and Cα and Cβ of the preceding (i-1) residue. In HNCACB, the chemical shift is observed for N, HN, Cα, and Cβ of its residue and Cα, and Cβ of preceding residue. For this reason, this 3D NMR pair is very helpful in walking through the protein backbone to assign the backbone nuclei unambiguously.

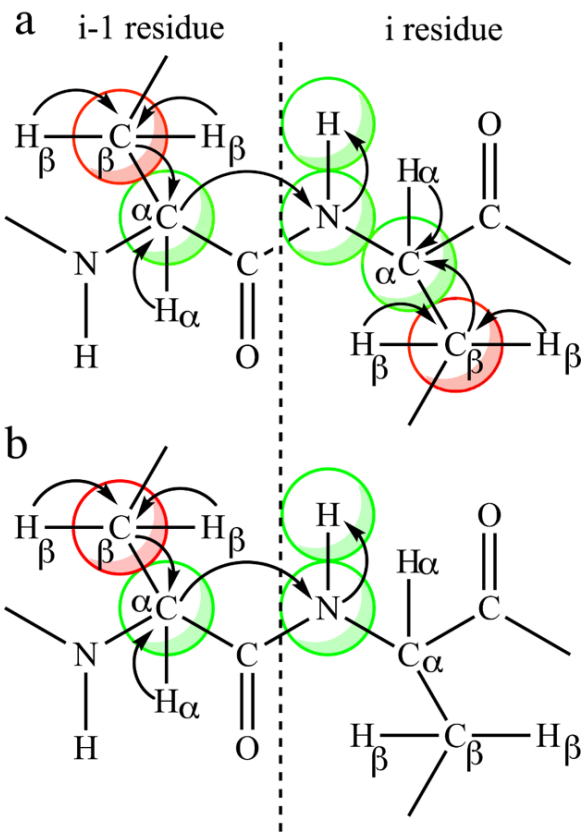


Figure 1.7: Magnetization transfer process and chemical shift evolvement in (a), HNCACB and (b), CBCA(CO)NH experiments. HNCACB provides CA and CB resonances of its own and its preceding residues whereas CBCA(CO)NH provides CA and CB of the preceding residue only. This pair is very essential for walking through the backbone of a protein sequence.

The HNCA and HN(CO)CA pair:

In HNCA, the chemical shift is evolved on N, HN, and Cα of its i residue and Cα of the i-1 residue (Figure 1.8). This spectrum is useful if the HNCACB and CBCA(CO)NH are of bad quality. Similarly, HN(CO)CA provides chemical Cα chemical shift of the preceding residue only (Figure 1.8).

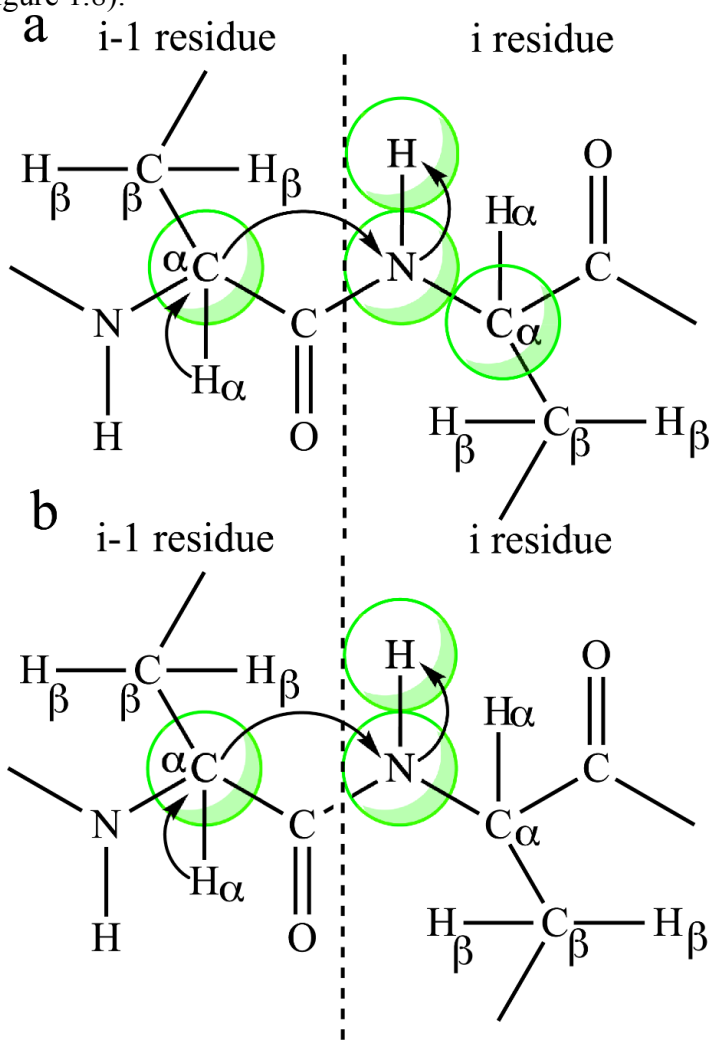


Figure 1.8: Magnetization transfer and chemical shift evolvement in (a), HNCA and (b), HN(CO)CA. Green circles indicate the nuclei for which chemical shift is observed.

The HNCO and HN(CA)CO pair:

In this pair, HNCO provides chemical shifts of carbonyl (CO) of the preceding residue. The HN(CA)CO gives the CO chemical shifts of its own and the preceding residue.

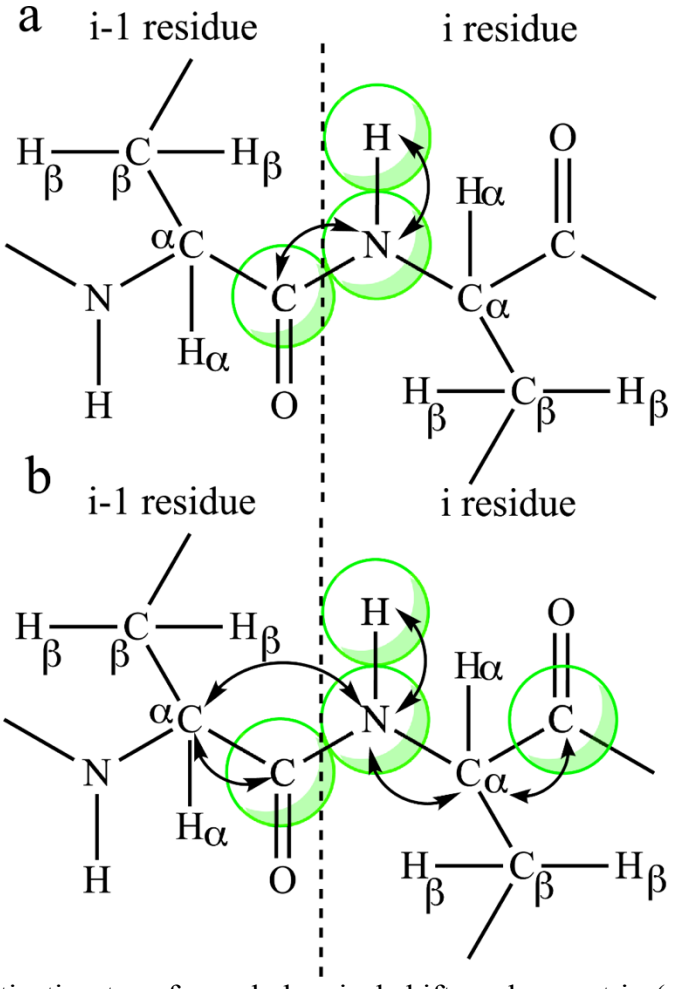


Figure 1.9: Magnetization transfer and chemical shift evolvement in (a), HNCO and (b), HN(CA)CO. Green circles indicate the nuclei for which a chemical shift is observed.

This pair is used to link the residues during the backbone assignment. Figure 1.9 provides the scheme of magnetization transfer in this pair.

After complete backbone assignment of the residues of a protein, the sidechain of the residues is assigned using different types of TOCSY experiments. The commonly used

TOCSY experiments are ^{15}N-edited 3D HSQC-TOCSY (*17*) and 3D H(CCO)NH-TOCSY (*18*). Sometimes, 3D HCCH-TOCSY (*19*), a carbon-based 3D TOCSY, is used to get the sidechain protons.

1.1.6.6.2 Side-chain chemical shift assignments

In the ^{15}N-edited 3D HSQC-TOCSY experiment, the magnetization transfers between ^{1}H nuclei *via* isotropic mixing followed by the magnetization transfer to the neighboring ^{15}N nuclei and back to ^{1}H for detection (Figure 1.10 a). In this experiment, the chemical shift evolves on the amide nitrogen, proton, and the side-chain proton of the i residue (Figure 1.10 a). This is an important experiment for the identification of the amino acid type and side-chain protons. Apart from HSQC-TOCSY, H(CCO)NH TOCSY and HCCH TOCSY are additional sets of experiments that are used for assigning side-chain protein nuclei. The former is ^{15}N-based TOCSY and the later one is ^{13}C-based TOCSY. In H(CCO)NH TOCSY, the magnetization is transferred from the ^{1}H nuclei to their attached ^{13}C nuclei. Then, the magnetization is transferred to all side-chain carbon nuclei via isotropic mixing from where the magnetization is finally transferred to the carbonyl carbon, amide nitrogen and proton for the detection. The chemical shift is evolved on amide the ^{15}N and ^{1}H nuclei of the i residue and the side-chain ^{1}H nuclei of the i-1 residue (Figure 1.10 b). In ^{13}C HCCH TOCSY, the magnetization is transferred from ^{1}H nuclei to their attached ^{13}C nuclei and then to other carbon nuclei *via* isotropic mixing. Finally, the magnetization is transferred to the side chain ^{1}H nuclei for detection. The chemical shift is evolved on side-chain ^{13}C and ^{1}H nuclei of i residue (Figure 1.10 c).

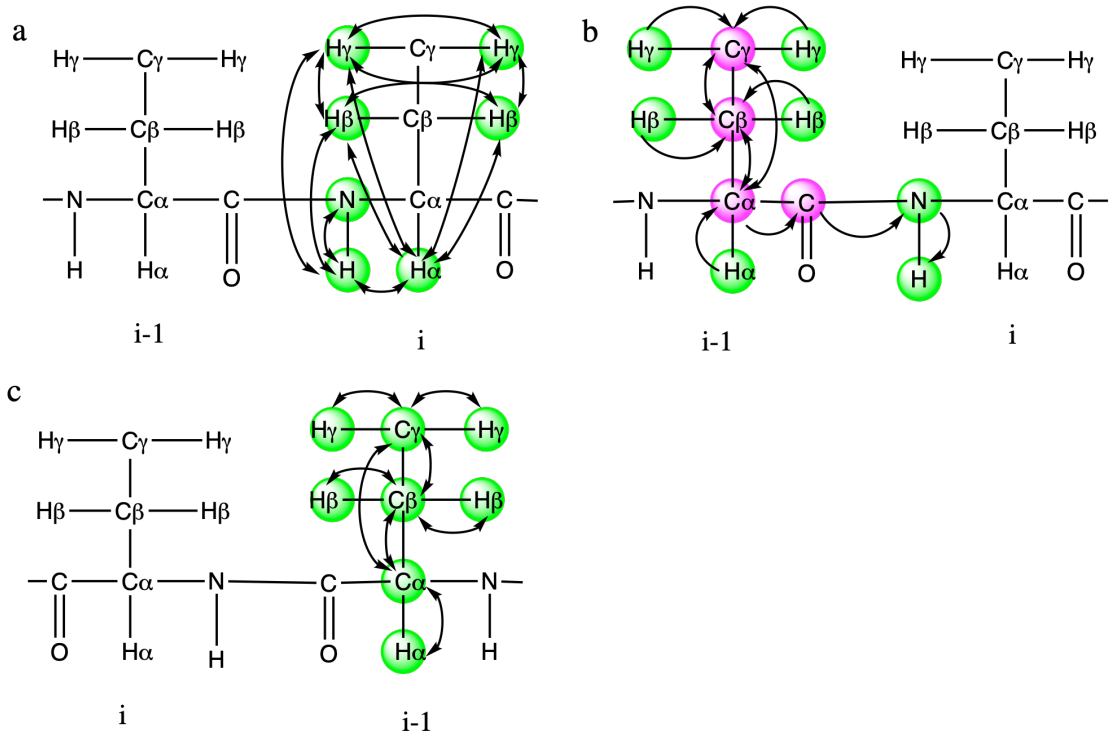


Figure 1.10: Magnetization transfer and chemical shift evolvement in 3D ^{15}N HSQC TOCSY(a), 3D ^{15}N H(CCO)NH TOCSY (b), and 3D ^{13}C HCCH TOCSY. Green circles indicate the nuclei on which a chemical shift is observed. Purple circles indicate the nuclei that are used for magnetization transfer.

1.1.6.6.3 NOE assignments

After the backbone protons are assigned, the final step of the assignment is the NOE assignment. The intensity of a cross peak between a pair of protons provides the distance between them. This is achieved by assigning the NOE peaks. Experiments such as ^{15}N-edited 3D HSQC-NOESY (*17, 20, 21*) and ^{13}C- edited 3D HSQC-NOESY (*21*) are used to get the NOEs between the protons. Once the NOEs are assigned, the structure calculation of the protein can be initiated by creating different restraint files. The files include the

chemical shift list, dihedral angle restraints, and the NOE restraints. By using these files as an input, the structure calculation is performed using CYANA (*22*). In addition to CYANA, other software packages such as XPLOR-NIH (*23, 24*) and ARIA (*25*) can also be used to calculate the structure. The final step of the structure calculation is the water refinement of the structure to lower the target function. The details of the structure calculation procedures will be discussed in Chapter 4. This can be performed by using software such as CNS, ARIA, YASRA, and XPLOR-NIH. The NMR structure contains an ensemble of structures having a low root mean square deviation (RMSD) among the defined set of structures. A good structure should contain a low energy function, and a few or no angle and distance restraint violations. The structure can be validated by using the program PROCHECK (*26*).

1.1.7 Solid-state NMR spectroscopy

In solution-state NMR spectroscopy, the molecules in solution are in rapid random motion. As a result, the environment around a molecule in the solution is isotropic. This isotropy is achieved only in the solution state due to the averaging of the anisotropic NMR interactions as the molecules tumble rapidly in the solution. Due to this isotropy, the NMR signals in solution NMR are sharp. In contrast to the solution sample, the solid samples have an anisotropic environment. Therefore, the solid-state NMR signals are a lot broader compared to those in solution-state NMR. The resolution of NMR spectra in solid-state NMR can be improved by applying specific techniques such as cross-polarization, magic angle spinning, enhanced probe electronics, and special 2D and 3D NMR experiments.

The introduction of magic angle spinning (MAS) in solid-state NMR (ssNMR) has helped to make advancement in the elucidation of the molecular structure of complex macromolecules including membrane proteins. MAS was used in ssNMR to remove the anisotropic dipolar interactions in the solid samples. The anisotropy can be suppressed by rotating the sample at an angle of 54.74^0 with respect to the applied external magnetic field. This is termed magic angle spinning (MAS). MAS in ssNMR mimics the molecular motion phenomenon observed by the solution NMR. The fast-isotropic tumbling behavior of solutions can average the dipolar and chemical shielding component ($3\cos^2\Theta - 1$) to zero. But in ssNMR, this $3\cos^2\Theta - 1$ term can be made zero by spinning the sample at an angle of 54.74^0 with respect to the applied magnetic field. For this, the solid powder samples are packed in a small rotor and then spun at a high spinning rate to get high-resolution spectra. The sensitivity of the ssNMR can be enhanced by using another technique called cross-polarization (CP). In CP, the polarization is transferred from the abundant spins such as ^1H to less abundant spins such as ^{13}C and ^{15}N.

For the determination of the 3D structure of a membrane protein by using ssNMR, different MAS ssNMR experiments are utilized. (i) 2D dipolar assisted rotational resonance (DARR), (ii) 3D N-C-C correlation experiments such as NCACX, NCOCX, and CAN(CO)CX, (iii) hCHHC, and (iv) hNCHHC.

(i) 2D dipolar assisted rotational resonance (DARR): This experiment is a 2D ssNMR experiment that is based on ^{13}C-^{13}C recoupling, mechanical rotation of the sample, and ^1H irradiation (*27, 28*).

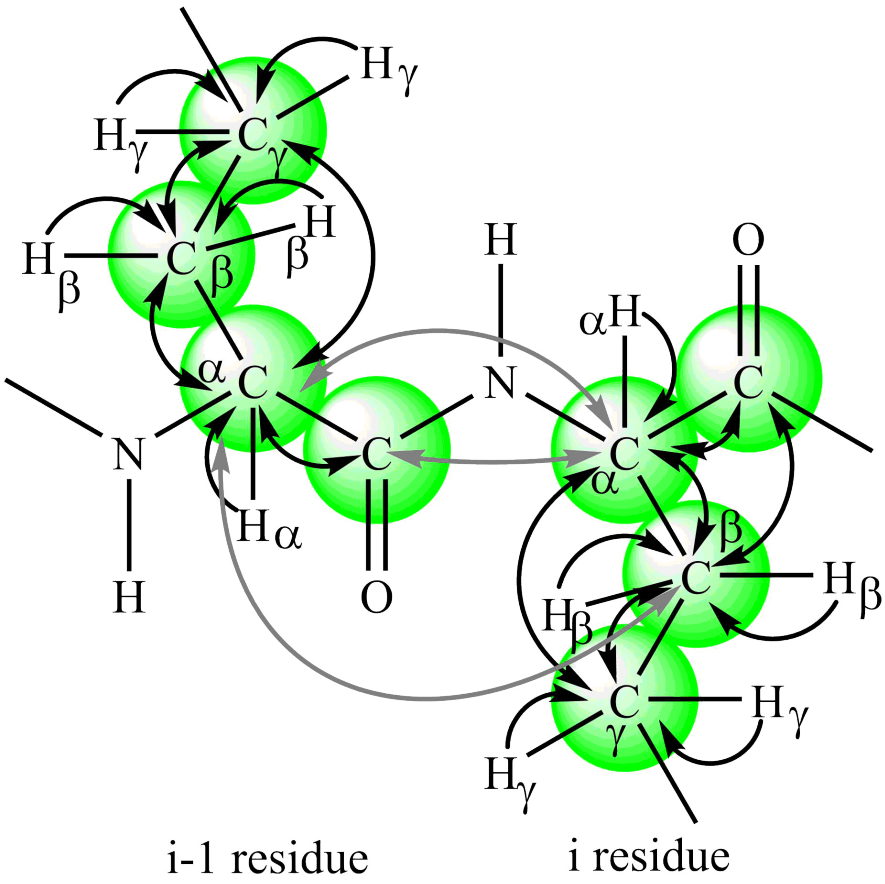


Figure 1.11: Magnetization transfer process and chemical shift evolvement in a CC2D DARR experiment. Chemical shift is observed on the ^{13}C nuclei that are surrounded by green circles. The black arrowheads indicate the magnetization transfer when the mixing time is 10 -50 ms and the grey arrows represent the magnetization transfer when the mixing time is 200-500 ms.

Here, the magnetization is transferred from ^{1}H to the ^{13}C nuclei connected to it and then to the ^{13}C nuclei that are close by in space (Figure 1.11). This experiment when acquired with a mixing time of 10 to 50 ms provides intra-residue contacts. If recorded with a mixing

time of 200-500 ms, this can provide inter-residue contacts for assignment and structure calculation.

(ii) 3D N-C-C correlation experiments such as NCACX, NCOCX, and CAN(CO)CX:

These 3D experiments are acquired in order to walk through the protein backbone in ssNMR. These experiments are based on dipolar coupling dependent polarization transfer between N, CO, Cα, Cβ, and side-chain carbon atoms to provide information for sequence-specific backbone assignment. Each of these experiments is described below.

1.1.7.1 NCACX

This ssNMR spectrum is very helpful to determine the intra-residue correlations. In this experiment, the magnetization is transferred from ^{1}H to ^{15}N by cross-polarization and then to Cα and finally, from Cα, the magnetization is then transferred to any carbon nuclei nearby in space via the DARR step (*29*). The spectra generate intra-residue cross-peaks when acquired with shorter (10-50 ms) mixing times and can give rise to inter-residue cross-peaks if collected with longer (200-500 ms) mixing times (Figure 1.12).

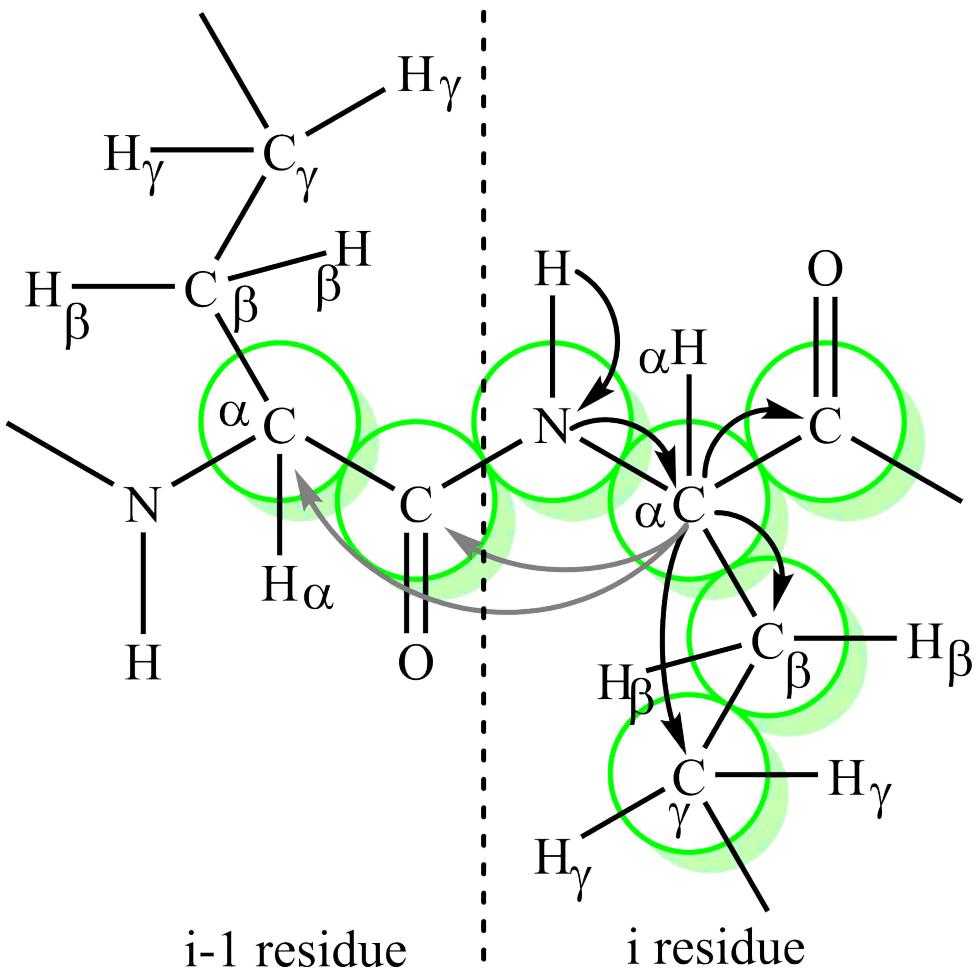


Figure 1.12: Magnetization transfer process and chemical shift evolvement in a 3D NCACX experiment. Chemical shift is observed on the ^{13}C nuclei that are surrounded by green circles. The black arrowheads indicate the magnetization transfer when the mixing time is 10 -50 ms and the grey arrows represent the magnetization transfer when the mixing time is 200-500 ms.

1.1.7.2 NCOCX

This experiment is also used to determine the intra-residue correlations for the sequence-specific backbone assignment of protein residues. Similar to that in the NCACX

experiment, the magnetization, in NCOCX (*29*), is transferred from ^{1}H to ^{15}N of i residue via cross-polarization at the first step. But at the second step, the magnetization is then transferred to carbonyl (CO) of the i-1 residue via a specific CP step and finally to the carbon nuclei of i-1 via the DARR step. At short mixing (10-50 ms) time the chemical shift evolves on the N nucleus of its residue and CO, Cα, Cβ, Cγ etc. nuclei of the preceding residue (Figure 1.13). At higher mixing time, the chemical shift can evolve on carbon nuclei of its residue as well.

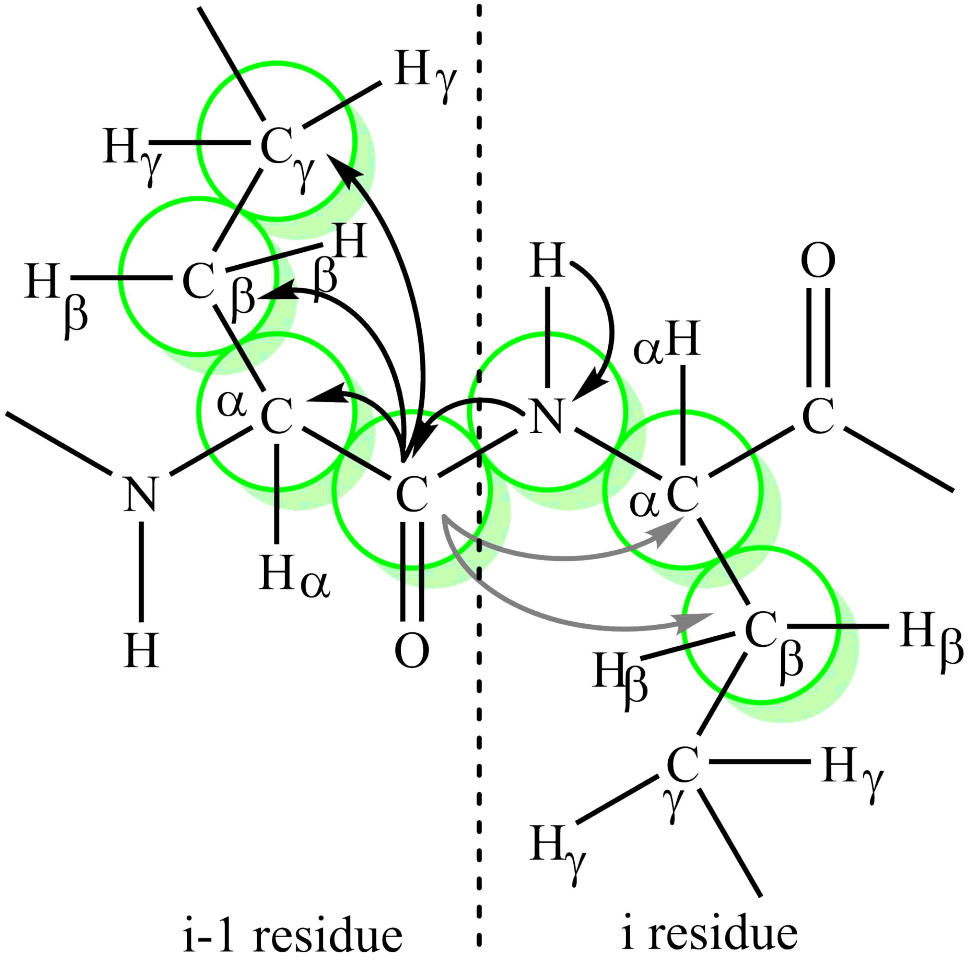


Figure 1.13: Magnetization transfer process and chemical shift evolvement in a 3D NCOCX experiment. Chemical shift is observed on the ^{13}C nuclei that are surrounded by

green circles. The black arrowheads indicate the magnetization transfer when the mixing time is 10 -50 ms and the grey arrows represent the magnetization transfer when the mixing time is 200-500 ms.

1.1.7.3 CANCOX

This is also a very important ssNMR experiment that is used for the sequence-specific backbone assignment of a protein. In this experiment, the magnetization is transferred from Hα to CA via cross-polarization and then to the ^{15}N nucleus from where the magnetization is transferred to the CO of i-1 residue via selective cross-polarization. At the final step, the magnetization is then transferred to the ^{13}C nuclei nearby via proton-driven spin diffusion (PDSD) (*30*). This experiment establishes the connection between two sequential residues. In this experiment, the chemical shift evolves on CA and ^{15}N of i residue and CO and sidechain ^{13}C of i-1 residue (Figure 1.14).

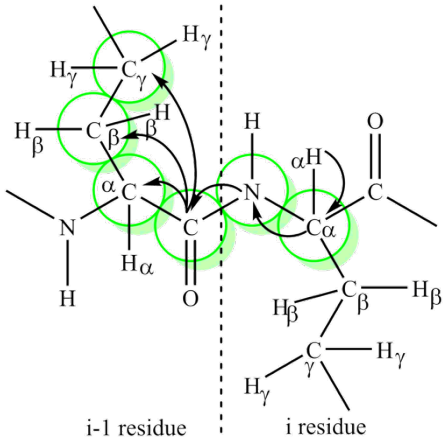


Figure 1.14: Magnetization transfer process and chemical shift evolvement in a CANCOCX experiment. Chemical shift is observed on the ^{13}C nuclei that are surrounded by green circles. The black arrowheads indicate the magnetization transfer.

1.1.7.4 2D CHHC

This experiment is a 2D carbon-carbon experiment that gives the chemical shifts of ^{13}C nuclei but encodes the information about the inter-proton distances attached to the carbon atoms. This experiment is used to derive the proton-proton distance restraints required for the protein structure calculation. In this experiment, the magnetization is transferred from the ^{1}H nucleus to ^{13}C by the cross-polarization process (Figure 1.15). After that, the magnetization is transferred back to the initial proton in three successive steps and then to another ^{1}H nucleus present close to it. Finally, the magnetization is transferred back to the ^{13}C nuclei for detection (*31*). Therefore, the chemical shift evolves on the ^{13}C nucleus that is attached to the initial proton and then detected on the nearby carbon.

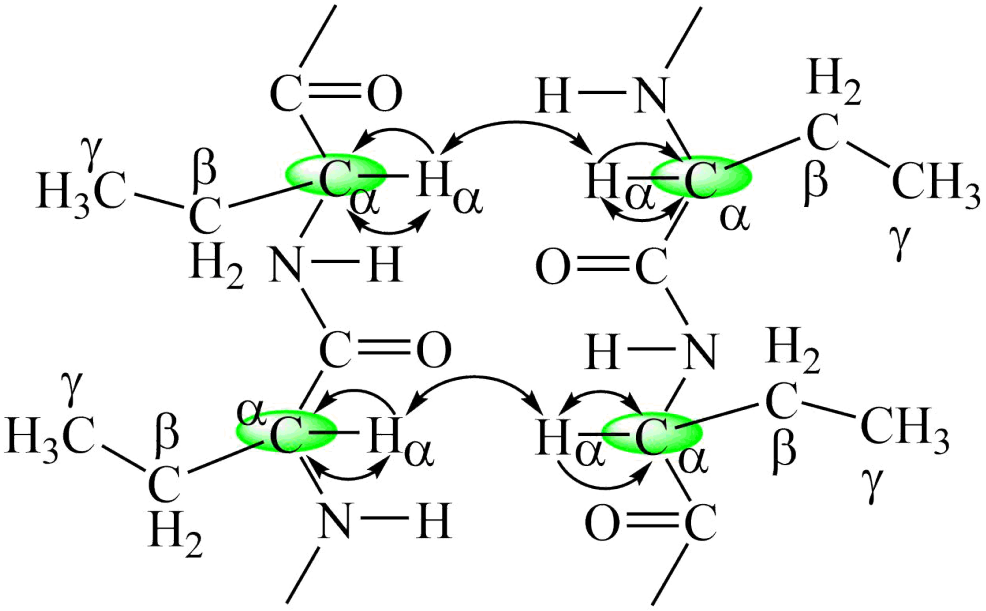


Figure 1.15: Magnetization transfer process and chemical shift evolvement in a 2D ChhC experiment. Chemical shift is observed on the ^{13}C nuclei that are surrounded by green circles. The black arrowheads indicate the magnetization transfer.

1.1.7.5 2D NHHC

This is a 2D experiment having carbon and nitrogen dimensions. In this experiment, the chemical shifts are detected for ^{13}C and ^{15}N nuclei but encode for proton-proton distances attached to them. This experiment also provides inter-proton distance restraints necessary for protein structure calculation. In this experiment, the magnetization is transferred from ^{1}H nuclei to ^{15}N via cross-polarization and then in three successive steps, the magnetization is transferred back to the ^{1}H and then to the close by ^{1}H nuclei and finally to the ^{13}C nuclei for detection (Figure 1.16) (*31*).

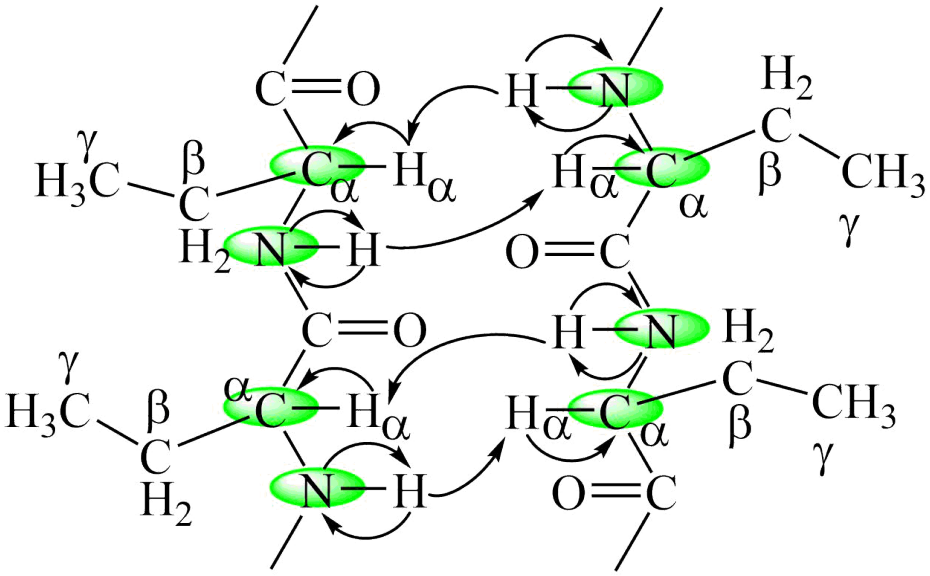


Figure 1.16: Magnetization transfer process and chemical shift evolvement in a 2D NhhC experiment. Chemical shift is observed on the ^{13}C nuclei that are surrounded by green circles. The black arrowheads indicate the magnetization transfer.

1.2 Oligosaccharyltransferase

Oligosaccharyltransferase (OST) is a membrane-associated enzyme complex that catalyzes an essential and highly conserved asparagine-linked glycosylation reaction. This reaction

is also known as *N*-linked glycosylation that occurs in the lumen of the endoplasmic reticulum (ER). It occurs in most eukaryotes and some prokaryotes (*32-35*). In contrast to eukaryotes, *N*-linked glycosylation in prokaryotes is nonessential but assists for survival and pathogenicity. The OST enzyme is monomeric in bacteria, archaea, and protozoa but is multimeric in higher eukaryotes such as animals, plants, and fungi (*36*). OST facilitates the transfer of an oligosaccharide donor substrate that is composed of three glucose (Glc), nine mannose (Man) and two *N*-acetylglucosamine (GlcNAc) monomers, ($Glc_3Man_9GlcNAc_2$). This sugar molecule is attached to the ER membrane through a lipid which is generally a dolichol pyrophosphate (DolPP). The sugar molecule and the DolPP are together known as lipid-linked oligosaccharide (LLO). The sugar moiety of the LLO donor is transferred to the side chain of an asparagine residue of a consensus sequence -N-X-T/S- (where X ≠ proline) nascent polypeptide (*37, 38*). This modification is necessary for proper protein folding and other biological processes such as molecular recognition, stability, cell-cell communication, and subcellular targeting (*36, 39, 40*). A general mechanism of *N*-linked glycosylation in higher eukaryotes is described in Figure 1.17. Genetic defects in *N*-linked glycosylation leads to a class of diseases collectively known as congenital disorders of glycosylation (CDG) with clinical symptoms that include but are not limited to developmental delay, mental retardation, liver dysfunction anorexia, gastrointestinal disorders, and dysmorphic features (*42, 43*). *N*-linked glycans are necessary for the proper folding of the nascent polypeptide (*44-46*). Any changes to the *N*-glycan structure on proteins can flag them for proteasome degradation by the ER quality control system (*40*).

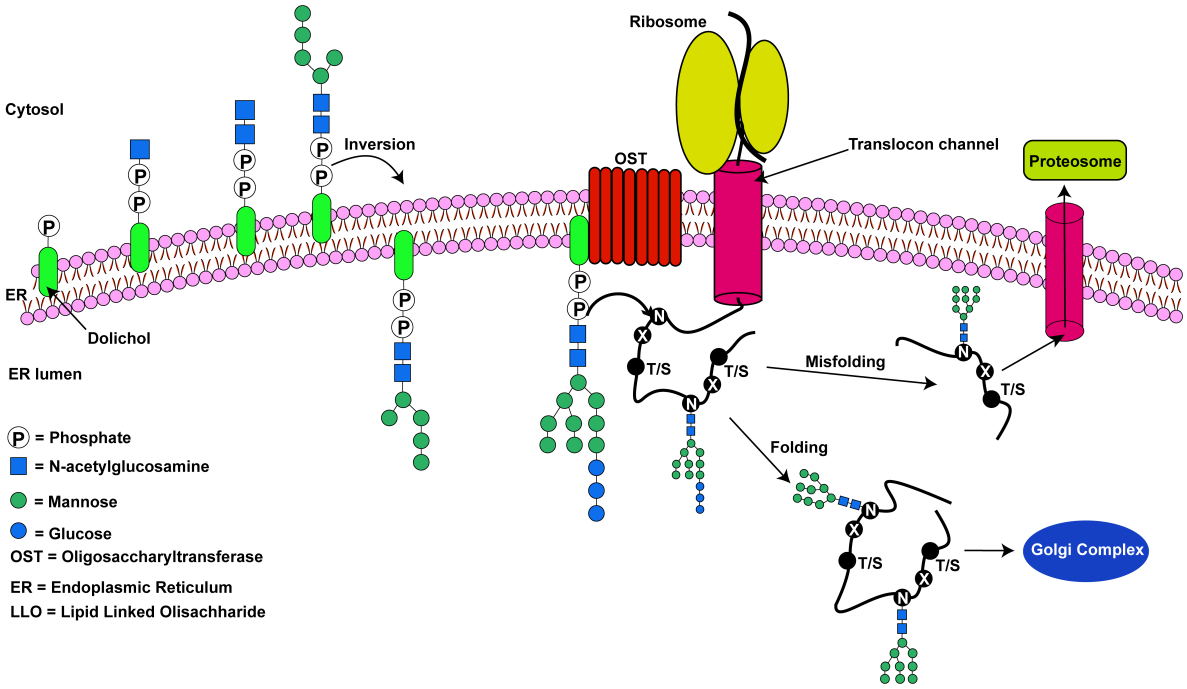


Figure 1.17: A general mechanism of the *N*-linked glycosylation of proteins in higher eukaryotes: pyrophosphate and monosaccharides are added to the dolichol lipid on the cytosolic side of the endoplasmic reticulum. The lipid-linked oligosaccharide (LLO) is inverted to the luminal side of the endoplasmic reticulum (ER). Additional monosaccharides are added to form the mature LLO. Oligosaccharyltransferase (OST) catalyzes the transfer of the oligosaccharide from the LLO to the side-chain of an asparagine residue in -N-X-T/S- consensus sequence within a protein. Protein folding occurs after *N*-linked glycosylation. The three terminal glucose residues are trimmed before translocating to the Golgi apparatus for sorting. Misfolded proteins are targeted for degradation by proteasomes. The figure is adapted from reference (*41*).

Unlike prokaryotic OST enzyme that contains a membrane-embedded single subunit: archaeal glycosylation B (AglB) for archaea, and protein glycosylation B (PglB) for bacteria, eukaryotic OST is composed of multiple non-identical protein subunits (*37*) (Figure 1.18).

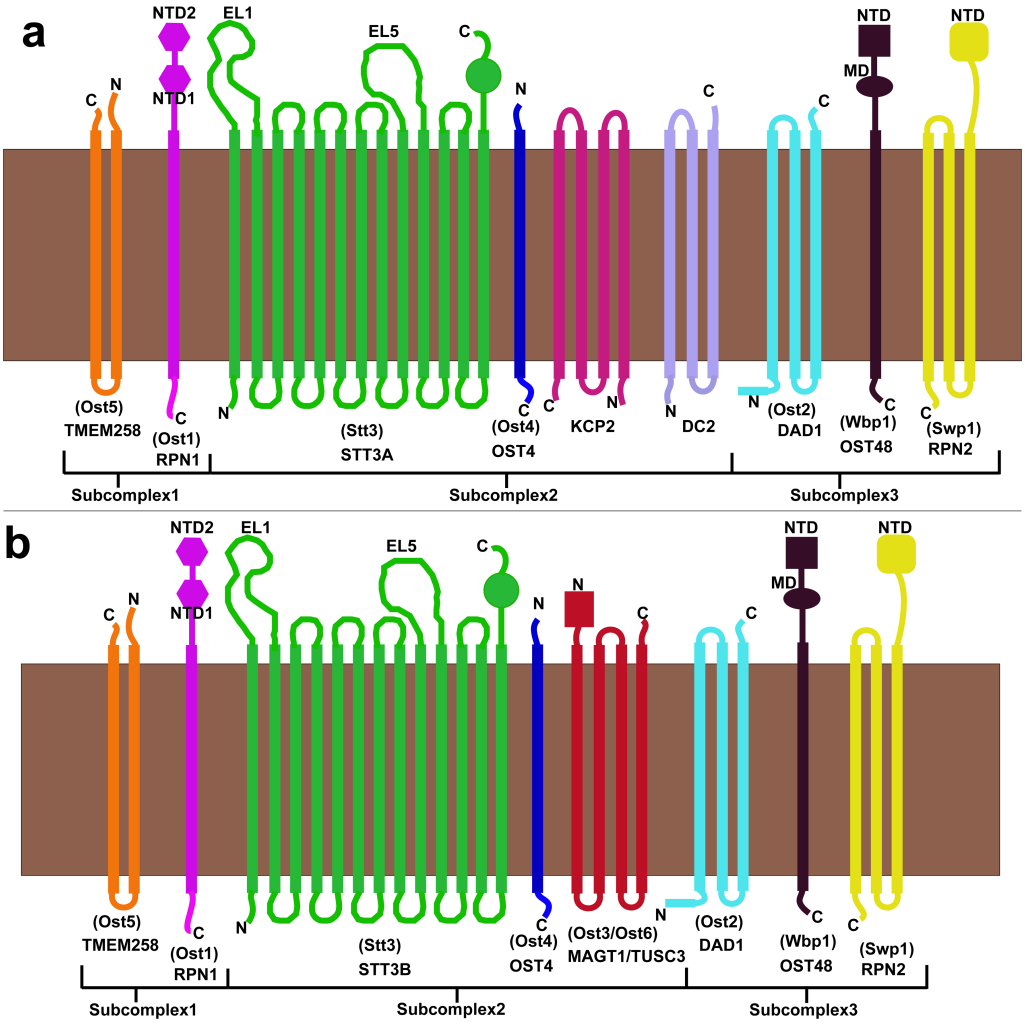


Figure 1.18: Subunit organization of the human and yeast OST complex in ER membrane. (a), OST-A complex. (b), OST-B complex. Subunits are labeled by mammalian names with yeast subunit names shown in parentheses. Mammalian OST-A complex is homologous to the yeast OST complex, while the yeast OST lacks KCP2 and DC2 subunits found exclusively in the OST-B complex. The figure is adapted from the reference (*41*).

The OST complex in humans and other mammals is divided into two distinct isoforms known as OST-A and OST-B. Each isoform performs completely different roles in the *N*-linked glycosylation of proteins. OST-A is attached directly to the Sec61 of the translocation channel in the ER membrane and glycosylates the newly synthesized unfolded polypeptide chain emerging from the ribosome co-translationally (*47, 48*). The OST-A glycosylates the majority of glycoproteins in mammals (*49*). In contrast to OST-A, OST-B plays a proofreading role to glycosylate the sites, in partially folded proteins or proteins that contain disulfide bonds, that are missed by OST-A. OST-A is very similar from the OST-B complex. However, they are dissimilar from each other to some extent. Both complexes share the following subunits: OST 4-kDa subunit (OST4), ribophorin 1 (RPN1), ribophorin 2 (RPN2), defender against cell death 1 (DAD1), transmembrane protein 258 (TMEM258), and OST 48-kDa subunit (OST48) (*50*). The only difference between these two complexes is that OST-A contains keratinocyte-associated protein 2 (KCP2) and/or DC2 in place of OST3/OST6 from yeast and the catalytic subunit is STT3A and OST-B and contains either tumor suppressor candidate 3 (TUSC3) or MAGT1 subunits as redox capable yeast OST3/OST6 homologs, respectively (*48, 51, 52*). KCP2 and DC2 subunits were reported to have a connection to the ribosome via the translocation channel Sec61 (*47, 48*). However, the structure of OST-A shows that it did not contain the KCP2 subunit as a part of subcomplex II, but it was found to bind an additional protein called malectin (*53*). The malectin takes part in the quality control of glycoproteins in the ER (*54*). The unfolded glycoproteins were found to increase the association between malectin and ribophorin I (*55*). This indicates that the OST-B complex may encounter more unfolded glycoproteins than OST-A.

Yeast, *Saccharomyces cerevisiae*, contains two functional isoforms that share seven subunits: Ost1, Ost2, Ost4, Ost5, Stt3, Swp1, and Wbp1. Each isoform consists of either an Ost3 or Ost6 subunit. Therefore, Ost3 and Ost6 subunits are homologues. In recent structural studies, genetic, and biochemical characterizations have shown these subunits to be into three subcomplexes: subcomplex I (Ost1-Ost5), subcomplex II (Wbp1-Swp1-Ost2), and subcomplex III (Ost4-Stt3-Ost3/Ost6) (*56-63*). Recent advancements in atomic resolution structural techniques such as nuclear magnetic resonance (NMR) spectroscopy, crystallography, and cryo-electron microscopy have provided high-resolution structures of the individual subunits or the whole OST complex. Crystal structures of the luminal domain of Ost6 (*64, 65*), NMR structures of Ost4 (*66, 67*) and Stt3 (*68*), and low resolution cryo-EM structures of mammalian and yeast OST complex (*69, 70*) have contributed to the understanding of the OST enzyme complex and the overall mechanism of *N*-linked glycosylation reaction (Table 1.1). Recent high-resolution cryo-EM structures of the yeast OST complex (*62, 63*) and both of the human OST complexes (*53*) have transformed our understanding of this enzyme. Table 1.1 provides the structures of OSTs and their subunits determined so far by using various methods.

Table 1.1: List of all the high-resolution structures determined by various methods thus far with their protein data bank (PDB) code. The table is adapted from reference (*41*).

Prokaryotic Oligosaccharyltransferase			
Bacterial Oligosaccharyltransferase			**References**
Protein	**Method**	**PDB ID**	
Campylobacter lari PglB with acceptor peptide	X-ray	3RCE	(*71*)
Campylobacter lari PglB with acceptor peptide and LLO analog	X-ray	5OGL	(*72*)

Campylobacter lari PglB with inhibitory peptide and reactive LLO	X-ray	6GXC	(*73*)
C-terminal domain of *Campylobacter jejuni* PglB	X-ray	3AAG	(*74*)
Archaeal Oligosaccharyltransferase			
The C-terminal soluble domain from *Archaeoglobus fulgidus*	X-ray	3VU0	(*75*)
The C-terminal soluble domain from *Pyrococcus horikoshii*	X-ray	3VU1	(*75*)
C-term globular domain as MBP fusion from *Archaeoglobus fulgidus*	X-ray	3WAI	(*76*)
Archaeoglobus fulgidus AglB	X-ray	3WAK	(*77*)
Archaeoglobus fulgidus AglB with Zn and sulfate	X-ray	3WAJ	(*77*)
Archaeoglobus fulgidus AglB with acceptor peptide	X-ray	5GMY	(*78*)
Eukaryotic Oligosaccharyltransferase			
(1) Yeast Oligosaccharyltransferase (OST)			
Yeast OST subunit Ost4p	Solution NMR	1RKL	(*67*)
Oxidized Ost6L	X-ray	3G7Y	(*65*)
Reduced Ost6L	X-ray	3G9B	(*65*)
Photo-reduced Ost6L	X-ray	3GA4	(*65*)
C-terminal domain of Stt3p subunit	Solution NMR	2LGZ	(*68*)
OST complex	Cryo-EM	6EZN	(*63*)
OST complex	Cryo-EM	6C26	(*62*)
(2) Human Oligosaccharyltransferase			
Soluble *N*-terminal domain of N33/Tusc3 subunit	X-ray	4M90, 4M91, 4M92, and 4M8G	(*64*)
Ost4 subunit	Solution NMR	2LAT	(*66*)
OST-A complex	Cryo-EM	6S7O	(*53*)
OST-B complex	Cryo-EM	6S7T	(*53*)

Despite a number of research reports on the role of subunits of the yeast OST enzyme, the exact functions of these subunits are still not clear. The subcomplexes in yeast are described below.

Subcomplex I is composed of two subunits, Ost1 and Ost5. Ost1 contains two similar *N*-terminal luminal domains and contains primarily β-sheets (*62, 63*). Ost1 binds only glycosylated peptides, which suggests that it may prevent the sliding of the newly glycosylated peptide back into the catalytic site (*62, 79*). And Ost5 of this subcomplex has been proposed to support Ost1 (*62*).

Subcomplex II is formed of subunits Ost2, Swp1, and Wbp1. Swp1 and Wbp1 subunits contain one and two large luminal *N*-terminal domains respectively. The Ost2 subunit mediates the contacts between Stt3 and transmembrane helices of Wbp1 and Swp1. Swp1 and Wbp1 most likely are involved in recruiting LLO or act as a docking platform for the recruitment of other accessory proteins acting on nascent glycoproteins (*62, 63*). Although a plethora of reports is published on the functions of these proteins, their role in substrate binding and catalysis is still unclear (*80-82*). Initially, Swp1 and Wbp1 along with Ost1 were proposed to act as chaperones for protein folding and glycosylation (*80, 83*). The recent structures of the OST complex have shown them not to have chaperone-like folds (*62, 63*).

Subcomplex III is composed of the smallest subunit Ost4, the catalytic subunit Stt3, and either Ost3 or Ost6. The NMR structure of the yeast Ost4 subunit in mixed aqueous-organic solvent shows a well-formed kinked helix (*67*). Mutation of any residue at positions 18-24 to a charged residue in Ost4 results in severe growth defects in yeast (*84*). These mutations destabilize the Stt3-Ost4-Ost3 sub-complex (*84, 85*). Ost3, a subunit homologous to Ost6, consists of four transmembrane helices and Ost3 acts as an LLO docking site (*62*).

Ost4 is reported to play a critical role in the incorporation of Ost3 or Ost6 in the OST complex (*62, 63, 86*). Ost4 is localized between Stt3 and Ost3 in the Stt3-Ost4-Ost3 sub-complex, acting as a bridge stabilizing this complex (*85*). Ost4p is a mini-membrane protein having only one transmembrane domain containing 36 amino acid residues (*85*). Sequence alignment has shown that most yeast OST subunits have high sequence similarity with OST subunits identified in higher eukaryotes, while Ost4 has the highest sequence similarity (Fig. 1.19) in *C. elegans*, *M. musculus*, and *H. sapiens* (*57, 67*). Mutagenesis studies have shown that substitution of any residue in positions 18-24 in Ost4 with a charged residue results in temperature sensitivity, impaired cell growth and disrupts *in vitro* OST activity (*84, 85*).

```
Yeast                MISDEQLNSLAITFGIVMMTLIVIYHAVDSTMSPKN----36
Human, House mouse   MITDVQLAIFANMLGVSLFLLVVLYHYVAVNNPKKQE---37
Fission yeast        M-TDVQLQNIVTTFGISMMLLIILYHYLSRPQA-------32
Nematode             MISDVQLGIAANILGIAMLMLVVLFHYLNANQKNK-----35
Clawed frog          MISDVQLAIFANMLGVSLFL-VVLYHYVSVNNPKKLD---36
Thale cress          MIDDQDLGFIANFLGIFIFALVIAYHYVTADPKYEAT---37
Fruit fly            MITDVQLAIFSNVLGVFLFLLVVAYHYINANTGKPSAKAK40
```

Figure 1.19: Sequence alignment of Ost4p from the *S. cerevisiae* (yeast) and analogs of Ost4 from other species: *Homo sapiens* (human), *Mus musculus* (house mouse), *Schizosaccharomyces pombe* (fission yeast), *C. elegans* (nematode), *Xenopus laevis* (clawed frog), *Arabidopsis thaliana* (thale cress), and *Drosophila melanogaster* (fruit fly). The residues highlighted in various colors are identical/similar across different species.

The three-dimensional structure of chemically synthesized Ost4 was determined in our laboratory in mixed aqueous-organic solvents (*67*). Ost4 is folded into a well-formed kinked helix in this system. This structure explained the results of mutational studies.

Mutation of any residue present in positions Met^{18} to Ile^{24} in helix-2 to a charged residue in Ost4 resulted in severe growth defects in yeast affecting OST activity (*84*). These mutations were reported to cause destabilization of the Stt3-Ost4-Ost3 sub-complex (*84, 85*). In contrast, mutation of any residue from Ile^{2} to Val^{17} did not affect OST activity or the stability of the complex (*84*). In the Stt3-Ost4-Ost3 sub-complex, Ost4 was reported to interact with Ost3 through residues M19, T20, I22 and V23, and with Stt3p through residues M18, L21 and I24 (*67, 84*). Thus, mutation of any of these hydrophobic residues to charged residue disrupted the interactions of Ost4 to either Ost3 or Stt3 in the complex (*84, 85*). Mutation of Met^{18} to Lys or Val^{23} to Asp resulted in a severe cell growth defect but mutation of Met^{18} to Leu or Val^{23} to Gly did not affect cell growth suggesting the importance of hydrophobic residues in these positions for maintaining the stability of the OST complex (*84*). Based on analysis of the NMR structure of yeast Ost4, it was suggested that the α2 helix of Ost4 interacts with Stt3 on one side and Ost3 on the other side via a i+4 "ridges-into grooves" helix packing mechanism (*67*). The NMR structure of human Ost4 determined in mixed aqueous-organic solvent shows a similar kinked helix as in yeast Ost4p (*66, 67*). Thus, point mutation of any residue in the α2 helix might disrupt the "ridges-into-grooves" fit resulting in the disruption of the Stt3-Ost4-Ost3 sub-complex.

Although the high-resolution NMR structures of chemically synthesized Ost4 from yeast and human are available, the effect of functionally important mutations on the 3-D structure of this protein remains unanswered. For example, what would happen to the "ridges into groove" interaction between Ost4 and Stt3/Ost3 upon mutation of any residue in the α2 helix? What is the impact of point mutation on the overall 3-D structure of Ost4? The

structure-function studies of Ost4 and its functionally important mutants would provide insight into the protein-protein interactions involved in stabilization of the Stt3-Ost4-Ost3 sub-complex and consequently the OST complex.

In this dissertation, the NMR structures of the Ost4 subunit and its critical point mutant Ost4V23D are to be determined by solution NMR in dodecyl phosphocholine (DPC) micelles and by ssNMR in the lipid bilayer. Additionally, the structures of Ost4 and its critical mutant will be compared in different membrane mimicking systems. The comparison of Ost4 structures determined in various membrane mimicking systems will be reported.

1.3 References

1. A. U. Rahman, *Nuclear Magnetic Resonance: Basic Principles, Springer-Verlag*. (New York, 1986).
2. J. P. Korb, R. G. Bryant, Magnetic field dependence of proton spin-lattice relaxation times. *Magn Reson Med* **48**, 21-26 (2002).
3. D. D. Traficante, Relaxation. Can T2, be longer than T1? *Concepts in Magnetic Resonance* **3**, 171-177 (1991).
4. N. S. John Cavanagh, Wayne Fairbrother, Mark Rance, Arthur Palmer, III, *Protein NMR spectroscopy: principles and practice*. (Academic Press, ed. 2nd, 2006).
5. J. Jeener. (Ampere International Summer School, Basko Polje, Jugoslavia, 1971).

6. W. P. Aue, E. Bartholdi, R. R. Ernst, Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *The Journal of Chemical Physics* **64**, 2229-2246 (1976).

7. A. W. Overhauser, Polarization of Nuclei in Metals. *Physical Review* **92**, 411-415 (1953).

8. K. Murata, M. Wolf, Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. *Biochim Biophys Acta Gen Subj* **1862**, 324-334 (2018).

9. C. Li, M. Liu, Protein dynamics in living cells studied by in-cell NMR spectroscopy. *FEBS Letters* **587**, 1008-1011 (2013).

10. C. Li, C. Tang, M. Liu, Protein dynamics elucidated by NMR technique. *Protein Cell* **4**, 726-730 (2013).

11. M. Osawa, K. Takeuchi, T. Ueda, N. Nishida, I. Shimada, Functional dynamics of proteins revealed by solution NMR. *Curr Opin Struct Biol* **22**, 660-669 (2012).

12. S. Grzesiek, A. Bax, An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *Journal of Magnetic Resonance (1969)* **99**, 201-207 (1992).

13. B. T. Farmer, R. A. Venters, L. D. Spicer, M. G. Wittekind, L. Müller, A refocused and optimized HNCA: Increased sensitivity and resolution in large macromolecules. *Journal of Biomolecular NMR* **2**, 195-202 (1992).

14. S. Grzesiek, A. Bax, Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *Journal of Magnetic Resonance (1969)* **96**, 432-440 (1992).

15. L. E. Kay, M. Ikura, R. Tschudin, A. Bax, Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *Journal of Magnetic Resonance (1969)* **89**, 496-514 (1990).

16. D. R. Muhandiram, L. E. Kay, Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity. *Journal of Magnetic Resonance, Series B* **103**, 203-216 (1994).

17. D. Marion *et al.*, Overcoming the overlap problem in the assignment of proton NMR spectra of larger proteins by use of three-dimensional heteronuclear proton-nitrogen-15 Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1.beta. *Biochemistry* **28**, 6150-6156 (1989).

18. S. Grzesiek, J. Anglister, A. Bax, Correlation of Backbone Amide and Aliphatic Side-Chain Resonances in 13C/15N-Enriched Proteins by Isotropic Mixing of 13C Magnetization. *Journal of Magnetic Resonance, Series B* **101**, 114-119 (1993).

19. A. Bax, G. M. Clore, A. M. Gronenborn, 1H□1H correlation via isotropic mixing of 13C magnetization, a new three-dimensional approach for assigning 1H and 13C spectra of 13C-enriched proteins. *Journal of Magnetic Resonance (1969)* **88**, 425-431 (1990).

20. D. Marion, L. E. Kay, S. W. Sparks, D. A. Torchia, A. Bax, Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins. *Journal of the American Chemical Society* **111**, 1515-1517 (1989).

21. E. R. P. Zuiderweg, S. W. Fesik, Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. *Biochemistry* **28**, 2387-2391 (1989).

22. P. Güntert, L. Buchner, Combined automated NOE assignment and structure calculation with CYANA. *Journal of Biomolecular NMR* **62**, 453-471 (2015).

23. C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* **160**, 65-73 (2003).

24. C. D. Schwieters, J. J. Kuszewski, G. Marius Clore, Using Xplor–NIH for NMR molecular structure determination. *Progress in Nuclear Magnetic Resonance Spectroscopy* **48**, 47-62 (2006).

25. J. P. Linge, M. Habeck, W. Rieping, M. Nilges, ARIA: automated NOE assignment and NMR structure calculation. *Bioinformatics* **19**, 315-316 (2003).

26. R. A. Laskowski, J. A. Rullmannn, M. W. MacArthur, R. Kaptein, J. M. Thornton, AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**, 477-486 (1996).

27. K. Takegoshi, S. Nakamura, T. Terao, 13C–1H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chemical Physics Letters* **344**, 631-637 (2001).

28. K. Takegoshi, S. Nakamura, T. Terao, 13C–1H dipolar-driven 13C–13C recoupling without 13C rf irradiation in nuclear magnetic resonance of rotating solids. *The Journal of Chemical Physics* **118**, 2325-2341 (2003).

29. J. Pauli, M. Baldus, B. van Rossum, H. de Groot, H. Oschkinat, Backbone and Side-Chain 13C and 15N Signal Assignments of the α-Spectrin SH3 Domain by

Magic Angle Spinning Solid-State NMR at 17.6 Tesla. *ChemBioChem* **2**, 272-281 (2001).

30. W. T. Franks, K. D. Kloepper, B. J. Wylie, C. M. Rienstra, Four-dimensional heteronuclear correlation experiments for chemical shift assignment of solid proteins. *Journal of Biomolecular NMR* **39**, 107-131 (2007).

31. A. Lange, S. Luca, M. Baldus, Structural Constraints from Proton-Mediated Rare-Spin Correlation Spectroscopy in Rotating Solids. *Journal of the American Chemical Society* **124**, 9704-9705 (2002).

32. R. E. Dempski, Jr., B. Imperiali, Oligosaccharyl transferase: gatekeeper to the secretory pathway. *Curr Opin Chem Biol* **6**, 844-850 (2002).

33. R. Kornfeld, S. Kornfeld, Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem* **54**, 631-664 (1985).

34. A. Larkin, B. Imperiali, The expanding horizons of asparagine-linked glycosylation. *Biochemistry* **50**, 4411-4426 (2011).

35. M. Wacker *et al.*, N-linked glycosylation in Campylobacter jejuni and its functional transfer into E. coli. *Science* **298**, 1790-1793 (2002).

36. E. Mohorko, R. Glockshuber, M. Aebi, Oligosaccharyltransferase: the central enzyme of N-linked protein glycosylation. *J Inherit Metab Dis* **34**, 869-878 (2011).

37. R. Knauer, L. Lehle, The oligosaccharyltransferase complex from yeast. *Biochim Biophys Acta* **1426**, 259-273 (1999).

38. J. K. Welply, P. Shenbagamurthi, W. J. Lennarz, F. Naider, Substrate recognition by oligosaccharyltransferase. Studies on glycosylation of modified Asn-X-Thr/Ser tripeptides. *J Biol Chem* **258**, 11856-11863 (1983).

39. C. G. Gahmberg, M. Tolvanen, Why mammalian cell surface proteins are glycoproteins. *Trends Biochem Sci* **21**, 308-311 (1996).

40. A. Helenius, M. Aebi, Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* **73**, 1019-1049 (2004).

41. S. Mohanty, B. P. Chaudhary, D. Zoetewey, Structural Insight into the Mechanism of N-Linked Glycosylation by Oligosaccharyltransferase. *Biomolecules* **10**, (2020).

42. H. H. Freeze, Understanding human glycosylation disorders: biochemistry leads the charge. *J Biol Chem* **288**, 6936-6945 (2013).

43. T. Hennet, J. Cabalzar, Congenital disorders of glycosylation: a concise chart of glycocalyx dysfunction. *Trends Biochem Sci* **40**, 377-384 (2015).

44. A. Helenius, How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. *Mol Biol Cell* **5**, 253-265 (1994).

45. J. C. Paulson, Glycoproteins: what are the sugar chains for? *Trends in Biochemical Sciences* **14**, 272-276 (1989).

46. M. R. Wormald, R. A. Dwek, Glycoproteins: glycan presentation and protein-fold stability. *Structure* **7**, R155-R160 (1999).

47. K. Braunger *et al.*, Structural basis for coupling protein transport and N-glycosylation at the mammalian endoplasmic reticulum. *Science* **360**, 215-219 (2018).

48. S. Shrimal, N. A. Cherepanova, R. Gilmore, DC2 and KCP2 mediate the interaction between the oligosaccharyltransferase and the ER translocon. *J Cell Biol* **216**, 3625-3638 (2017).

49. N. A. Cherepanova, S. V. Venev, J. D. Leszyk, S. A. Shaffer, R. Gilmore, Quantitative glycoproteomics reveals new classes of STT3A- and STT3B-dependent N-glycosylation sites. *J Cell Biol* **218**, 2782-2796 (2019).

50. D. J. Kelleher, D. Karaoglu, E. C. Mandon, R. Gilmore, Oligosaccharyltransferase isoforms that contain different catalytic STT3 subunits have distinct enzymatic properties. *Mol Cell* **12**, 101-111 (2003).

51. N. Cherepanova, S. Shrimal, R. Gilmore, N-linked glycosylation and homeostasis of the endoplasmic reticulum. *Curr Opin Cell Biol* **41**, 57-65 (2016).

52. C. Ruiz-Canada, D. J. Kelleher, R. Gilmore, Cotranslational and posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms. *Cell* **136**, 272-283 (2009).

53. A. S. Ramirez, J. Kowal, K. P. Locher, Cryo-electron microscopy structures of human oligosaccharyltransferase complexes OST-A and OST-B. *Science* **366**, 1372-1375 (2019).

54. K. Takeda, S.-Y. Qin, N. Matsumoto, K. Yamamoto, Association of malectin with ribophorin I is crucial for attenuation of misfolded glycoprotein secretion. *Biochemical and Biophysical Research Communications* **454**, 436-440 (2014).

55. S.-Y. Qin *et al.*, Malectin Forms a Complex with Ribophorin I for Enhanced Association with Misfolded Glycoproteins. *Journal of Biological Chemistry* **287**, 38080-38089 (2012).

56. D. Karaoglu, D. J. Kelleher, R. Gilmore, The highly conserved Stt3 protein is a subunit of the yeast oligosaccharyltransferase and forms a subcomplex with Ost3p and Ost4p. *J Biol Chem* **272**, 32513-32520 (1997).

57. D. J. Kelleher, R. Gilmore, An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* **16**, 47R-62R (2006).

58. S. Mueller *et al.*, Protein degradation corrects for imbalanced subunit stoichiometry in OST complex assembly. *Mol Biol Cell* **26**, 2596-2608 (2015).

59. U. Spirig *et al.*, The STT3 protein is a component of the yeast oligosaccharyltransferase complex. *Molecular and General Genetics MGG* **256**, 628-637 (1997).

60. S. Te Heesen, B. Janetzky, L. Lehle, M. Aebi, The yeast WBP1 is essential for oligosaccharyl transferase activity in vivo and in vitro. *EMBO Journal* **11**, 2071-2075 (1992).

61. S. te Heesen, R. Knauer, L. Lehle, M. Aebi, Yeast Wbp1p and Swp1p form a protein complex essential for oligosaccharyl transferase activity. *The EMBO journal* **12**, 279-284 (1993).

62. L. Bai, T. Wang, G. Zhao, A. Kovach, H. Li, The atomic structure of a eukaryotic oligosaccharyltransferase complex. *Nature* **555**, 328-333 (2018).

63. R. Wild *et al.*, Structure of the yeast oligosaccharyltransferase complex gives insight into eukaryotic N-glycosylation. *Science* **359**, 545-550 (2018).

64. E. Mohorko *et al.*, Structural basis of substrate specificity of human oligosaccharyl transferase subunit N33/Tusc3 and its role in regulating protein N-glycosylation. *Structure* **22**, 590-601 (2014).

65. B. L. Schulz *et al.*, Oxidoreductase activity of oligosaccharyltransferase subunits Ost3p and Ost6p defines site-specific glycosylation efficiency. *Proc Natl Acad Sci U S A* **106**, 11061-11066 (2009).

66. S. Gayen, C. Kang, Solution structure of a human minimembrane protein Ost4, a subunit of the oligosaccharyltransferase complex. *Biochem Biophys Res Commun* **409**, 572-576 (2011).

67. S. Zubkov, W. J. Lennarz, S. Mohanty, Structural basis for the function of a minimembrane protein subunit of yeast oligosaccharyltransferase. *Proc Natl Acad Sci U S A* **101**, 3821-3826 (2004).

68. C. Huang, R. Bhaskaran, S. Mohanty, Eukaryotic N-glycosylation occurs via the membrane-anchored C-terminal domain of the Stt3p subunit of oligosaccharyltransferase. *J Biol Chem* **287**, 32450-32458 (2012).

69. H. Li, M. Chavan, H. Schindelin, W. J. Lennarz, H. Li, Structure of the oligosaccharyl transferase complex at 12 A resolution. *Structure* **16**, 432-440 (2008).

70. S. Pfeffer *et al.*, Structure of the mammalian oligosaccharyl-transferase complex in the native ER protein translocon. *Nat Commun* **5**, 3072 (2014).

71. C. Lizak, S. Gerber, S. Numao, M. Aebi, K. P. Locher, X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **474**, 350-355 (2011).

72. M. Napiórkowska *et al.*, Molecular basis of lipid-linked oligosaccharide recognition and processing by bacterial oligosaccharyltransferase. *Nature Structural & Molecular Biology* **24**, 1100-1106 (2017).

73. M. Napiórkowska, J. Boilevin, T. Darbre, J.-L. Reymond, K. P. Locher, Structure of bacterial oligosaccharyltransferase PglB bound to a reactive LLO and an inhibitory peptide. *Scientific Reports* **8**, 16297 (2018).

74. N. Maita, J. Nyirenda, M. Igura, J. Kamishikiryo, D. Kohda, Comparative Structural Biology of Eubacterial and Archaeal Oligosaccharyltransferases. *Journal of Biological Chemistry* **285**, 4941-4950 (2010).

75. J. Nyirenda *et al.*, Crystallographic and NMR evidence for flexibility in oligosaccharyltransferases and its catalytic significance. *Structure* **21**, 32-41 (2013).

76. S. Matsumoto, A. Shimada, D. Kohda, Crystal structure of the C-terminal globular domain of the third paralog of the Archaeoglobus fulgidus oligosaccharyltransferases. *BMC Struct Biol* **13**, 11 (2013).

77. S. Matsumoto *et al.*, Crystal structures of an archaeal oligosaccharyltransferase provide insights into the catalytic cycle of N-linked protein glycosylation. *Proc Natl Acad Sci U S A* **110**, 17868-17873 (2013).

78. S. Matsumoto, Y. Taguchi, A. Shimada, M. Igura, D. Kohda, Tethering an N-Glycosylation Sequon-Containing Peptide Creates a Catalytically Competent Oligosaccharyltransferase Complex. *Biochemistry* **56**, 602-611 (2017).

79. Q. Yan, G. D. Prestwich, W. J. Lennarz, The Ost1p subunit of yeast oligosaccharyl transferase recognizes the peptide glycosylation site sequence, -Asn-X-Ser/Thr. *J Biol Chem* **274**, 5021-5025 (1999).

80. R. Pathak, T. L. Hendrickson, B. Imperiali, Sulfhydryl Modification of the Yeast Wbp1p Inhibits Oligosaccharyl Transferase Activity. *Biochemistry* **34**, 4179-4185 (1995).

81. Q. Yan, W. J. Lennarz, Studies on the function of oligosaccharyl transferase subunits. Stt3p is directly involved in the glycosylation process. *J Biol Chem* **277**, 47692-47700 (2002).

82. S. Shrimal, R. Gilmore, Oligosaccharyltransferase structures provide novel insight into the mechanism of asparagine-linked glycosylation in prokaryotic and eukaryotic cells. *Glycobiology*, (2018).

83. M. Chavan, A. Yan, W. J. Lennarz, Subunits of the translocon interact with components of the oligosaccharyl transferase complex. *J Biol Chem* **280**, 22917-22924 (2005).

84. H. Kim, H. Park, L. Montalvo, W. J. Lennarz, Studies on the role of the hydrophobic domain of Ost4p in interactions with other subunits of yeast oligosaccharyl transferase. *Proc Natl Acad Sci U S A* **97**, 1516-1520 (2000).

85. H. Kim, Q. Yan, G. Von Heijne, G. A. Caputo, W. J. Lennarz, Determination of the membrane topology of Ost4p and its subunit interactions in the oligosaccharyltransferase complex in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **100**, 7460-7464 (2003).

86. U. Spirig, D. Bodmer, M. Wacker, P. Burda, M. Aebi, The 3.4-kDa Ost4 protein is required for the assembly of two distinct oligosaccharyltransferase complexes in yeast. *Glycobiology* **15**, 1396-1406 (2005).

CHAPTER 2

EXPRESSION, PURIFICATION, AND BIOPHYSICAL CHARACTERIZATION OF A MINI-MEMBRANE PROTEIN, Ost4 A SUBUNIT OF YEAST OLIGOSACCHARYLTRANSFERASE AND IT'S FUNCTIONALLY IMPORTANT MUTANT, Ost4V23D.

2.1 Introduction

Although the structural biology field has made tremendous advancements, structural and functional characterizations of integral membrane proteins remain a challenging task. The main limiting factors for structure determination at atomic resolution are the production of pure, homogeneous membrane proteins and their reconstitution in a suitable membrane mimetic. Of 150916 available protein structures, only 1178 (0.78%) PDB structures belong to unique membrane proteins (Figure 2.1a and Figure 2.1b). However, 30% of the genome is composed of membrane proteins and over 60% of current drug targets are membrane proteins (*1, 2*). This indicates the significance of membrane proteins. Purification of proteins to their homogeneity is the key to the structure determination of membrane proteins as well as for water-soluble proteins. Therefore, the target protein must be expressed and then purified to its homogeneity. However, the purification protocol is different for different proteins. Membrane proteins in particular are difficult to handle because of their hydrophobic surfaces.

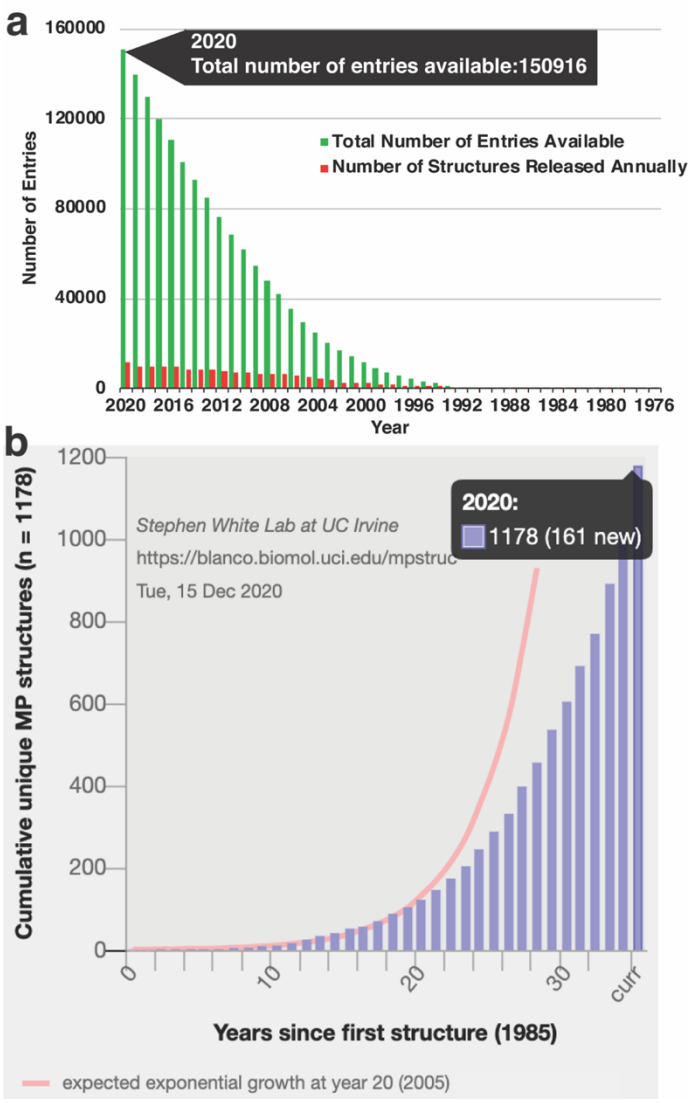


Figure 2.1:PDB statistics of protein available protein structures as of 15th December 2020. (a) The total number of protein structures (green bars) and the total number of protein structures released annually (red bars). (b) The cumulative number of unique membrane proteins available since the first membrane protein structure in 1985. The PDB statistics for total protein entries and unique membrane proteins were obtained from www.rcsb.org/stats/growth/growth-protein and www.blanco.biomol.uci.edu/mpstruc, respectively.

Even though subunits of yeast OST were identified, cloned and sequenced over 36 years ago, there is a gap in our knowledge and understanding of the individual role and most important function of each subunit of OST in the N-glycosylation protein modification reaction. The major obstacle in membrane protein research is the heterologous expression and purification of pure, homogenous integral membrane proteins (IMPs) and their reconstitution in a suitable membrane mimetic. As a result, biochemical, biophysical and high-resolution structural characterizations of IMPs remain challenging.

The heterologous expression and purification of recombinant Ost4 have already been reported from our laboratory (*3*). In this chapter, the overexpression, purification, reconstitution and biophysical characterization of a functionally important mutant protein, Ost4V23D will be reported. Additionally, a comparison of the secondary structure and conformation of Ost4V23D with Ost4 suggests that the mutation affects both secondary and tertiary structure of the wild type protein.

2.2 Materials and methods

2.2.1 Transformation of GB1 – Ost4 into *E. coli* bacterial cell

GB1 – Ost4 was transformed into *E. coli* BL21DE3pLysS cells according to following procedures. The GB1 – OST4 plasmid from a -20 ºC freezer and BL21DE3pLysS competent cells from a -80 ºC freezer were thawed on ice for 45 minutes. 2 µL of the thawed plasmid was added to the e-tube containing competent cells followed by incubation for 45 minutes on ice. The heat shock for 45 sec was given by keeping the mixture of the plasmid and the competent cells on water bath at 42 ºC. The culture was incubated for

another 45 minutes on ice. Exactly 200 μL of SOC (super optimal broth with catabolite repression) media was added to the e-tube containing the mixture of plasmid and cells followed by incubation for 20 minutes at 37 ºC with shaking. At the end, 20 μL of the bacterial cell culture was plated on an LB (Liquid Broth) plate containing ampicillin antibiotic. The plate was then incubated overnight at 37 ºC. The plating was generally performed in the evening so as to obtain fully grown bacterial colonies by the next day morning. Once the bacterial colony were fully grown on the plate, the plate was stored at 4 ºC in the cold room for later use.

2.2.2 Preparation of overnight culture of GB1 – Ost4 transformed *E. coli* cells

For 1 L of expression culture, 25 mL of overnight culture was prepared. For the preparation of overnight culture, 25 μL of 100 mg/mL of ampicillin was added to a conical flask 25 mL of LB media. A single isolated colony was selected from the LB ampicillin plate and the colony was picked up by using a sterile pipette tip added into the conical flask containing the LB media and ampicillin antibiotic. The culture was allowed to grow overnight.

2.2.3 Expression of GB1 – Ost4 protein

6X-His-tagged GB1-Ost4 was expressed in *E. coli* BL21DE3pLysS cells following the previously described procedure (*3*). For expression, the overnight culture of the GB1-Ost4 transformed *E. coli* cells were diluted to OD_{600} of 0.06 using M_9 media containing 4 g/L ^{13}C-glucose, 1.2 g/L $^{15}N–NH_4Cl$, 1 mg/L thiamine, 100 μg/mL ampicillin 2 mM

$MgSO_4$, 50 µM $CaCl_2$, and 100 µM trace elements. Generally, 1 L of expression media was prepared at a time using four flasks each containing 250 mL of M_9 media. The following amount of ingredients (Table 2.1) were added for 250 ml of M_9 media to produce a ^{13}C, ^{15}N double-labelled protein.

Table 2.1: The concentration and volume of ingredients added for 250 ml of double-labelled protein expression M_9 media.

SN	Ingredient	Amount added	Final concentration
1	^{13}C glucose	5 mL	4 g/L
2	Trace elements	250 µL	100 µM
3	Thiamine	1.25 mL	1 mg/L
4	$MgSO_4$	500 µL	2 mM
5	$CaCl_2$	125 µL	50 µM
6	Ampicillin	250 µL	100 µg/mL
7	Yeast extract	500 µL	0.2 g/L

The expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at OD_{600} of 0.4-0.6 and grown at 30 ºC for 8 hr. The cells were then harvested by centrifugation at 9000 rpm for 25 min at 4 ºC. The supernatant liquid was discarded and the cell lysate was stored at 4 ºC freezer for purification when needed.

2.2.4 Purification of GB1 – Ost4 protein

For the purification of GB1 – Ost4 protein, the harvested cells from 250 mL of cell culture were suspended in 15 mL of lysis buffer containing 50 mM sodium phosphate buffer, pH

7.4, 200 mM sodium chloride, 0.01% sodium azide). The cells were lysed by sonicating the mixture solution using a 6-second sonication pulse followed by a 1-minute equilibration on the ice water bath 8 times. The cell debris was separated from the protein solution by centrifugation at 12000 rpm for 20 minutes. The cell debris was discarded and the supernatant liquid was collected for further purification of GB1 – Ost4 protein.

Since the recombinant GB1 – Ost4 protein contained a 6X-His-tag at the C-terminus, the protein was purified to homogeneity by using Ni – NTA chromatography (ThermoFisher Scientific) For the purification, 10 mM imidazole and 300 mM NaCl were added to the supernatant liquid containing 6X-His-tagged GB1-Ost4. To each column, 4 mL of Ni-NTA resin was added followed by extensive washing of the resin with nano pure water to remove the ethanol used for storing the resin. Each column was equilibrated 3 times with equilibration buffer (20 mM sodium phosphate buffer, pH 7.5, 500 mM sodium chloride, 0.01% sodium azide and 10 mM imidazole). During equilibration, each time 4 mL of equilibration buffer was added followed by equilibration for 1 h at room temperature with shaking. Then, 4 ml of cell lysate was loaded added to each column followed by washing 10 times with wash buffer (20 mM sodium phosphate buffer, pH 7.5, 500 mM sodium chloride, 0.01% sodium azide and 50 mM imidazole). For washing purpose, 1 mL of wash buffer was added and equilibrated for 1 h with shaking. Finally, the protein was eluted with elution buffer (20 mM sodium phosphate buffer, pH 7.5, 500 mM sodium chloride, 0.01% sodium azide and 500 mM imidazole). For eluting the protein, 500 μL of elution buffer was added followed by equilibration for 1 h with shaking. The elution was monitored by

measuring OD280 of the eluted protein and was considered completed once OD280 of the elute reached ~ 0.05. The eluted protein was stored at 4 ºC for further purification of Ost4.

2.2.5 Removal of GB1 – tag from GB1 – Ost4

The purified GB1 – Ost4 was buffer exchanged with 1X cleavage buffer (25 mM Tris-HCl buffer, pH 7.4, 150 mM NaCl and 2.5 mM $CaCl_2$) 3 times. Each time, 10 mL of 1X cleavage buffer was added to a 15 mL 3K molecular weight cutoff ultrafiltration unit containing 2 mL of protein. Finally, the centrifugation was carried out until the final volume reached ~800 µL. From the stock of 800 µL protein, 200 µL of protein was used for setting up one cleavage reaction. Therefore, multiple cleavage reactions were set up at the same time. The cleaving capacity of the thrombin enzyme (VWR) was 100 µg/unit. The concentration of available thrombin enzyme was 1 unit/µL. Therefore, 80 µL of the thrombin was used which could cleave about 8 mg of GB1 – Ost4 protein. Since only 200 µL of GB1 – Ost4 which would contains about 6.68 mg of protein, 80 µL of thrombin is sufficient to complete the cleavage reaction. Each cleavage reaction was set by adding the components as listed in Table 2.2.

Table 2.2: Reaction components for the removal of GB1 tag from GB1 – Ost4 protein

Component	Volume
GB1 – Ost4 Protein	200 µL (6.68 mg)
10X cleavage buffer	18 µL
Thrombin (VWR)	80 µL

0.1% SDS	2.8 μL
Total volume	300.8 μL

After mixing the reaction components in a 1.5 mL Eppendorf tube, the reaction mixture was incubated for 5 days with shaking at room temperature (RT). Upon completion of the cleavage reaction, the supernatant containing the GB1 tag was separated from the protein pellet by using centrifugation. The Ost4 pellet was washed with nano pure water 4 times. Each wash was carried out by adding 200 μL of water followed by suspending the pellet by pipetting, centrifugation for 15 minutes and removal of supernatant water. The protein pellet was dried by using a speedvac vacuum concentrator. The dried powder of Ost4 protein was stored at 4 ºC for later use.

2.2.6 Production of GB1 – Ost4V23D protein

2.2.6.1 Mutagenesis of OST4 gene to obtain OST4V23D

OST4 gene was previously cloned into GEV 2 vector (*3*). The mutagenesis of OST4 to obtain Ost4V23D mutant was carried out previously in Dr Mohanty's laboratory. The Ost4V23D mutant was already available for my project. Briefly, the mutation of valine 23 to aspartate was carried out by using the Quickchange Site-Directed Mutagenesis Kit from Stratagene. The following forward and reverse primer were used: 5' GTGATGATGACTTTAATTGACATTTACCATGCTGTTGACTCC 3' (forward) and 5' GGAGTCAACAGCATGGTAAATGTCAATTAAAGTCATCATCAC 3' (reverse). The PCR amplified reaction mixture was treated with an enzyme DpnI to digest the template DNA. The incorporation of the mutation was verified by DNA sequencing.

2.2.6.2 GB1 – Ost4V23D expression

Similar to the expression procedure of Ost4 protein, unlabeled and ^{15}N labeled GB1 – Ost4V23D mutant proteins were produced in *E. coli* BL21DE3pLysS cells (Strategene). Expression of the Ost4V23D in the GEV2 vector was under the control of an IPTG (isopropyl β-D-thiogalactopyranoside) inducible promoter. Briefly, the overnight starter culture was diluted to an OD_{600} of 0.06 in fresh LB medium containing 100 μg/mL ampicillin and grown at 37 ^{o}C to an OD_{600} of 0.4-0.6. At that point, the temperature was reduced to 30 ^{o}C and the expression of the protein was induced by the addition of IPTG to a final concentration of 1 mM. After 4 hours, the cells were harvested by incubating on ice for 30 minutes followed by centrifugation at 9000 rpm for 20 min at 4 ^{o}C. The cell pellet was frozen at -20 ^{o}C until needed. For NMR studies, ^{15}N uniformly labeled protein was produced by using M_9 minimal media containing ^{15}N ammonium chloride (Cambridge Isotope Laboratories, MA). The cell culture was grown for 8 hours after induction with IPTG and the cells were harvested the same as unlabeled protein. The protein expression was verified by Tris-Tricine gel electrophoresis.

2.2.6.3 GB1 – Ost4V23D purification

2.2.6.3.1 Lysis of GB1 – Ost4V23D cells

Prior to lysis of the GB1 – Ost4V23D cells, the cell pellet was freeze-thawed six times, by freezing it in liquid nitrogen for 5 minutes and thawing on ice. The cell pellet was suspended in lysis buffer (50 mM sodium phosphate buffer, pH 6.5, 200 mM sodium

chloride, 0.01% sodium azide) and sonicated eight times with a 6-sec pulse followed by incubating 1 minute on the ice water bath. The cell debris was separated from cell lysate by centrifuging at 12000 rpm for 20 minutes at 4 ^{o}C. The cell lysate containing GB1-Ost4V23D was stored at 4 °C for further purification.

2.2.6.3.2 Purification of GB1 – Ost4V23D by Ni- NTA chromatography

The purification of GB1 – Ost4V23D protein followed a similar procedure as explained for GB1 – Ost4 protein. Briefly, Ni-NTA affinity chromatography was used for the purification of GB1 – Ost4V23D protein. The protein was purified as follows under gravity. For 4 mL Ni -NTA (ThermoFisher Scientific) resin was packed into one column for 4 mL of cell lysate. Each column was washed with nano-pure water 10 times to remove storage ethanol. The column was equilibrated with binding buffer (50 mM sodium phosphate buffer, pH 6.5, 200 mM sodium chloride, 0.01% sodium azide and 10 mM imidazole). A 4 mL GB1 – Ost4V23D of cell lysate containing 10 mM imidazole was loaded onto the column. The binding of the protein to the resin was continued overnight in a shaker. The unbound proteins in the lysate were removed by letting the unbound protein solution flow through the column. The impurities were removed with wash buffer (50 mM sodium phosphate buffer, pH 6.5, 200 mM sodium chloride, 0.01% sodium azide and 30 mM imidazole) several (10) times. The bound protein was eluted with elution buffer (50 mM sodium phosphate buffer, pH 6.5, 200 mM sodium chloride, 0.01% sodium azide and 500 mM imidazole). Completion of the protein elution was monitored by measuring OD280 of the protein.

2.2.6.3.3 Cleavage of GB1 Tag from Ost4V23D protein

The removal of the GB1 – tag from GB1-Ost4V23D protein was carried out by following a procedure similar to that as described for Ost4 protein. Briefly, the pure GB1-Ost4V23D protein was concentrated and the buffer was exchanged with cleavage buffer (25 mM Tris-HCl buffer, pH 7.4, 150 mM NaCl and 2.5 mM $CaCl_2$). The cleavage of GB1 – Ost4V23D was carried out as reported previously (*24*) except for minor changes in the buffer conditions. In the previous study, the cleavage was performed using a buffer containing 20 mM Tris-HCl buffer, pH 8.4, 150 mM NaCl, 2.5 mM $CaCl_2$, and 0.01% SDS. Here, the cleavage was carried out using a cleavage buffer containing 25 mM Tris-HCl buffer, pH 7.4, 150 mM NaCl, 2.5 mM $CaCl_2$, and 0.01% SDS. The cleavage reaction was set by following a procedure similar to that done for the Ost4 protein as shown in Table 2.2. The cleavage reaction mixture was incubated for 5 days at RT with shaking. The completion of the cleavage reaction was monitored by SDS PAGE. Soluble GB1 was removed from insoluble Ost4V23D protein by centrifugation of the reaction mixture at 13200 rpm at room temperature after the completion of the reaction. The protein pellet was washed with nano-pure water to remove any contaminating GB1 protein. Pure Ost4V23D containing the 6x – Histidine tag was reconstituted in an appropriate concentration of either unlabeled or deuterated dodecylphosphocholine (DPC) micelles at a suitable buffer condition for biophysical characterization using CD and NMR as described below.

2.2.7 Circular dichroism

2.2.7.1 Sample preparation for Circular Dichroism (CD) experiments:

Approximately 6.7 mL of 15 mM sodium phosphate buffer (pH 6.5) was diluted with 3.3 mL of nano pure water to obtain 10 mL of 10 mM phosphate buffer (pH 6.5). To 10 mL of this buffer, 351 mg of unlabeled dodecylphosphocholine (DPC, MW 351.5 g/mol) was added and dissolved completely yielding a CD buffer (10 mM phosphate (pH 6.5) containing 100 mM DPC micelles). 1.7 mg and 1.2 mg of dry pellets of unlabeled Ost4 and Ost4V23D, respectively, were dissolved in 1 mL of CD buffer (10 mM phosphate buffer, pH 6.5 containing 100 mM DPC micelles) yielding a stock solution of 1.7 mg/mL (343.5 μM) Ost4 and 1.2 mg/mL (242.5 μM) Ost4V23D. The molecular weights of Ost4 and Ost4V23D were 5.3501 kDa and 5.3657 kDa, respectively. The stock solution of each protein was separately diluted with 10 mM phosphate buffer (pH 6.5) containing 100 mM DPC micelles to a final concentration of 48.5 μM for all CD experiments.

2.2.7.2 CD experiments

All circular dichroism (CD) experiments except for the Ost4 were carried out with a Jasco J – 810 automatic recording spectropolarimeter in our laboratory in the Department of Chemistry at Oklahoma State University (Stillwater, OK). The Ost4 data were collected on a Jasco J – 715 in the Department of Biochemistry and Molecular Biology at Oklahoma State University (Stillwater, OK). All CD data were collected at room temperature with a 0.05 cm quartz cell. During data acquisition, the data were averaged over 5 scans with a response time of 4 seconds and with a scan speed of 50 nm/min. All spectra were corrected by subtracting the blank spectra (buffer containing an appropriate amount of DPC). CD ellipticity values were converted to normalized values (mean molar ellipticity per residue) by the standard method available in the spectra manager software of the CD instrument.

The secondary structure contents of Ost4 and Ost4V23D were quantified through deconvolution of CD spectra by using CDSSTR, CONTINLL and SELCON3 programs incorporated in the CDPro software package (*4*).

2.2.8 NMR experiment

2.2.8.1 NMR sample preparation

NMR samples were prepared by dissolving 1.2 mg and 1.3 mg of dry pellet of Ost4 and Ost4V23D proteins, respectively in 500 μL of NMR buffer (50 mM phosphate buffer, pH 6.5 containing 1 mM EDTA, 0.01% NaN_3) containing 100 mM deuterated DPC micelles. The weight of dry protein pellet was obtained by measuring the weight of empty e-tube and subtracting the weight of the empty e-tube from the weight of the e-tube containing dry protein pellet. The protein pellet was suspended in the NMR buffer by vortexing and then water bath sonicated at 35 ºC until all the protein completely dissolved in the NMR buffer containing DPC micelles. The water bath temperature was maintained < 45 ºC by adding ice to the water bath throughout the sonication process. Approximately all protein dissolved in the NMR buffer in about 24 hours. Finally, the protein solution was centrifugated at 13200 RPM for 15 minutes to remove any undissolved particles. The supernatant liquid containing reconstituted protein was transferred into a Shigemi tube for NMR data acquisition.

2.2.8.2 NMR data acquisition and processing

All the NMR experiments were collected using the Varian INOVA 600 MHz spectrometer equipped with an effective three-channel system with linear amplifiers, waveform generators, Z pulsed-field gradient capability, and FTS chiller for temperature control at Oklahoma State University. The NMR samples used for [^{1}H, ^{15}N] HSQC experiments consisted of 400 µL of 0.750 -1.3 mM of uniformly ^{15}N labeled wild type Ost4 or Ost4V23D mutant protein in 50 mM phosphate buffer, pH 6.5 containing 1 mM EDTA, 0.01% NaN_3, 100 mM DPC, and 5% D_2O (used as a lock solvent) in a Shigemi tube. 2D [^{1}H, ^{15}N] heteronuclear single quantum coherence (HSQC) spectra were collected for Ost4 and Ost4V23D protein samples at 35 °C. The HSQC spectra for both the proteins were collected with 256 and 2048 complex points in the ^{15}N dimension (t_1 time domain) and ^{1}H dimension (t_2 time domain), respectively. The data were zero filled to 512 x 4096 and apodized using a Gaussian window function. Finally, the data was Fourier transferred for the spectra analysis. The NMR data were processed by using Nmrpipe (*5*) and analyzed with Sparky(*6*).

2.3 Results

2.3.1 Overexpression and purification of GB1-Ost4 protein

GB1-Ost4 protein was successfully overexpressed and purified by following a previously published protocol (*3*). The expression, purification, and cleavage of wildtype protein were monitored by SDS page gel as shown in Figure 2.2.

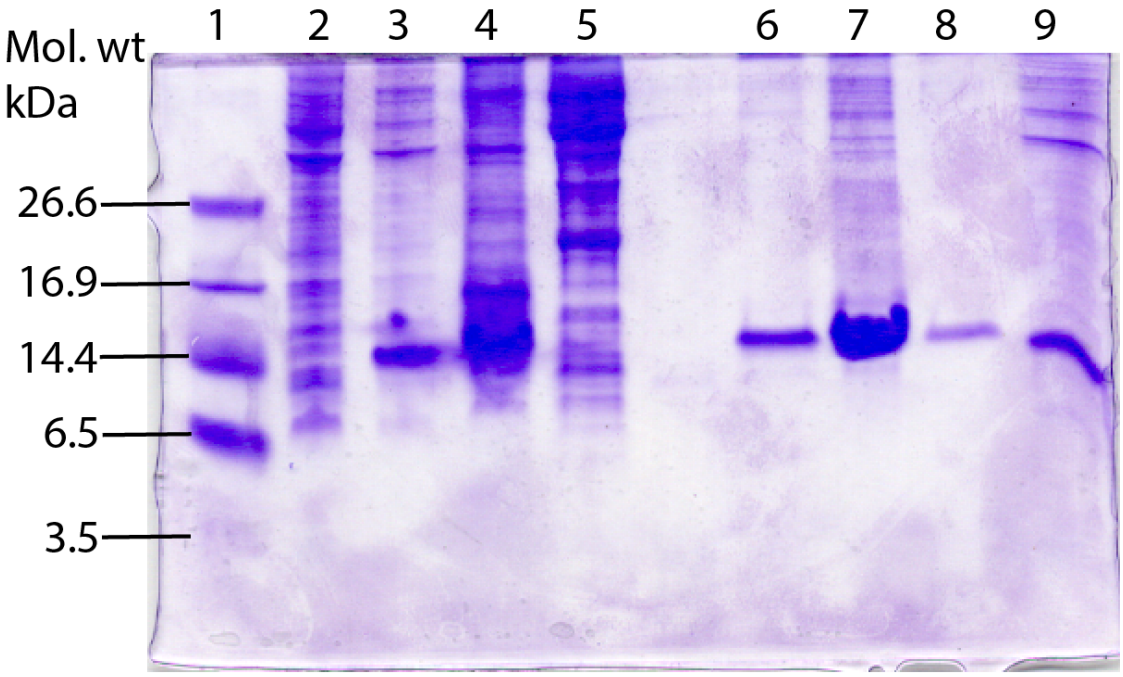


Figure 2.2: SDS-PAGE gel picture of the expression and purification profile of ^{15}N-labeled Ost4. Lane 1: protein marker, lane 2: cell lysate at the time of induction (T_0), lane 3: cell lysate after 8 hours of induction (T_8) with 1mM IPTG (10 µL), lane 4: cell lysate after 8 hours of induction (T_8) with 1mM IPTG (25 µL), lane 5: Flow-through after binding, lane 6: empty column, lane 7: eluted GB1-Ost4 (5 µL), lane 8: eluted GB1-Ost4 (25 µL), lane 9: Initial wash, lane 10: cell lysate after 8 hours of induction (T_8) with 1mM IPTG (5 µL).

2.3.2 Mutagenesis, overexpression and purification of GB1-Ost4V23D

The pGEV2-Ost4V23D mutant plasmid was constructed from the pGEV2-Ost4 vector by site-directed mutagenesis. The incorporation of mutation was confirmed by DNA sequencing (Figure 2.3). Native as well as mutant plasmids were transformed into *E. coli* BL21 (DE3) pLysS cells. Protein expression was optimized by altering temperatures and IPTG concentrations. The GB1-Ost4V23D mutant protein was expressed as a soluble protein. The mutant protein was successfully purified using 50 mM sodium phosphate buffer (pH 6.5) 200 mM NaCl and an appropriate amount of imidazole using Ni-NTA column chromatography. The recombinant GB1-Ost4V23D was found to be very stable under these conditions and could be purified to homogeneity (Figure 2.4 a and b). Pure Ost4V23D was obtained when the GB1 tag was cleaved successfully with thrombin (Figure 2.4 c). In the cleavage reaction, Ost4V23D separated as a white precipitate leaving the GB1 tag in solution eliminating the necessity for further purification. The addition of 0.01% SDS was found to be necessary for efficient cleavage of GB1 from Ost4V23D. The C-terminal 6X-His-tag was not cleaved from either Ost4 or Ost4V23D. A total of 21.5 mg and 18.2 mg of pure GB1-Ost4 and GB1-Ost4V23D proteins, respectively, were obtained from one L of cell culture.

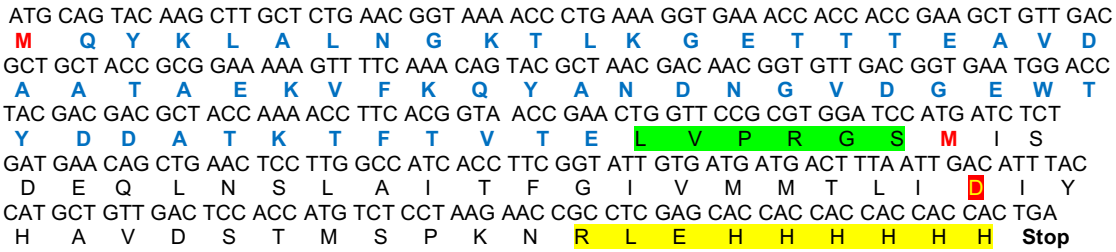
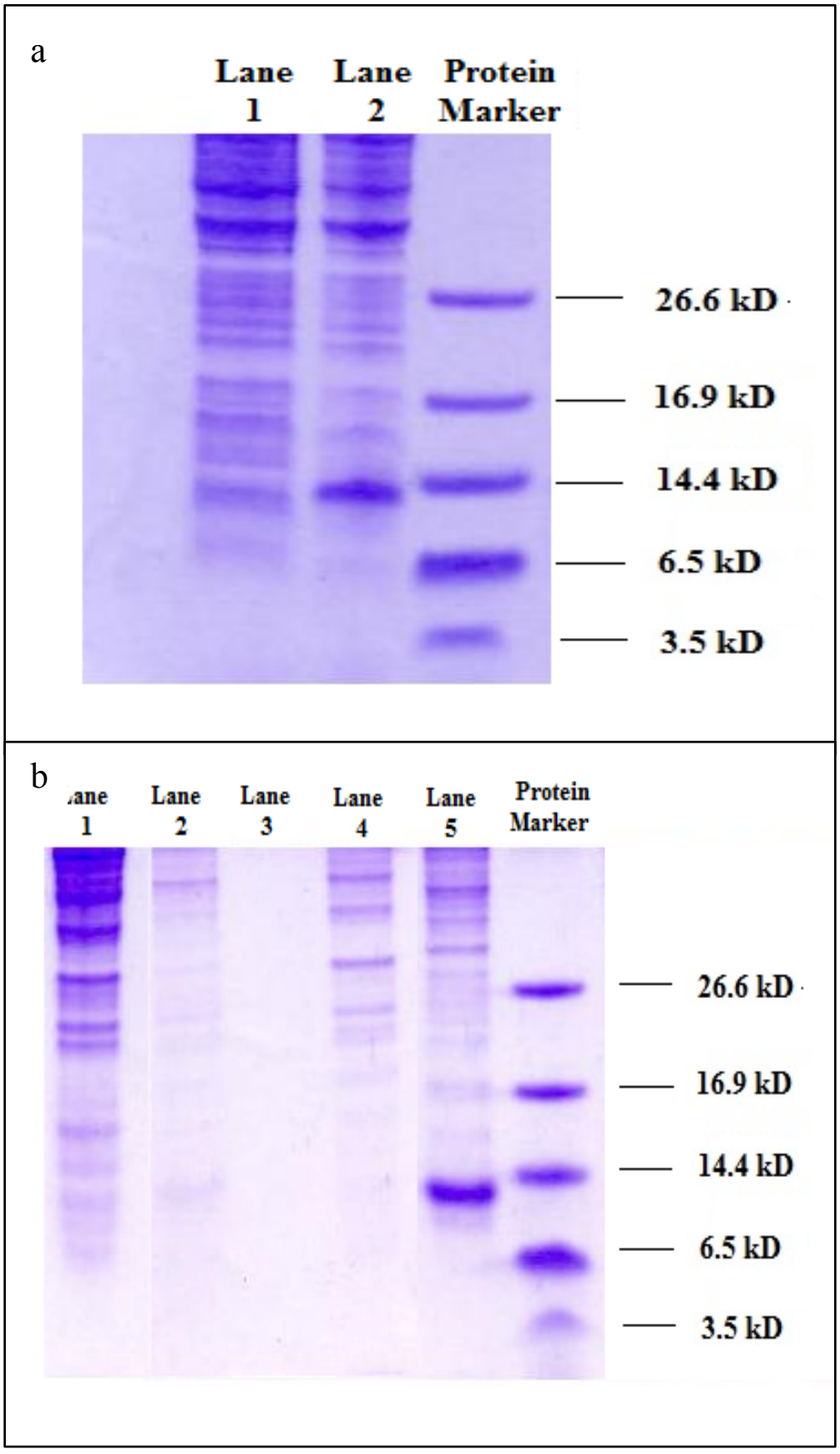


Figure 2.3: Nucleotide (top) and corresponding amino acid (bottom) sequences of GB1-Ost4V23D. Amino acid sequences of GB1 protein are represented by blue letters. Amino acids in red are the first amino acids in the GB1 or Ost4 sequence. Amino acids in green background represent thrombin cleavage site. Amino acids in yellow background are those which are not present originally in Ost4 sequence including the hexa-histidine tag. The mutation site is represented by yellow letter highlighted in red.


a
Lane 1
Lane 2
Protein Marker
26.6 kD
16.9 kD
14.4 kD
6.5 kD
3.5 kD
b
ane 1
Lane 2
Lane 3
Lane 4
Lane 5
Protein Marker
26.6 kD
16.9 kD
14.4 kD
6.5 kD
3.5 kD

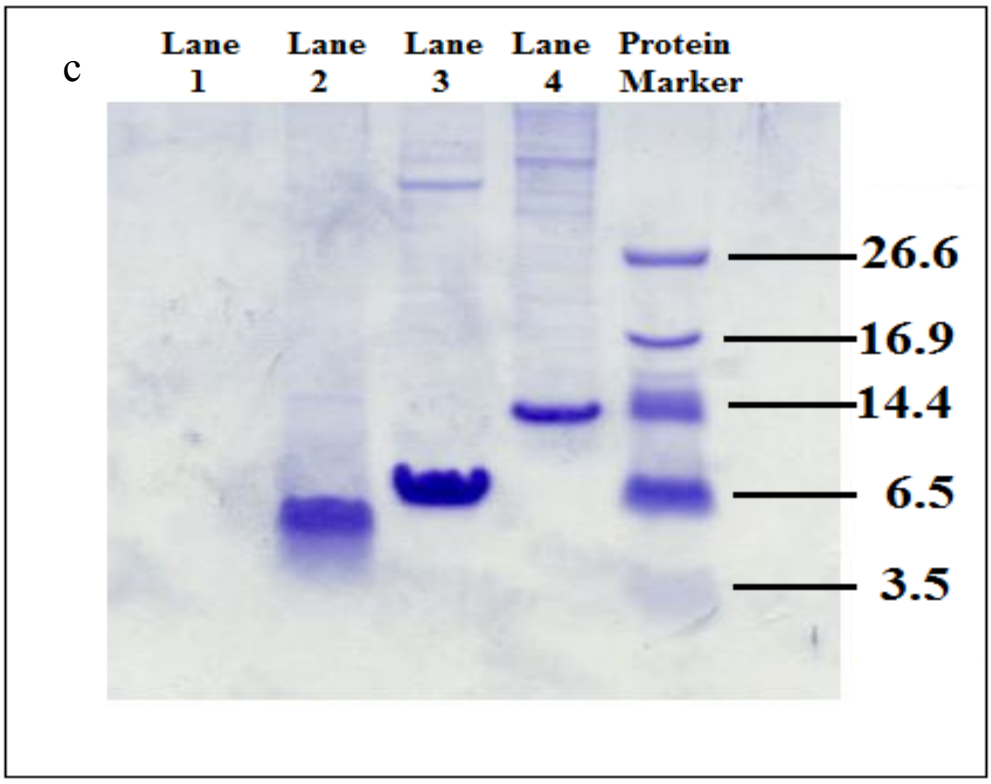


Figure 2.4: SDS-PAGE gel picture of the expression (a), purification (b), and cleavage (c) profile of ^{15}N-labeled Ost4V23D. The protein marker lane is labeled in the figure. In panel (a), lane 1: cell lysate at the time of induction (T_0), lane 2: cell lysate after 8 hours of induction (T_8) with 1 mM IPTG. In panel (b), lanes 1-4: Washings of cell lysate loaded on Ni-NTA, lane 5: eluted protein. In panel (c), lane 1: pallet wash, lane 2: pallet after cleavage, lane 3: supernatant liquid after cleavage, lane 4: eluted protein before cleavage.

The C-terminus of the chemically synthesized Ost4 is disordered and has not been shown to have any functional or structural significance (*7*). Therefore, the tag was left intact in the recombinant Ost4 and Ost4V23D proteins.

2.3.3 Characterization of the Ost4 and Ost4V23D by far-UV CD spectroscopy

Far UV-CD spectroscopy was carried out to explore the secondary structure of Ost4 and Ost4V23D under identical conditions. DPC is a mild detergent that generally does not denature proteins (*8*). DPC has been used as a membrane mimetic for the determination of NMR structures of OmpA (177 residues) (*9*), outer membrane enzyme PagP (164 residues) (*10*), phospholamban (52 residues) (*11*) and CHIF (a member of Na^+-, K^+-ATPase regulatory membrane proteins, 67 residues) (*12*) etc. To determine optimal DPC concentration, the far UV-CD region was used for detergent screening. Far-UV CD spectra on Ost4 and Ost4V23D were collected at 50 mM, 100 mM, 200 mM, 300 mM and 400 mM of DPC concentration (Figure 2.5).

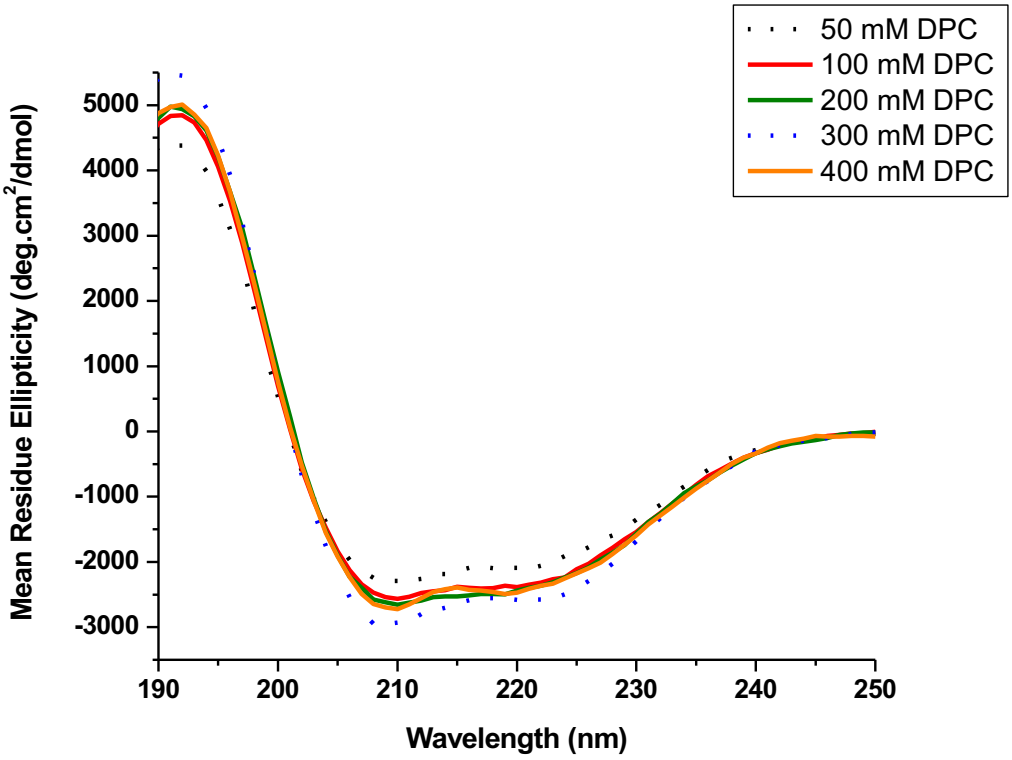


Figure 2.5: Far UV-Circular dichroism (CD) spectroscopic analysis of the Ost4 (wild type) at different 50 mM, 100 mM, 200 mM, 300 mM and 400 mM DPC micelle concentrations

respectively. The protein concentration was 48.5 μM in each of the DPC micelle concentrations. Characteristic minima at 208 and 222 nm in these DPC micelle concentrations show that Ost4 is in a well-folded state in all of the above DPC concentrations but 100 mM DPC is the best concentration having high helical content.

CD data demonstrate that Ost4 and Ost4V23D are helical in DPC micelles with two characteristic minima at 208 and 222 nm (Figure 2.6).

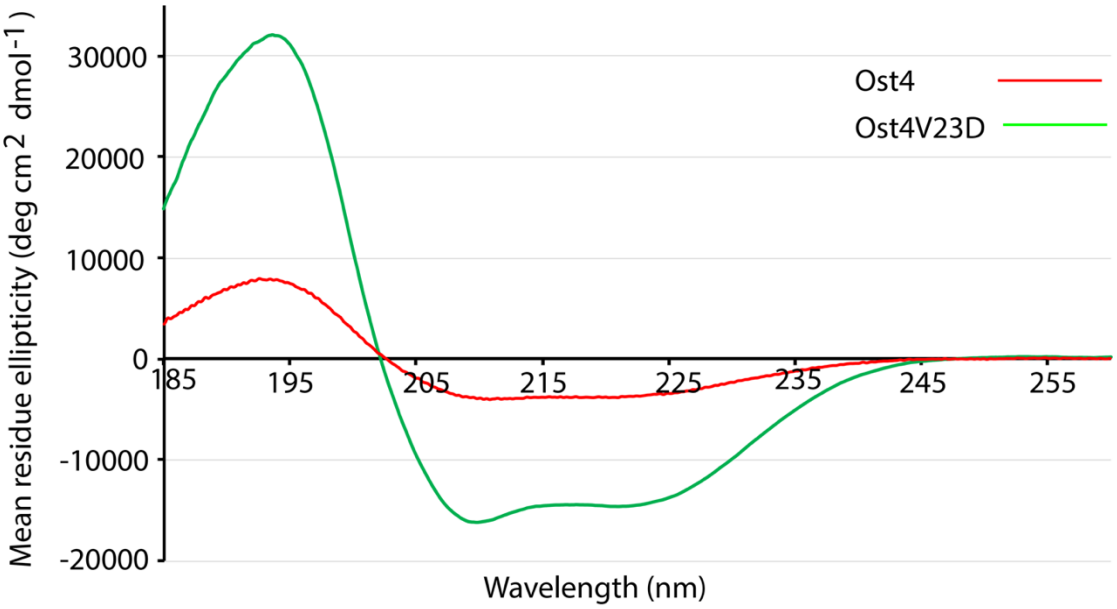


Figure 2.6: Far UV-circular dichroism (CD) spectroscopic analysis of the Ost4 (red) and Ost4V23D (green) mutant protein in 100 mM DPC micelle concentration. The protein concentration for both the proteins was 48.5 μM. Characteristic minima at 208 and 222 nm indicate that Ost4 protein has lower alpha-helical content than Ost4V23D.

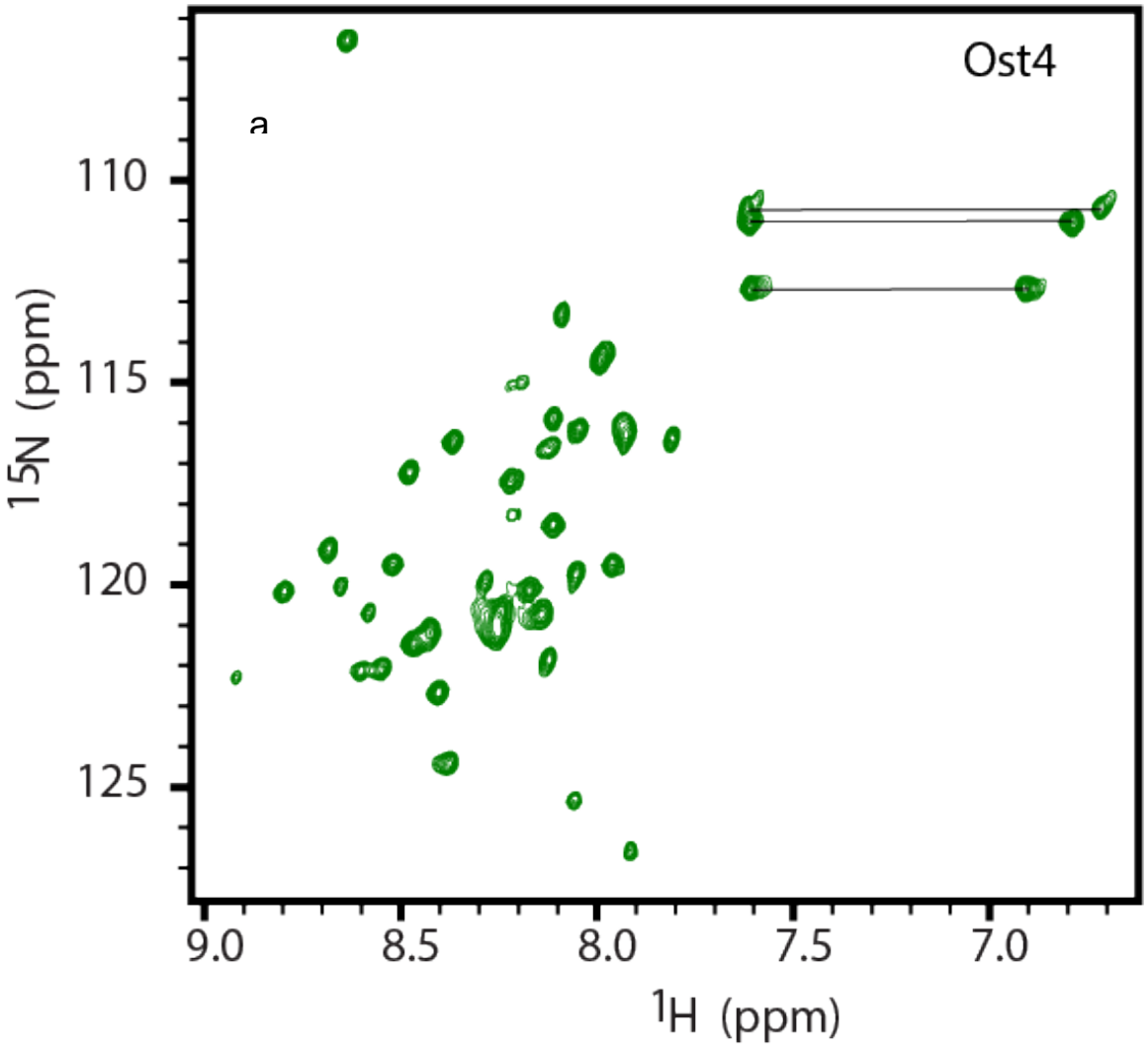
The CD spectra of Ost4 were very similar when the DPC concentration was varied from 50 - 400 mM suggesting that Ost4 maintains its structure under various DPC concentrations (Figure 2.5). The CD spectra for Ost4V23D (Figure 2.5) were identical for 50 mM, 100 mM, and 200 mM DPC but the helicity was reduced when the DPC concentration was

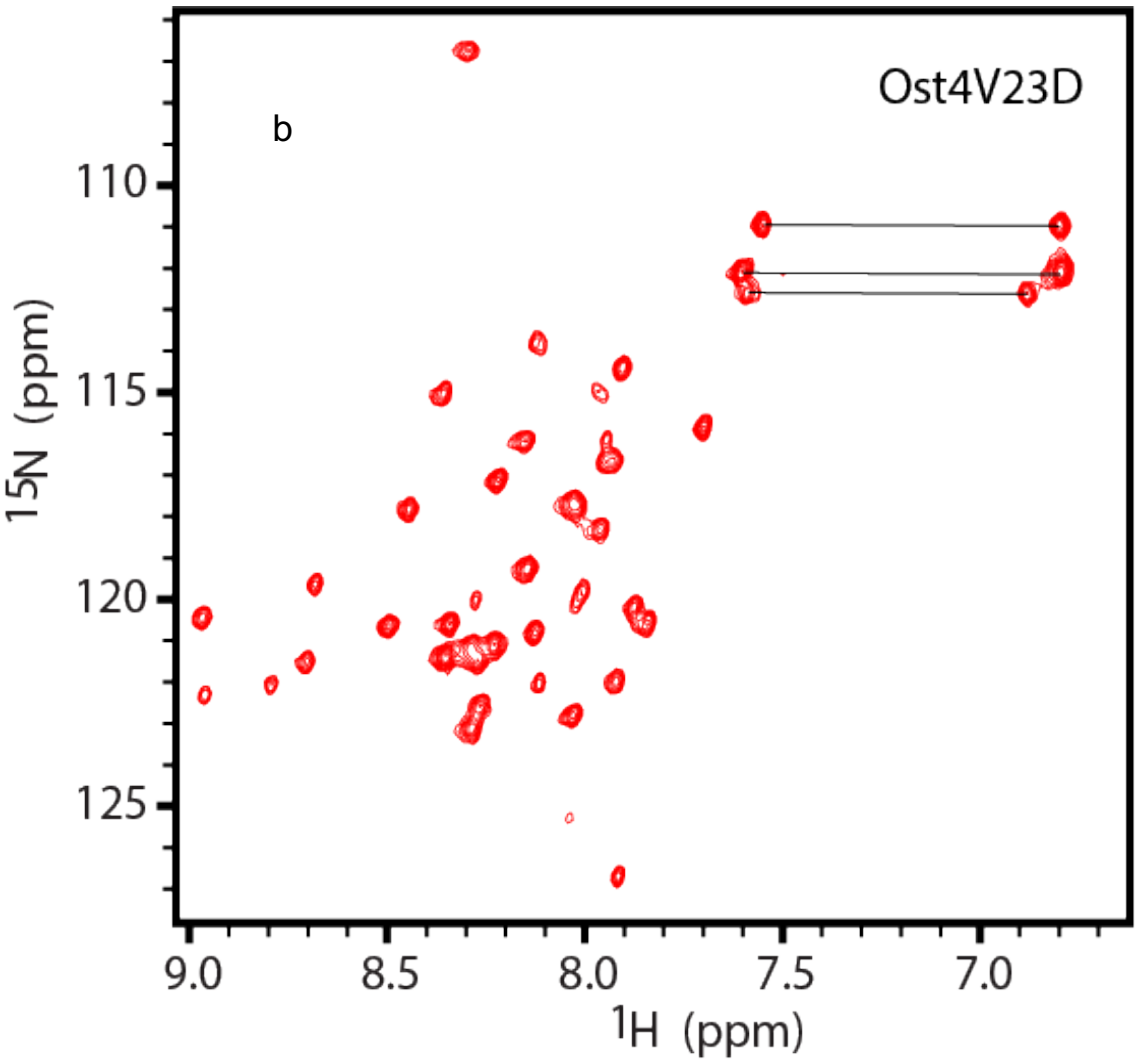
increased to 400 mM. Based on the CD data, 100 mM DPC was chosen as the condition for further characterization of these proteins. To investigate the structural impact of the Val^{23} to Asp mutation in Ost4, biophysical characterization with far UV-CD was carried out under identical conditions. The far UV-CD spectra (Figure 2.6) indicated that although Ost4V23D had a typical α-helical structural characteristic in 100 mM DPC micelle, it was quite different from Ost4 under similar conditions. The Ost4V23D mutant protein had more helical propensity than the wild type protein under identical conditions suggesting that there was a significant change in secondary structure with the point mutation of Val^{23} to Asp (Figure 2.6).

2.3.4 Characterization of Ost4 and Ost4V23D by NMR

The 2D [^{1}H, ^{15}N] HSQC is a very sensitive NMR experiment correlating ^{15}N with the attached amide proton for each residue in the protein backbone except for the prolines. The 2D HSQC spectrum essentially manifests the fingerprint region of a protein providing information on its conformation under the experimental conditions (*13-17*). Any changes in the experimental conditions such as pH or salt or titration of a ligand or mutation/s etc. can cause changes in the resonance positions of amino acid residue/s in the fingerprint region of the protein. A change in chemical shift positions in HSQC is an indication of conformational change in the protein due to the perturbation in its environment. These conformational changes can either be global or local. The global conformational change includes several residues and local conformational change includes a few residues. The 2D [^{1}H, ^{15}N] HSQC spectra were collected on both Ost4 and Ost4V23D (Figures 2.7 a and 2.7b) under identical conditions. Both spectra display well-dispersed peaks indicating both

proteins are properly folded in DPC micelles. However, it is clear from Figure 2.7C that there is a drastic change in the fingerprint region indicating that the wild type and the V23D mutant proteins have distinct conformations.




b
Ost4V23D
110
115
120
125
15N (ppm)
9.0
8.5
8.0
7.5
7.0
1H (ppm)

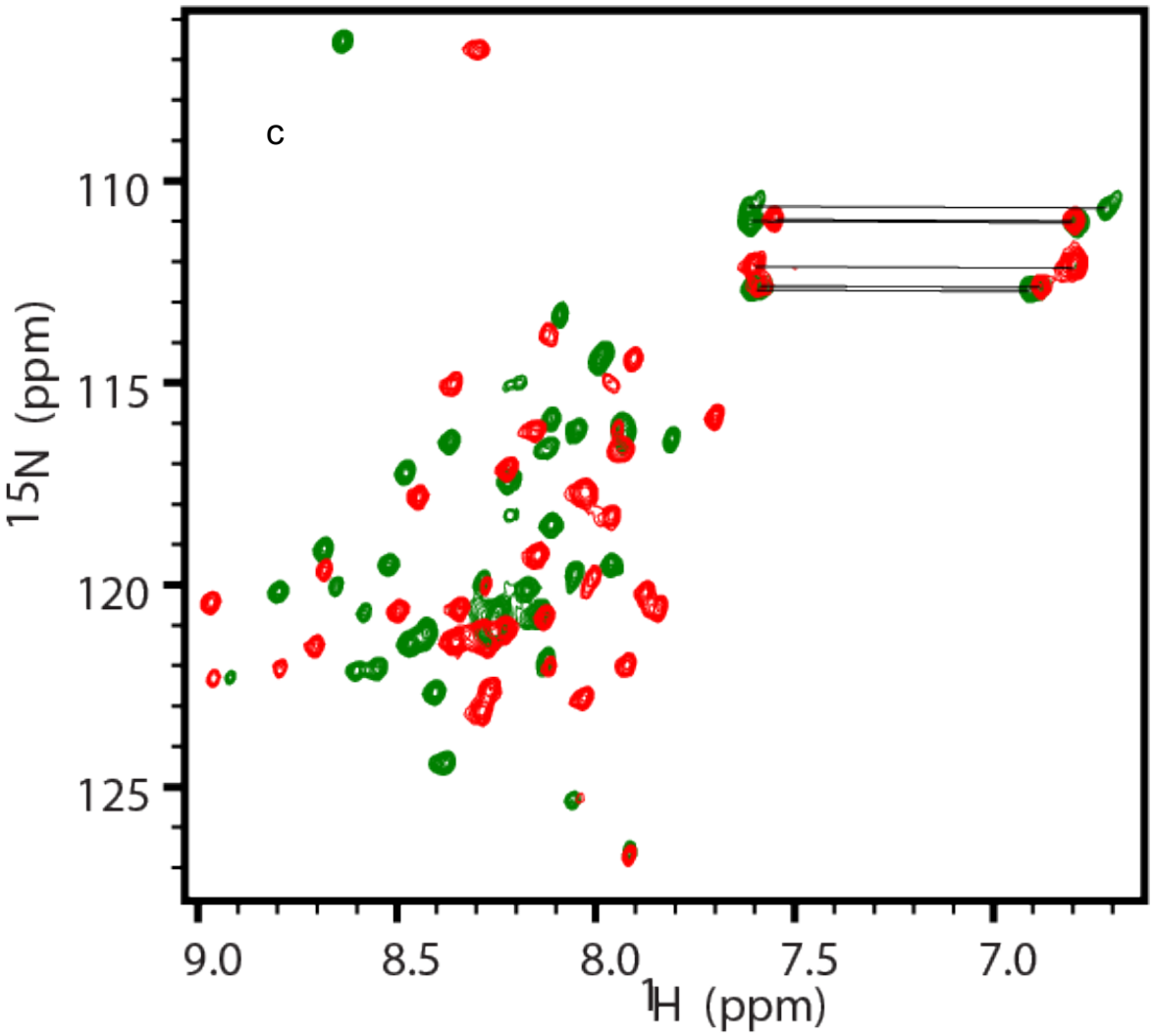


Figure 2.7: 2D [^{1}H, ^{15}N]-heteronuclear single quantum coherence (HSQC) spectrum of ^{15}N-labeled Ost4 and Ost4V23D. (a) HSQC spectrum of ^{15}N-labeled Ost4 in 100 mM DPC micelle containing 5% D_2O at pH 6.5. (b) HSQC spectrum of ^{15}N-labeled Ost4V23D in 100 mM DPC micelle containing 5% D_2O at pH 6.5. The peaks were well dispersed for both ^{15}N Ost4 (a) and ^{15}N-Ost4V23D (b) indicating that both the proteins were well-folded. The spectra of Ost4 (c) and Ost4V23D (b) indicate both the proteins are helical proteins due to narrow dispersion in the amide proton regions, which is typical for helical proteins.

(c) Overlay of [^{1}H ^{15}N] HSQC of wild type Ost4 and its point mutant Ost4V23D. The movement of peaks in the fingerprint region of Ost4 and Ost4V23D is an indication of crucial changes in the 3D structure of the protein when the point mutation is performed.

2.4 Discussion

Even though N-glycosylation is an essential and highly conserved protein modification reaction in all eukaryotic organisms, mechanistic details of eukaryotic OST subunits and their role is not yet clear. The key reason is the difficulties with the heterologous expression and production of milligram quantities of IMPs for structural and functional characterization. *E. coli,* which is most frequently used for the production of large quantities of recombinant proteins, has very limited success for the recombinant expression of IMPs (*18*). Therefore, very few methods for expression of recombinant IMPs in *E.coli* are reported.

The overexpression, purification, reconstitution, and biophysical characterization of Ost4V23D protein and its comparison with the WT Ost4 protein have been discussed in this chapter. The V23D mutation results in impaired cell growth and *in vitro* OST activity (*19, 20*). The effect of V23D mutation on the structure and function of Ost4 can be well understood by performing a comparative study on Ost4 and Ost4V23D proteins. The successful expression and purification for the production of the wild type Ost4 in high yield have been reported previously (*3*). The expression of the Ost4V23D follows the same protocol as the WT, except the cell culture were incubated for 4 h at 30 ºC after induction with IPTG for optimum expression as opposed to 37 ºC for 4 h for the WT. Surprisingly,

Ost4V23D does not bind to the Ni-NTA column under the same conditions of pH and NaCl concentration as used for Ost4. After many trials, we were able to optimize the pH and sodium chloride concentration necessary for effective binding of Ost4V23D to Ni-NTA resin. Ost4 binds effectively to the Ni-NTA column at pH 7.4 and 500 mM NaCl concentration whereas Ost4V23D binding occurs at pH 6.5 and 200 mM salt.

Circular dichroism (CD) is a very simple and reliable technique for the characterization of protein secondary structure and stability under various buffer conditions. CD was used to screen detergents as well as to compare the secondary structures of Ost4 and Ost4V23D under identical conditions. CD data collected on these proteins showed that both these proteins are helical when reconstituted in DPC micelles (Figure 2.5 and 2.6). Quantification of secondary structure content through deconvolution of CD spectra from various programs showed that Ost4 and Ost4V23D have approximately 30% and 62% helicity, respectively. Therefore, the CD data has shown that Ost4V23D has higher helical content than WT protein under identical conditions indicating that there is a significant impact in the tertiary structure of Ost4 when Val^{23} is mutated to Asp. The 2D [^{1}H, ^{15}N] HSQC is a very effective technique that indicates whether a protein is properly folded or not (*8, 13*). The suitability of a detergent micelle to be a surrogate for the natural membrane is determined based on various factors including solubility, the stability of the reconstituted membrane protein and the quality of the 2D-HSQC spectrum. 2D-HSQC is a very sensitive NMR experiment that correlates the amide proton to its corresponding nitrogen atom for each amino acid residue in a protein except proline and provides a map of the fingerprint region. The peaks of a well-folded protein are generally well-dispersed in an HSQC spectrum (*8, 9*). Thus, HSQC

can be used to monitor if the protein under study is well-behaved under provided experimental conditions such as pH, temperature, salt concentration etc. In addition, 2D HSQC is an excellent experiment to track the conformational change of a protein due to substrate binding or mutation/s or change in any other experimental conditions such as temperature or pH etc. Thus, change in the fingerprint region of a 2D HSQC spectrum of a protein due to any of the above reasons, is an indication of either local or global conformational change. The 2D [^{1}H, ^{15}N] HSQC spectra of Ost4 and Ost4V23D reconstituted in 100 mM DPC is well-dispersed suggesting that both the proteins are properly folded having stable tertiary structures. However, the fingerprint region of the Ost4V23D mutant is significantly different than that of Ost4 (Figure 2.6) suggesting that the mutation of Val^{23} to Asp affects the tertiary structure of the protein. The dispersion of resonances in the HSQC of the Ost4V23D is significantly narrower than that of the Ost4, which is an indication of higher helical content in the mutant protein. This observation is consistent with CD data as well. It is clear from both NMR and CD data that Ost4V23D has a distinct structure. Detailed structural analysis of both Ost4 and Ost4V23D can unravel the mechanism of Ost4 function and mode of interactions with Ost3 and Stt3 subunits. Ost4 subunit bridges the catalytic subunit, Stt3 with Ost3 in the Stt3-Ost4-Stt3 sub-complex (*19, 20*). These pieces of evidence show that Ost4V23D has a distinct structure from WT Ost4 and a detail structural study of Ost4V23D can provide the molecular basis of Ost4 function.

2.5 Conclusion

In this chapter, the overexpression, purification and biophysical characterization of Ost4 and Ost4V23D have been reported. The mini-membrane mutant protein, Ost4 was

expressed and purified successfully. Similarly, the critical mutant of Ost4, Ost4V23D, was successfully overexpressed as a soluble fusion protein in BL21DE3pLysS cells when fused to GB1 on its N-terminus along with a 6X-His-tag at its C-terminus. GB1 – tagged Ost4 and Ost4V23D recombinant proteins were purified using Ni-NTA column chromatography. The GB1 – tag was successfully cleaved off by thrombin enzyme yielding an insoluble protein precipitate. The pure insoluble Ost4 or Ost4V23D was efficiently separated from soluble GB1 – tag by centrifugation requiring no further purification step. The precipitated Ost4V23D was reconstituted using DPC micelles. Ost4V23D is a highly helical protein based on far UV-CD data. In fact, it is more helical than the wild type protein. The well-dispersed peaks in the 2D-HSQC data demonstrate that the protein is well-folded in DPC micelles and their structure can be determined in this membrane mimetic system. The 2D HSQC spectra of Ost4 and Ost4V23D suggest that mutation of Val^{23} to Asp affects either the conformation or the environment of the Ost4.

2.6 References

1. D. Vitkup, E. Melamud, J. Moult, C. Sander, Completeness in structural genomics. *Nat Struct Biol* **8**, 559-566 (2001).
2. E. Wallin, G. von Heijne, Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci* **7**, 1029-1038 (1998).
3. A. Kumar, P. Ward, U. V. Katre, S. Mohanty, A novel and simple method of production and biophysical characterization of a mini-membrane protein, Ost4p: a subunit of yeast oligosaccharyl transferase. *Biopolymers* **97**, 499-507 (2012).

4. N. Sreerama, R. W. Woody, Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* **287**, 252-260 (2000).

5. F. Delaglio *et al.*, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995).

6. W. Lee, M. Tonelli, J. L. Markley, NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325-1327 (2015).

7. S. Zubkov, W. J. Lennarz, S. Mohanty, Structural basis for the function of a minimembrane protein subunit of yeast oligosaccharyltransferase. *Proc Natl Acad Sci U S A* **101**, 3821-3826 (2004).

8. C. Huang, S. Mohanty, M. Banerjee, A novel method of production and biophysical characterization of the catalytic domain of yeast oligosaccharyl transferase. *Biochemistry* **49**, 1115-1126 (2010).

9. A. Arora, F. Abildgaard, J. H. Bushweller, L. K. Tamm, Structure of outer membrane protein A transmembrane domain by NMR spectroscopy. *Nat Struct Biol* **8**, 334-338 (2001).

10. P. M. Hwang *et al.*, Solution structure and dynamics of the outer membrane enzyme PagP by NMR. *Proc Natl Acad Sci U S A* **99**, 13560-13565 (2002).

11. J. Zamoon, A. Mascioni, D. D. Thomas, G. Veglia, NMR solution structure and topological orientation of monomeric phospholamban in dodecylphosphocholine micelles. *Biophysical journal* **85**, 2589-2598 (2003).

12. X.-M. Gong, C. M. Franzin, K. Thai, J. Yu, F. M. Marassi, Nuclear magnetic resonance structural studies of membrane proteins in micelles and bilayers. *Methods Mol Biol* **400**, 515-529 (2007).

13. A. M. Gronenborn, G. M. Clore, J. M. Louis, P. T. Wingfield, Is human thioredoxin monomeric or dimeric? *Protein Sci* **8**, 426-429 (1999).

14. M. Bieri *et al.*, Macromolecular NMR spectroscopy for the non-spectroscopist: beyond macromolecular solution structure determination. *Febs j* **278**, 704-715 (2011).

15. N. S. John Cavanagh, Wayne Fairbrother, Mark Rance, Arthur Palmer, III, *Protein NMR spectroscopy: principles and practice*. (Academic Press, ed. 2nd, 2006).

16. A. M. Gronenborn, M. K. Frank, G. M. Clore, Core mutants of the immunoglobulin binding domain of streptococcal protein G: Stability and structural integrity. *FEBS Letters* **398**, 312-316 (1996).

17. P. Selenko, Z. Serber, B. Gadea, J. Ruderman, G. Wagner, Quantitative NMR analysis of the protein G B1 domain in Xenopus laevis egg extracts and intact oocytes. *Proceedings of the National Academy of Sciences* **103**, 11904 (2006).

18. C. G. Tate, Overexpression of mammalian integral membrane proteins for structural studies. *FEBS Lett* **504**, 94-98 (2001).

19. H. Kim, H. Park, L. Montalvo, W. J. Lennarz, Studies on the role of the hydrophobic domain of Ost4p in interactions with other subunits of yeast oligosaccharyl transferase. *Proc Natl Acad Sci U S A* **97**, 1516-1520 (2000).

20. H. Kim, Q. Yan, G. Von Heijne, G. A. Caputo, W. J. Lennarz, Determination of the membrane topology of Ost4p and its subunit interactions in the oligosaccharyltransferase complex in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **100**, 7460-7464 (2003).

CHAPTER 3

BACKBONE AND SIDE CHAIN RESONANCE ASSIGNMENTS AND SECONDARY STRUCTURE CALCULATION OF YEAST OLIGOSACCHARYLTRANSFERASE SUBUNIT Ost4 AND ITS FUNCTIONALLY IMPORTANT MUTANT Ost4V23D

3.1 Introduction

About two-thirds of all the drug targets are membrane proteins. Therefore, membrane proteins are very important in the biomedical and biotechnology field. However, the number of membrane protein structures in the protein data bank (PDB) is significantly low due to various limiting factors and technical difficulties. One of the reasons for very few membrane protein structures in the PDB is the difficulty in the crystallization of these proteins. The membrane proteins are also difficult to purify to homogeneity due to proteolytic degradation and the presence of various impurities. Several efforts have been made to increase the yields from membrane protein expression in both prokaryotes and eukaryotic eukaryotes (*1, 2*). Additionally, the membrane proteins are comparatively more flexible than soluble proteins and have conformational heterogeneity and a high concentration of detergent is often used to extract and purify them which makes them difficult to crystallize (*3*). Therefore, solution NMR is the technique of choice for the determination of high-resolution structures of membrane proteins. Atomic resolution

structure of proteins by solution NMR requires several steps to be completed before we can determine the final refined structure. Such steps can be categorized as (a) preparation of ^{13}C and ^{15}N isotopically labeled protein samples, (b) acquisition of multidimensional homonuclear multinuclear to NMR data and NMR data processing, (c) backbone and sidechain resonance assignments followed by NOE assignment, (d) restraints generation such as NOE restraints, dihedral angle restraints, H-bond restraints, and (e) structure calculation, refinement and energy minimization using different computer programs.

Once all NMR data for a protein sample are collected and processed, the first step is to assign the backbone resonances by walking through the backbone sequence. This step follows the side – chain resonance assignment by using different TOCSY experiments. The final step of the resonance assignment is the NOE assignment. The assignment of NOE cross-peaks unambiguously is very important because NOE provides the structural constraints that are used for structure generation. Theoretically, the NOE cross-peaks can be assigned by using chemical shifts of backbone and side – chain resonances.

This chapter includes the backbone and side – chain resonances assignments of Ost4 and its functionally important mutant Ost4V23D.

3.2 Backbone assignments of Ost4 protein

3.2.1 Introduction

The initial step of the protein structure determination by NMR begins with the interpretation and assignment of NMR data for the protein. In the backbone assignment, the resonances of protein backbone nuclei such as amide H and the attached N, Cα, Cβ,

and Hα are assigned. Despite recent advances towards automated assignments, the assignments for membrane proteins by these programs are often not possible.

The backbone assignment of a protein can be accomplished by following the sequential assignment strategy. This strategy was developed by the Wüthrich group using a set of 2D NMR experiments on unlabeled protein samples about 35 years back (*4*). For larger globular proteins and membrane proteins, the sequential backbone assignment requires uniformly ^{13}C and ^{15}N-labeled protein samples to record various necessary and complimentary 3D NMR experiments (*5*). The most commonly used 3D NMR experiments for protein backbone assignment are HNCA, HN(CO)CA, HNCO, HN(CA)CO, CBCANH and CBCA(CO)NH. All of these 3D heteronuclear correlation experiments use one-bond $^{13}CO(i-1) - ^{15}N(i)$, $^{15}N(i)$-$^{13}C\alpha(i)$ and $^{13}C\alpha(i)$-$^{13}CO(i)$, as well as two-bond $13C\alpha(i-1)$-$^{15}N(i)$ scalar couplings. Therefore, the backbone resonances of its own (i), the preceding (i-1) residue or just residue (i) are correlated with the amide group of residues (i). And, the sequential assignment is accomplished and confirmed by connecting the resonances of one residue with those of its adjacent neighbor through a number of independent pathways (Cα, Cβ, and CO). The backbone resonances are correlated with amide groups because the amide groups are usually the best-resolved set of signals. Figure 3.1 shows the strategy of sequential backbone resonance assignment by linking one NH group next into a long chain and walking through the backbone of a protein by using HNCACB and CBCA(CO)NH pair of 3D experiments.

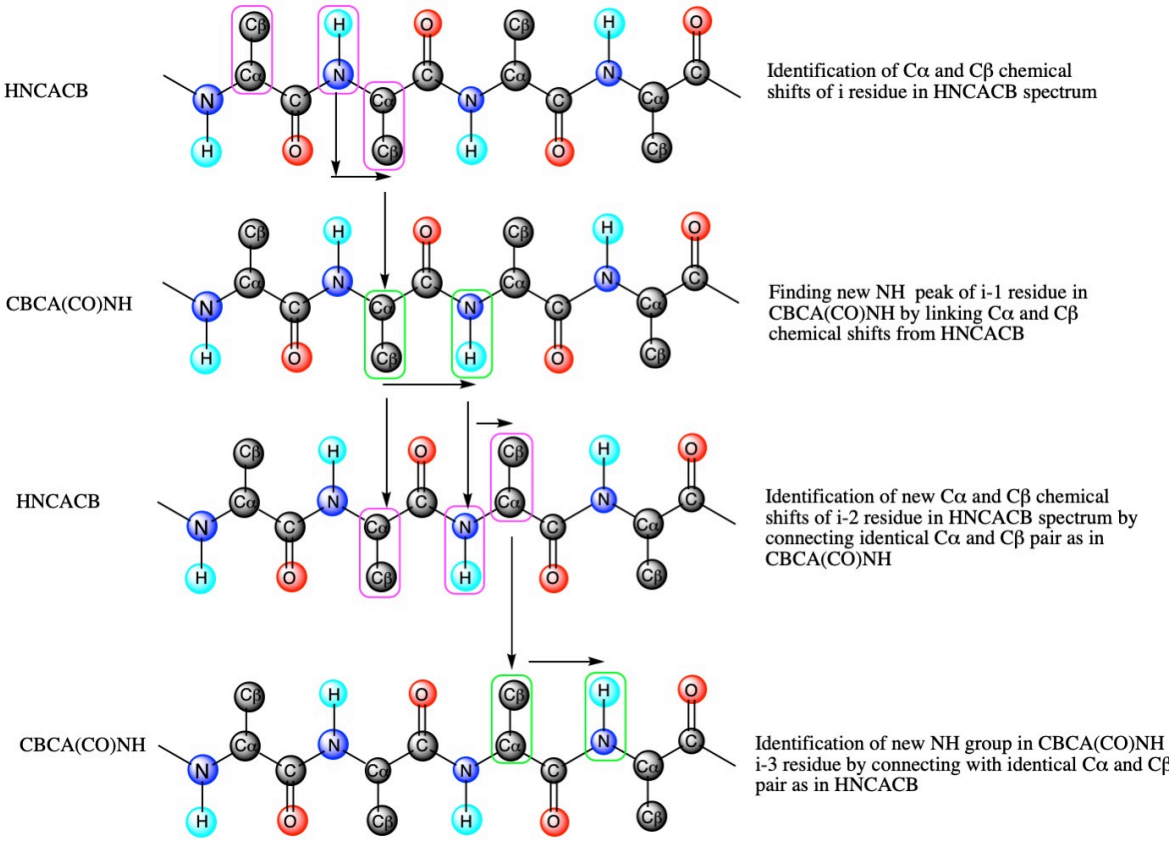


Figure 3.1: Schematic representation of the strategy for sequential backbone assignment by using 3D HNCACB and 3D CBCA(CO)NH pair of spectra.

In actual practice, the strip plot of HNCACB and CBCA(CO)NH spectra together will be as shown in Figure 3.2.

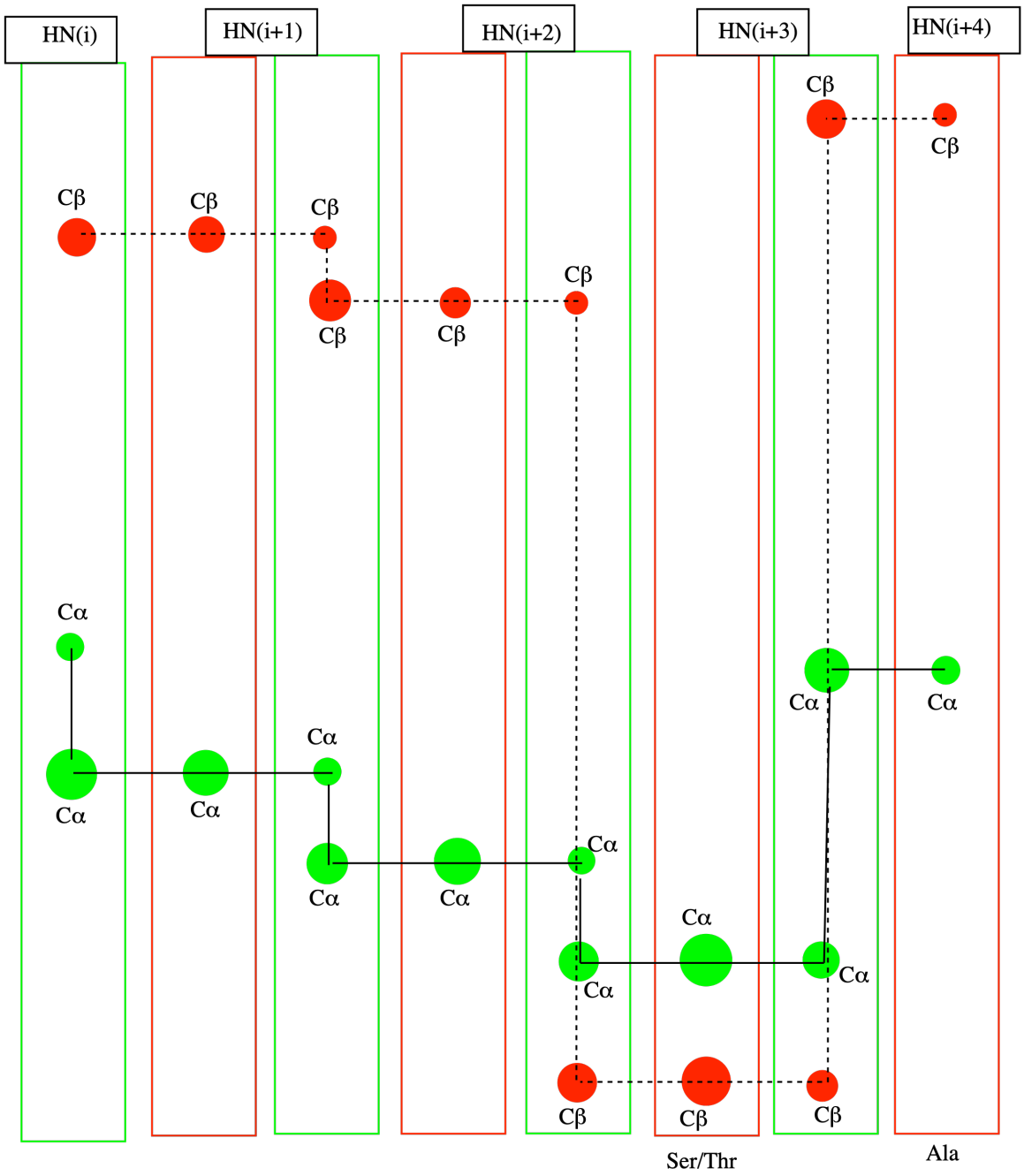


Figure 3.2: Schematic representation of a strip plot of HNCACB and CBCA(CO)NH pair of spectra. Cα (green) connectivity is shown with solid lines and Cβ (red) connectivity are shown with dotted lines.

Once the protein backbone resonances are assigned, the secondary structure of the protein can be calculated by using backbone chemical shift values by using TALOS+ and CSI (Chemical Shift Index) method developed by Wishart *et al*. (*6, 7*). The backbone chemical shift values of protein contain information about the secondary structure of proteins (*8-12*). From the CSI method, the secondary structure of proteins can be predicted with about 92% accuracy (*6*). However, the other programs such as TALOS+ (*13*), PSICSI (*14, 15*), CSI, and SSP (*16*) can also be used to derive the secondary structure information about the protein from the backbone chemical shift values.

In this section, the backbone resonance assignment of yeast oligosaccharyltransferase subunit Ost4 will be discussed.

3.2.2 Materials and methods

3.2.2.1 ^{13}C, ^{15}N – labeled protein overexpression and purification

Construction of GB1-OST4 has been reported previously (*17*). Ost4 was expressed and purified by following previously described methods (*17*). The details of production and purification of unlabeled and ^{15}N – labeled (single labeled) Ost4 protein has been described in Chapter 2. For backbone and side-chain assignment, ^{13}C, ^{15}N – labeled (double labeled) Ost4 protein was overexpressed and purified by following the same procedure as described in Chapter 2. This chapter describes overexpression and purification of ^{13}C, ^{15}N – labeled (double labeled) protein. Figure 3.3 demonstrates a summary of overall steps followed during the expression process of the double–labeled protein.

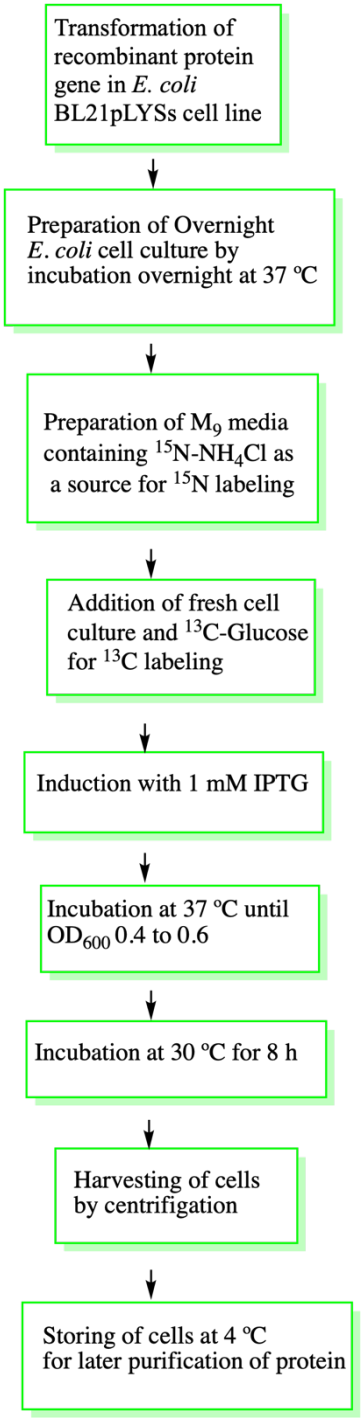


Figure 3.3: Schematic representation of steps followed for the expression of double labeled GB1 – Ost4 protein.

Briefly, GB1-Ost4 was transformed in *E. coli* BL21DE3pLysS cells (Stratagene). The overnight culture of each expression was diluted to an OD_{600} of 0.06 using fresh M_9 media supplemented with 4 g/L ^{13}C-glucose, 1.2 g/L ^{15}N-NH_4Cl, 1mg/L thiamine, 100 μg/mL ampicillin 2 mM $MgSO_4$, 50 μM $CaCl_2$, and 100 μM trace elements. The culture was then incubated at 37 ºC until OD_{600} reached between 0.4-0.6. Expression was induced at 30 ºC by the addition of 1mM isopropyl-β-D-thiogalactopyranoside (IPTG). After 8 hours, the cells were harvested by centrifugation at 9000 rpm. Approximately 2.0 g of cell pellet was obtained from 500 mL of cell culture. The cell pellets from a 500 mL of cell culture were resuspended in 15 mL of lysis buffer (50 mM sodium phosphate buffer (pH 7.4) 200 mM sodium chloride, 0.01% sodium azide). Then a 6-second sonication pulse followed by 1-minute incubation on ice water bath was performed to lyse the cells. After lysis by sonication, the cell debris was separated from the supernatant liquid by centrifugation for 20 minutes at 4 ºC. The purification of GB1 – Ost4 was done by using Ni – NTA affinity chromatography. For 4 mL of cell lysate, a column was packed with 4 mL of Ni – NTA resin. Prior to loading the cell lysate, the column was washed with nano pure water 10 times to remove the 20 % ethanol used for storage. The column was equilibrated sufficiently with an equilibration buffer. The cell lysate being in the same buffer as the equilibration buffer loaded onto a pre-equilibrated Ni-NTA column for purification taking advantage of the C-terminal 6X-His-tag. Finally, GB1 – Ost4 protein eluted under the flow of gravity by using elution buffer containing 500 mM imidazole. Completion of the elution was monitored by measuring the A280 of each elution. The elution was considered complete once the OD280 of the protein reached ~ 0.050 AU. The eluted protein was concentrated and the buffer exchanged with 1X cleavage buffer for the removal of the GB1

– tag from the protein. The removal of the GB1 tag from the protein was carried out by using thrombin cleavage as described in Chapter 2. A total of 21.5 mg of ^{13}C, ^{15}N double labeled Ost4 protein was obtained from 1 L of cell culture. The completion of the cleavage was monitored by SDS-PAGE gel electrophoresis.

3.2.2.2 NMR sample preparation

The NMR samples used for NMR experiments for sidechain and backbone assignments consisted of 400 μL of 0.400 – 0.750 mM of uniformly ^{13}C, ^{15}N labeled protein in 50 mM phosphate buffer (pH 6.5) containing 1 mM EDTA, 0.01% NaN_3, 100 mM DPC. The dried pellet of 2 – 4 mg pure protein was dissolved in 600 – 800 μL of NMR buffer containing 100 mM DPC micelles. The protein pellet was suspended in the buffer followed by water batch sonication until all the pellet dissolved in the buffer. Finally, the solution of Ost4 protein in NMR buffer containing 100 mM DPC micelles was centrifugated at 13200 rpm for 15 minutes to get rid of any undissolved impurities. Approximately 400 μL of supernatant liquid containing reconstituted Ost4 in DPC micelle was mixed with 10% D_2O (used as a lock solvent) and loaded in a Shigemi tube for NMR data collection.

3.2.2.3 NMR data collection

All NMR data were acquired using either a Bruker 800 MHz (Avance II) spectrometer having a 5 mm triple resonance pulsed-field gradient TCI cryoprobe at the National High Field Magnetic Laboratory, Tallahassee, Florida, or Varian Inova 600 MHz or 900 MHz spectrometers equipped with cryoprobes at the Department of Pharmacology, University

of Colorado School of Medicine, Colorado, or a Bruker 600 MHz (Avance III) or a Varian Inova 900 MHz spectrometers with cryoprobes at the University of Minnesota NMR center. The following NMR experiments were performed: 2D [^{1}H,^{15}N]- HSQC, 2D [^{1}H, ^{13}C]-HSQC, 3D HNCACB, 3D CBCACONH, 3D HNCA, 3D HN(CO)CA, 3D HNHAHB, 3D HBHA(CO)NH, 3D ^{13}C-edited (H)CCH-TOCSY, 3D (H)CCONH-TOCSY, and 3D ^{15}N-edited HSQC-TOCSY experiments were performed for backbone and sidechain assignment of Ost4. All the NMR data were acquired at 35 ºC. Additionally, 3D ^{15}N-edited Trosy-HSQC-NOESY and ^{15}N-edited HSQC-NOESY were collected for the structure calculation of this protein. The list of experiments with the important parameters used for data collection are listed in the Table 3.1.

Table 3.1: Summary of experiments with important parameters used for the data collection for Ost4 protein

Experiment	Dimension	Isotope	Carrier Frequency (ppm)	Spectral width (ppm)	Number of data points	Number of scans
2D [^{1}H,^{15}N]-HSQC	1	1H	4.70	15.00	2048	32
	2	15N	117.00	36.00	512	
2D [^{1}H, ^{13}C]-HSQC	1	1H	4.68	16.08	2048	32
	2	13C	41.65	70.00	256	
3D HNCACB	1	1H	4.68	15.00	1536	32
	2	15N	117.81	36.23	100	
	3	13C	47.70	70.00	60	
3D CBCACONH	1	1H	4.68	15.00	1534	32
	2	15N	117.81	36.23	64	
	3	13C	47.70	70.00	132	
3D HNCA	1	1H	4.68	15.00	1534	32
	2	15N	117.81	36.23	68	
	3	13C	47.70	75.00	192	
	1	1H	4.68	15.00	1534	

3D HN(CO)CA	2	15N	117.81	36.23	100	32
	3	13C	47.70	70.00	60	
3D HNHAHB	1	1H	4.70	11.71	2048	8
	2	15N	117.00	24.00	82	
	3	1H	4.70	5.85	256	
3D HBHA(CO)NH	1	1H	4.70	11.71	2048	8
	2	15N	117.00	24.00	82	
	3	1H	4.70	5.85	128	
3D ^{13}C-edited (H)CCH-TOCSY	1	1H	4.70	11.71	2048	2
	2	13C	38.00	76.00	140	
	3	13C	38.00	76.00	200	
3D (H)CCONH-TOCSY	1	1H	4.70	11.71	2048	8
	2	15N	117.81	36.00	62	
	3	13C	40.00	76.00	200	
3D ^{15}N-edited HSQC-TOCSY 85 ms	1	1H	4.70	13.02	2048	128
	2	15N	117.49	34.00	32	
	3	1H	4.70	13.02	128	
15N – edited Trosy HSQC – NOESY 120 ms	1	1H	4.70	10.06	2048	88
	2	15N	115.00	30.00	44	
	3	1H	4.70	10.06	128	
15N – edited HSQC – NOESY 120 ms	1	1H	4.70	12.00	2048	16
	2	15N	117.98	36.00	64	
	3	1H	4.70	12.00	180	

3.2.2.4 Data processing

The NMR data were processed by using NMR pipe (*18*). During the data processing, the residual water signal was minimized by a time domain deconvolution process. Prior to the

application of apodization with a cosine-bell window function and Fourier transformation, the time domain data in the ^{15}N dimension of all the spectra were doubled using mirror image linear prediction. Similarly, the time domain ^{13}C dimension of all of the spectra were doubled using mirror-image linear prediction and then apodised with cosine squared window functions. The linear prediction was performed only after all of the other spectral dimensions were Fourier transformed. The spectral width of the transformed data sets was decreased to include only the regions of interest. Finally, the data were converted to a SPARKY readable file and analyzed using the NMRFAM-SPARKY program (*19*).

3.2.3 Results and discussion

3.2.3.1 Backbone assignment of Ost4 protein

The backbone resonances of the protein were assigned sequentially by using HNCAB and CBCA(CO)NH 3D NMR experiments. HNCACB and CBCA(CO)NH experiments together provided the connectivity among the Cα and Cβ atoms of both (i) and (i-1) residues of the peptide chain (Figure 3.4 and 3.5). By using this pair of 3D NMR experiments, it was possible to walk through the backbone sequence of the protein. In addition, a HNCA 3D experiment was used to resolve any ambiguity that arose from overlap in the Cα of the (i) and (i-1) residues. This experiment provided connectivity only between Cα atoms of both (i) and (i-1) residues (Figure 3.6). The CBCA(CO)NH experiment provided only Cα and Cβ atoms of the (i-1) residues. This helped to reconfirm the assignments of the HNCACB experiment. The peak intensity of the (i-1) residue of the HNCACB spectrum was almost the same as that of the (i-1) residue of CBCA(CO)NH

spectrum (Figure 3.4 and 3.5). However, more often the peak intensity of the (i-1) residue of the HNCACB experiments is less than that of the (i-1) residue of CBCA(CO)NH

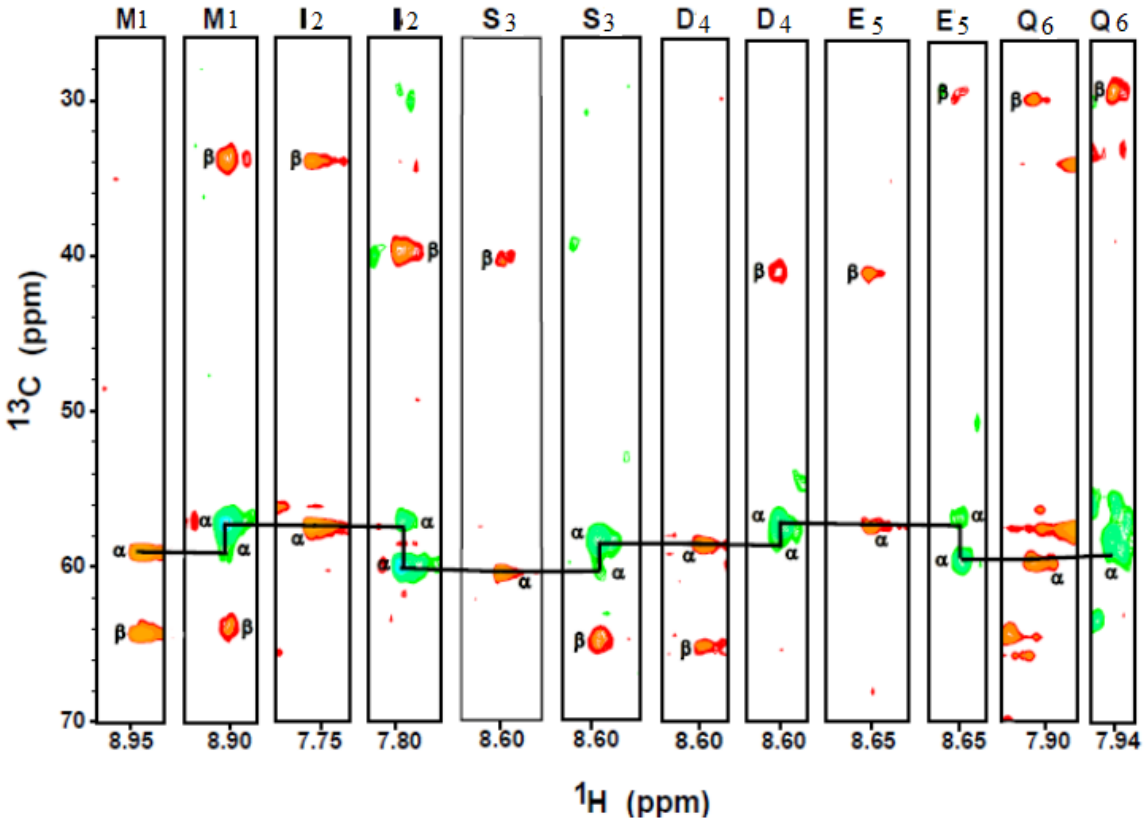


Figure 3.4: Strip plot of HNCACB and CBCACONH for sequential backbone assignment for residues M1 to Q6. Only Cα of the residues was connected with black lines to show the connectivity. Positive signals are shown in green and negative signals are shown in red color. The Cα was made positive and Cβ was made negative during data processing.

experiment due to the difference in the transfer of magnetization in these two different experiments. This feature provides an added confidence for the unambiguous identification of Cα and Cβ atoms of (i) and (i-1) residues of HNCACB. All these assignments were accomplished by continuously referring to the table of statistics calculated for all chemical

shifts from atoms of the 20 common amino acids (biological magnetic resonance data bank, BMRB, http://www.bmrb.wisc.edu/) which is always updated. The backbone chemical shifts were assigned using HNCACB and CBCA(CO)NH pair using a continuous connectivity strategy. Figure 3.5 demonstrates the Cα connectivity of residues Leu[10] – Asp[29] in HNCACB and CBCA(CO)NH pair.

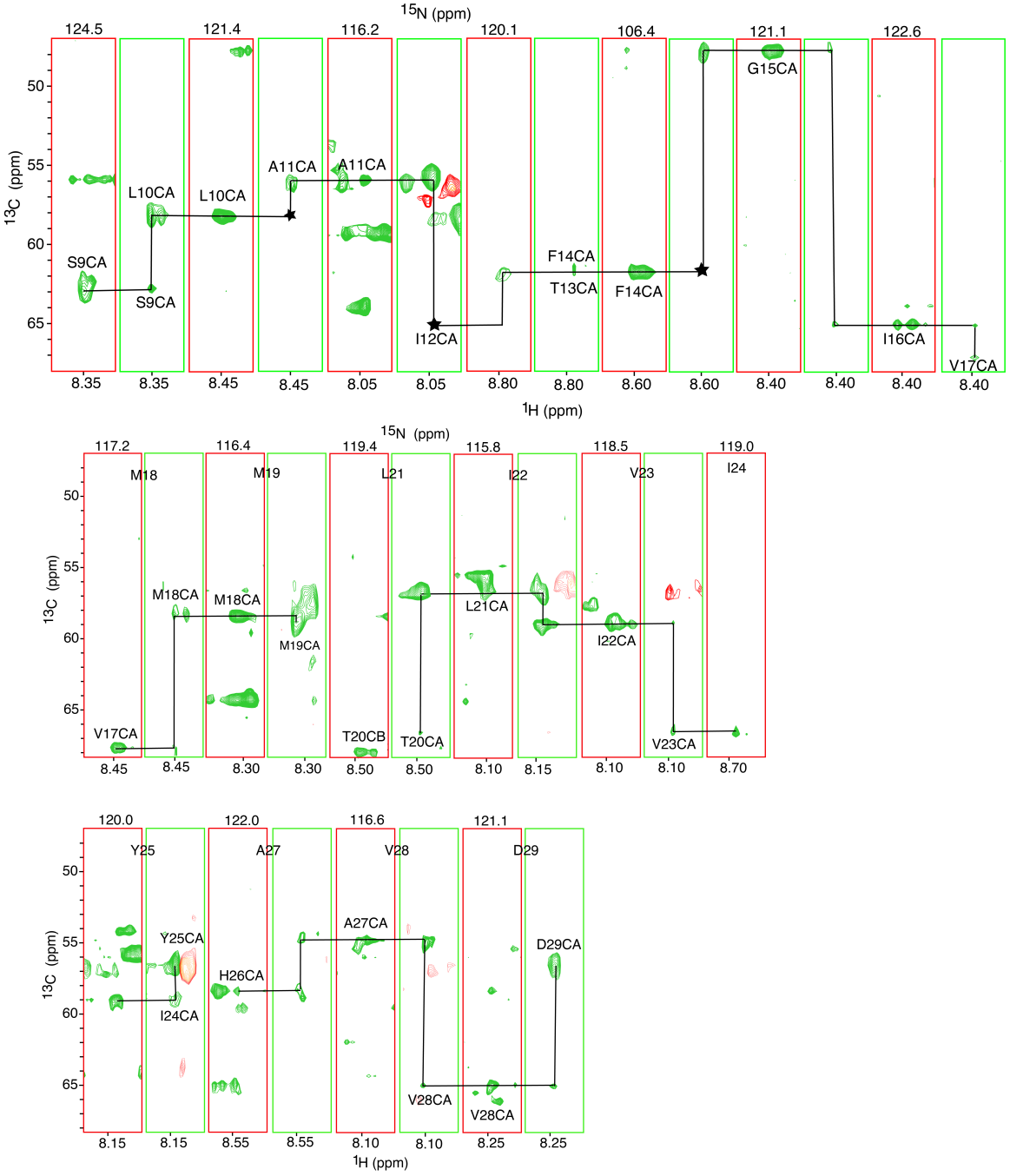


Figure 3.5: Strip plot of HNCACB (green) and CBCA(CO)NH (red) for sequential backbone assignment for residues Leu[10] – Asp[29]. The figure demonstrates the Cα connectivity of the residues present in the transmembrane domain of the protein. The resonances that exist at lower contour level are indicated by * in the plots.

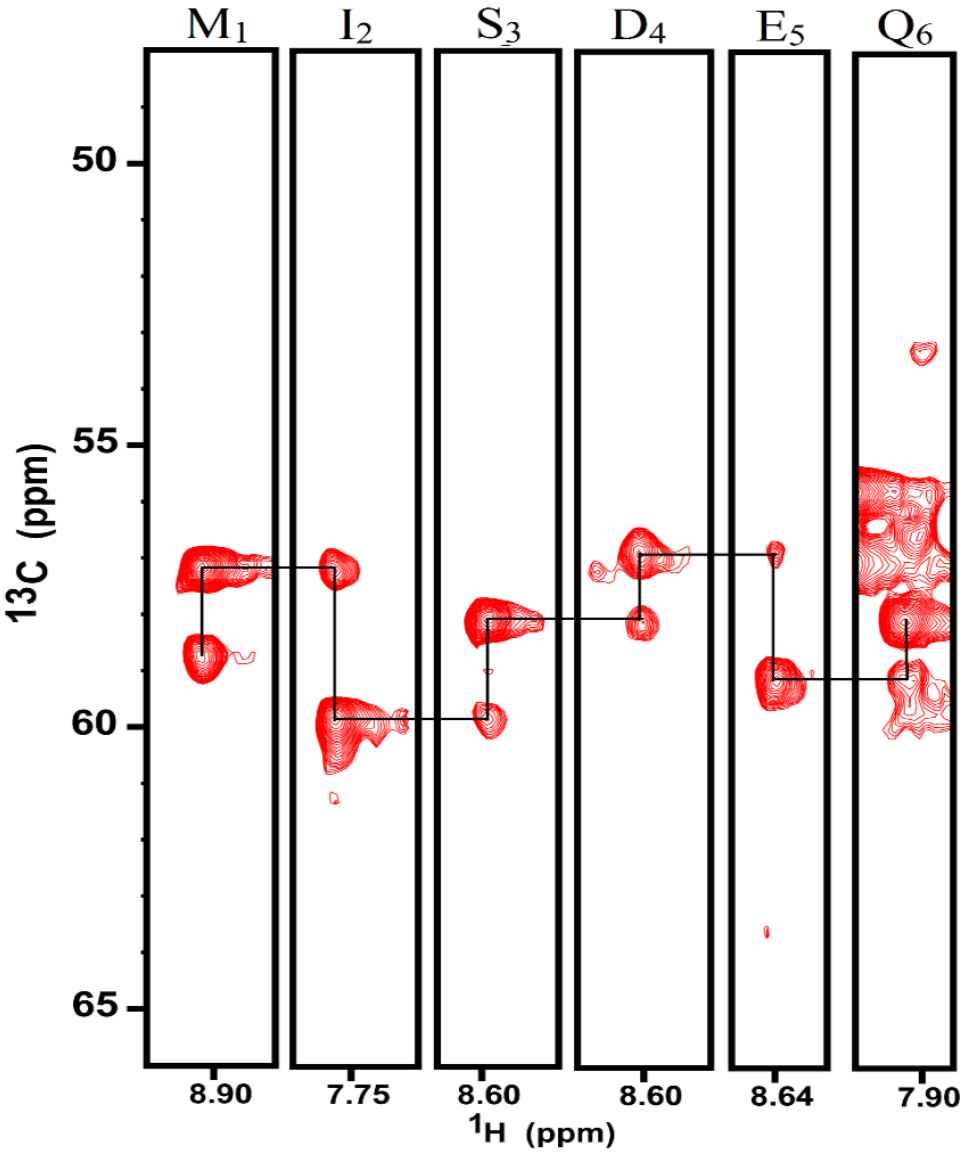


Figure 3.6: Strip plot of the HNCA 3D experiment for sequential backbone assignment for residues Met[1] – Gln[6].

Once the sequential assignment was completed, the [^{1}H, ^{15}N]-HSQC spectrum was assigned by using the information from the sequential assignment (Figure 3.7). Well-resolved resonances in the 2D [^{1}H, ^{15}N]- HSQC of the protein indicated that the proteins were well-folded in 100 mM DPC micelles (Figure 3.7). The ^{1}H, ^{15}N, and ^{13}C resonance assignments for Ost4 have been deposited into BMRB (http://www.bmrb.wisc.edu) with accession number 50159. The protein is composed of 45 residues including nine additional residues in the C-terminus. Besides the 6X-His- tag (His^{40}-His^{45}), both the proteins contain three additional residues: R37, L38, and E39.

The assignment of the backbone resonances (^{1}H, ^{15}N, $^{13}C_{\alpha}$, and $^{13}C_{\beta}$) of Ost4 was completed for all residues except ^{1}H and ^{15}N of T13, T20, and H26 in addition to C_{β} for I12 in the primary sequence (Met^{1}-Asn^{36}). In the [^{1}H, ^{15}N]-HSQC spectrum, the resonances ^{1}H and ^{15}N of T13, T20, and H26 were absent due to the exchange of the free amide protons with the deuterated solvent (10 % D_2O. The only proline residue at position 34 is present in both the cis and trans forms. Thus, two sets of resonances for residues S33 and K35 were observed for Ost4 and Ost4V23D in the 2D [^{1}H, ^{15}N]-HSQC spectra (Figure 3.7).

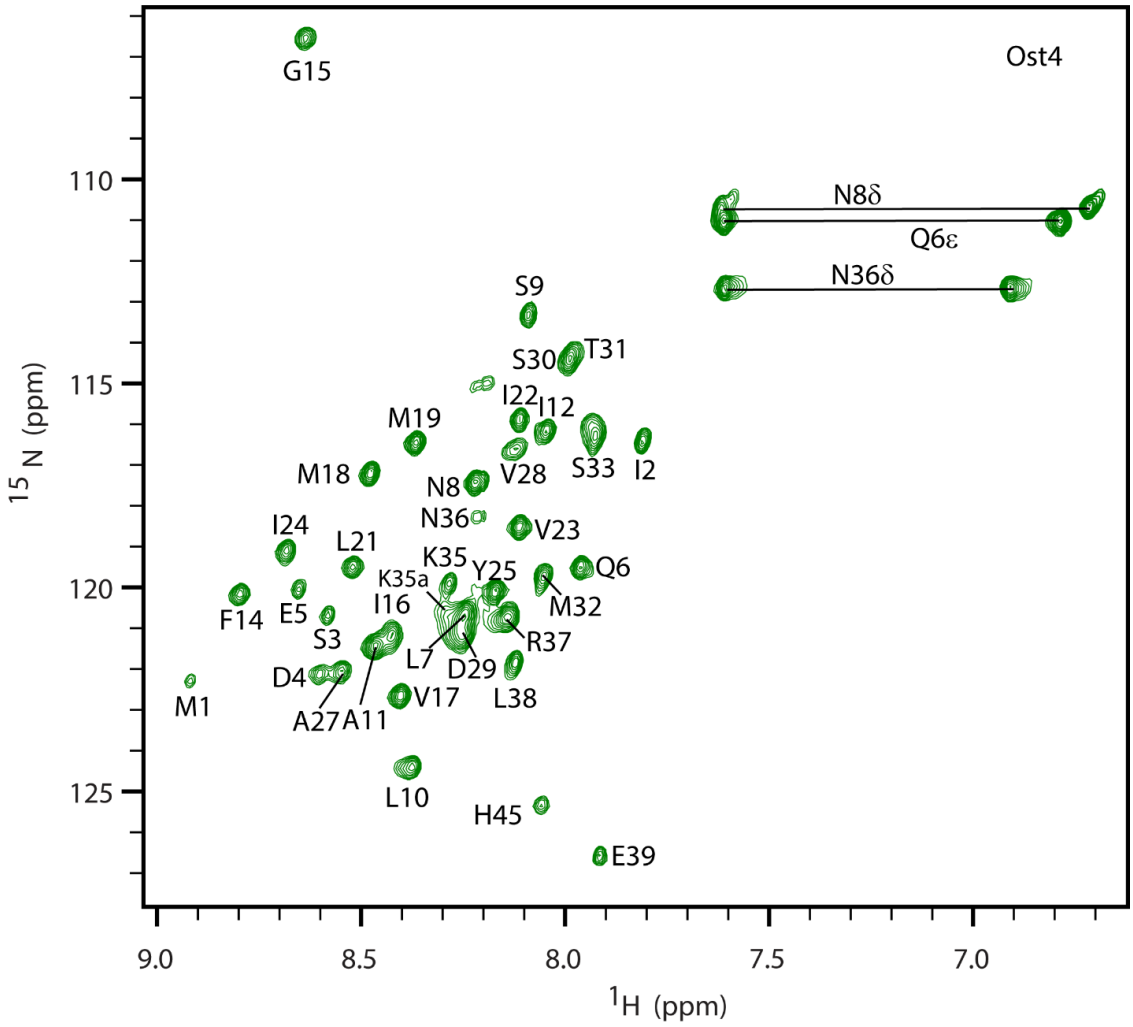


Figure 3.7: 2D [^{1}H, ^{15}N]-HSQC spectra of Ost4 at pH 6.5 at 35 ºC. The proteins contain 45 amino acid residues including nine additional residues in the C-terminal 6X-His-tag. Only the terminal histidine (H45) of the hexahistidine tag was assigned. The residues from Met^{1} - Asn^{36} belong to Ost4 and Ost4V23D. The residues that do not belong to the proteins are R37, L38, E39, and His^{40}-His^{45} in the C-terminus.

3.2.3.2 Chemical shift index (CSI) and the secondary structure of the protein

The secondary structures of Ost4 were calculated using backbone chemical shifts by using TALOS+ (*13*) and SSP programs (*16*). The deviation of the $^{13}C\alpha$ and $^{13}C\beta$ chemical shift from their mean random coil values was evaluated and the secondary structure was determined by using the secondary structure propensity (SSP) program (*16*). As shown in Figure 3.8, $\Delta\delta C_\alpha$ and $\Delta\delta C_\alpha$-$\Delta\delta C_\beta$ were plotted against the protein sequence. Both programs predicted a single helix encompassing residues Asp^4 - Met^{32} for Ost4 (Figure 3.8).

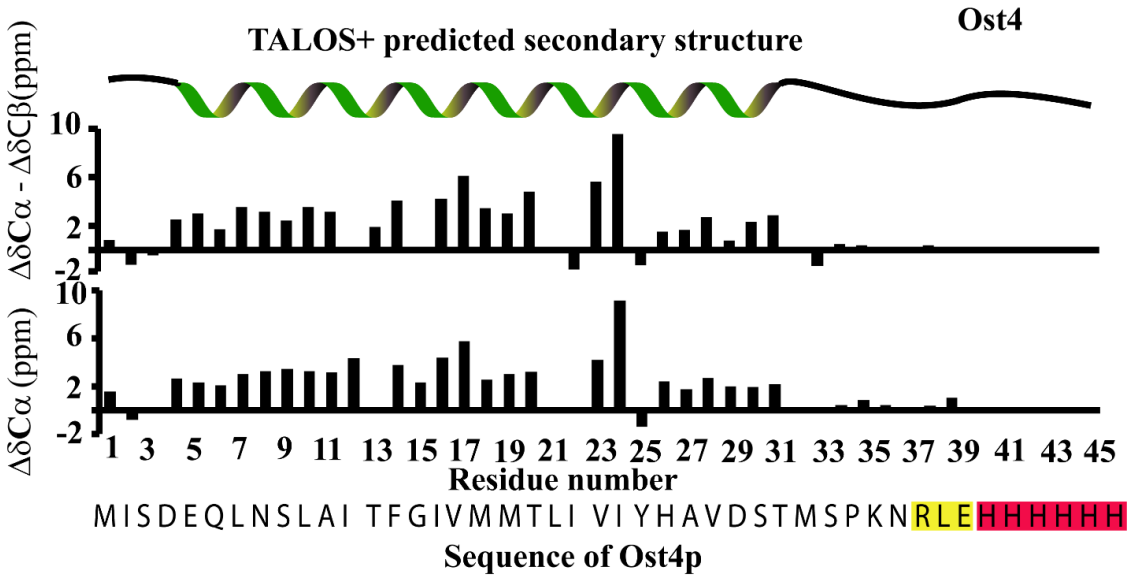


Figure 3.8: TALOS+ predicted secondary structure, $\Delta\delta C_\alpha$–$\Delta\delta C_\beta$, and $\Delta\delta C_\alpha$ secondary chemical shifts of Ost4. The TALOS+ predicted secondary structures of each protein are shown on the top. Positive values in secondary chemical shifts indicate α–structure propensity and negative values indicate β–structure or random coil propensity. The secondary chemical shifts were derived by using the SSP program (*16*). The residues R37,

L38, and E39 (highlighted in yellow) along with the C-terminal 6xHis-tag tag (highlighted in red) do not belong to Ost4.

3.2.3.3 Side chain assignment of the Ost4

Once the backbone assignments are complete, the next and essential step is the assignment of the side chain resonances of the amino acid residues in the protein. For the side chain assignment of the resonance for which backbone amide proton and nitrogen atoms are already assigned, we can use various TOCSY experiments such as ^{15}N HSQC TOCSY, HCCH TOCSY, HC(CO)NH TOCSY experiments.

The side chain protons of the (i) residue were assigned by using the ^{15}N HSQC TOCSY experiment. In these experiments, the side chain protons of the (i-1) residue could appear with a negative signal. HC(CO)NH TOCSY experiment provides a side chain proton of the (i-1) residue. The side chain carbon atoms can be found by using the (H)C(CO)NH TOCSY experiment (Figure 3.9). This experiment provides side-chain carbon atoms that belong to the (i-1) residues as shown in Figure 3.9.

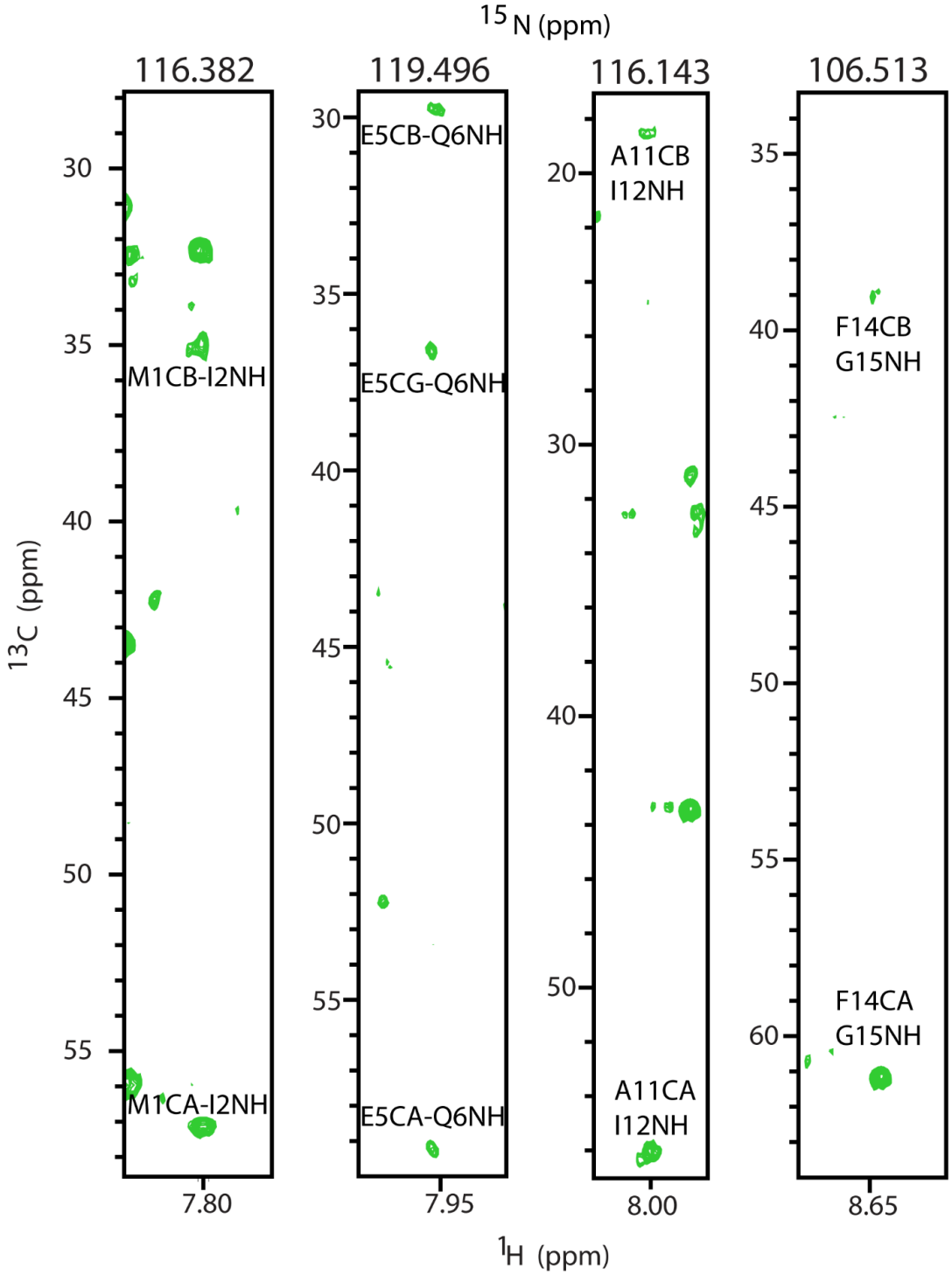


Figure 3.9: Strips of the (H)C(CO)NH TOCSY experiment for residues I2, Q6, I12, and G15. This experiment provides side chain carbon atoms of preceding residues.

For large proteins, it is better to use the [^{1}H, ^{15}N] HSQC TOCSY and H(CCO)NH TOCSY experiments together to assign the side chain ^{1}H resonances. However, for small proteins, the side chain ^{1}H resonances of the backbone residues can be accomplished by using either of these two experiments. Figure 3.10 provides an example of side chain ^{1}H resonance assignments for I2, E5, and V23 residues by using [^{1}H, ^{15}N] HSQC TOCSY. Another experiment to assign the side chain is HCCH TOCSY. Figure 3.11 shows strips for residues I2 and V28 as an example of HCCH TOCSY. This HCCH TOCSY experiment is used to assign the all the side chain proton from ^{13}C resonances. Apart from the HCCH TOCSY, ^{15}N-edited HSQC-NOESY or ^{13}C-edited HSQC-NOESY experiments can also be used to assign the degenerate protons occasionally. Additionally, [^{1}H, ^{13}C]-HSQC was used to assign and confirm the ^{1}H and ^{13}C side chain resonances of the protein residues. By using the above-mentioned experiments, it was possible to assign 91% of all the side chain resonances.

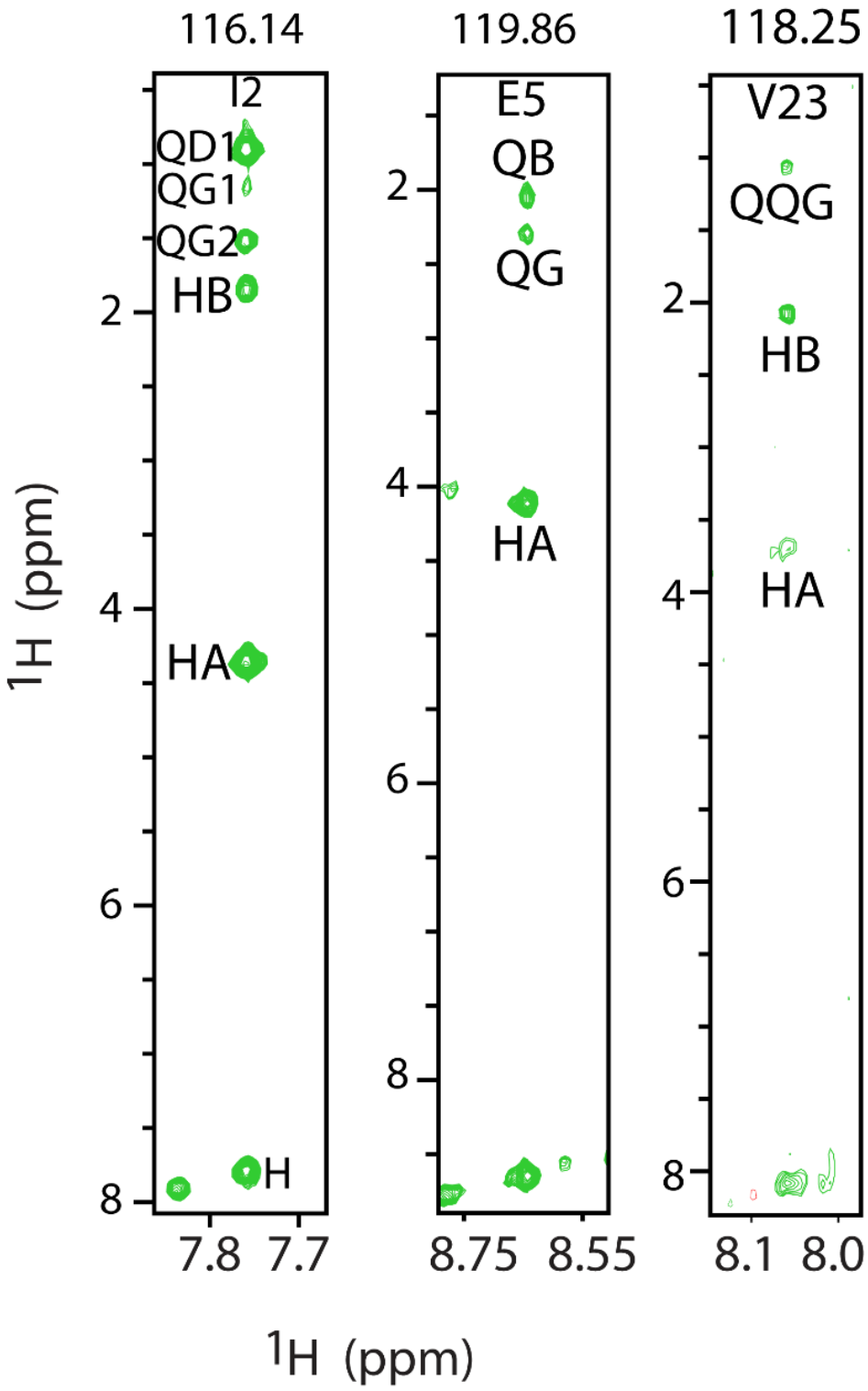


Figure 3.10: Strips of I2, E5, and V23 residues from the HCCH TOCSY experiment. This provides a side chain proton of the (i) residue.

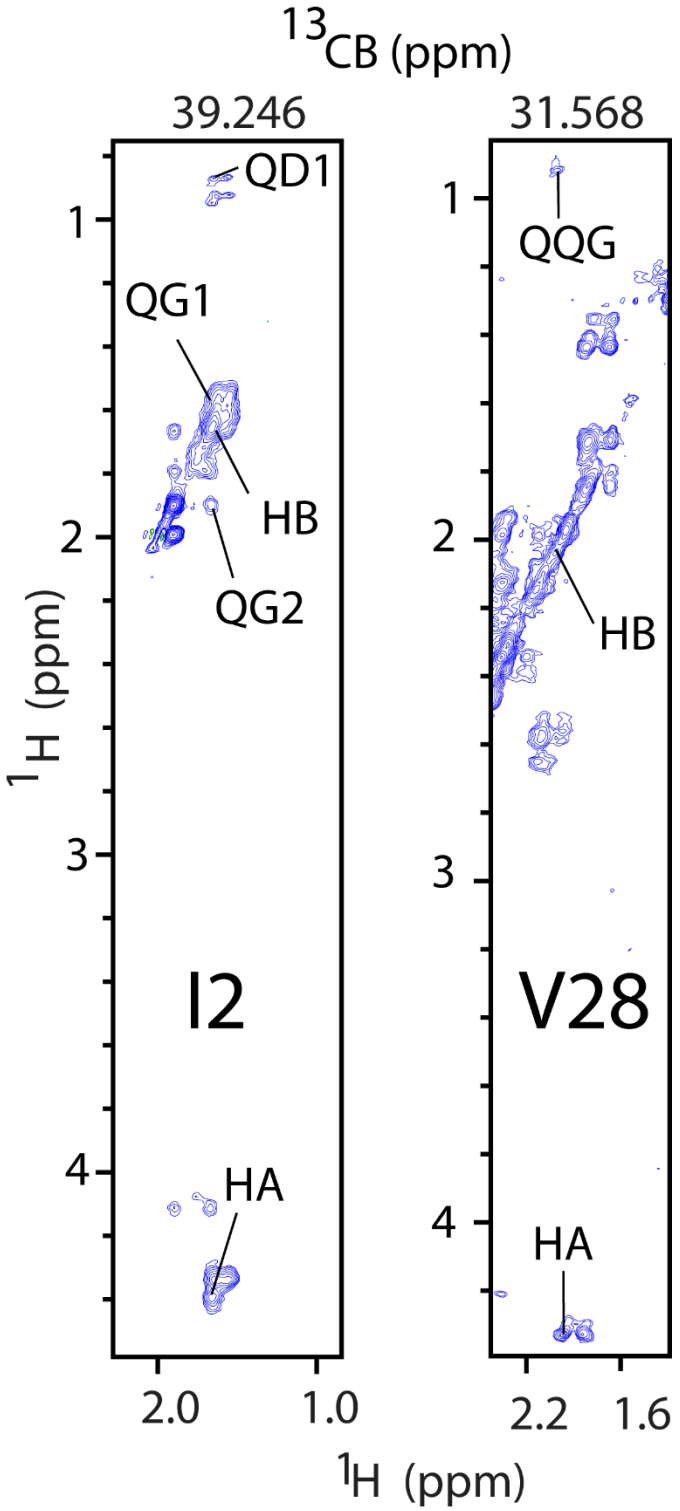


Figure 3.11: Strips of I2 and V28 residues from the HCCH TOCSY experiment. This provides a side chain proton of the (i) residue.

3.3 Backbone assignment of Ost4V23D protein

In this section, the backbone resonance assignment of yeast oligosaccharyltransferase subunit Ost4V23D will be reported.

3.3.1 Mutagenesis

The OST4 gene was previously cloned into a GEV 2 vector (*17*). The mutagenesis of the OST4 gene to obtain the OST4V23D gene has been previously described in Chapter 2.

3.3.2 GB1- Ost4V23D expression

Unlabeled and ^{15}N labeled GB1-Ost4V23D mutant proteins were produced in *E. coli* BL21DE3pLysS cells (Strategene) and details have been discussed in Chapter 2. The production and purification of ^{13}C, ^{15}N double labeled Ost4V23D will be explained in this chapter. Expression of the Ost4V23D in the GEV2 vector was under the control of an IPTG (isopropyl β-D-thiogalactopyranoside) inducible promoter. Briefly, the overnight starter culture was diluted to an OD_{600} of 0.06 in fresh LB medium containing 100 μg/mL of ampicillin and grown at 37 ^{o}C to an OD_{600} of 0.4-0.6. At that point, the temperature was reduced to 30 ^{o}C and protein production was induced by the addition of IPTG to a final concentration of 1 mM. After 4 hours, the cells were harvested by incubating on ice for 30 minutes followed by centrifugation at 9000 rpm for 20 min at 4 ^{o}C. The cell pellet was frozen at -20 ^{o}C until needed. For NMR studies, ^{15}N uniformly labeled protein was produced by using M_9 minimal media containing ^{15}N ammonium chloride (Cambridge Isotope Laboratories, MA). The cell culture was grown for 8 hours after induction with

IPTG and the cells were harvested in the same fashion as unlabeled protein. The protein expression was verified by Tris-Tricine gel electrophoresis.

3.3.3 ^{13}C, ^{15}N labeled GB1- Ost4V23D purification

The cell pellet from ^{13}C, ^{15}N labeled expression was freeze-thawed six times, by freezing it in liquid nitrogen for 5 minutes and thawing on ice. The cell pellet was suspended in lysis buffer (50 mM sodium phosphate buffer, pH 6.5, 200 mM sodium chloride, 0.01% sodium azide) and sonicated eight times with a 6-sec pulse followed by incubating 1 minute on the ice water bath. The cell debris was separated from cell lysate by centrifuging at 12000 rpm for 20 minutes at 4 ºC. The cell lysate containing GB1-Ost4V23D was stored at 4 °C pending further purification.

Ni-NTA affinity chromatography was used for the purification of the GB1-Ost4V23D protein. The protein was purified as follows under gravity. The column was packed with Ni-NTA resin (Thermo Fisher Scientific) and washed with nano-pure water several times. The column was equilibrated with binding buffer (50 mM sodium phosphate buffer, pH 6.5, 200 mM sodium chloride, 0.01% sodium azide and 10 mM imidazole). Cell lysate containing 10 mM imidazole was loaded to the column. The binding of the protein to the resin was continued overnight in a shaker.

The unbound proteins in the lysate were removed by draining the column. The impurities were removed with wash buffer (50 mM sodium phosphate buffer, pH 6.5, 200 mM sodium chloride, 0.01% sodium azide and 30 mM imidazole) several times. The bound protein was

eluted with elution buffer (50 mM sodium phosphate buffer, pH 6.5, 200 mM sodium chloride, 0.01% sodium azide and 500 mM imidazole).

3.3.4 Cleavage of GB1 tag from Ost4V23D protein

The pure ^{13}C, ^{15}N labeled GB1-Ost4V23D protein was concentrated and buffer exchanged with cleavage buffer (25 mM Tris-HCl buffer, pH 7.4, 150 mM NaCl and 2.5 mM $CaCl_2$). The cleavage of GB1-Ost4V23D was carried out as reported previously except for minor changes in the buffer conditions. Completion of the cleavage reaction was monitored by SDS PAGE after the completion of the reaction. Soluble GB1 was removed from insoluble Ost4V23D protein by centrifugation of the reaction mixture at 13200 rpm at room temperature. The protein pellet was washed with nano-pure water to remove any contaminating GB1 protein. A total of 9.2 mg of pure ^{13}C, ^{15}N labeled Ost4V23D protein was obtained from 500 mL of cell culture. Pure Ost4V23D protein containing the hexahistidine tag was stored for later use.

3.3.5 NMR spectroscopy

For the NMR data collection, the NMR sample of Ost4V23D protein was prepared by dissolving 2 – 4 mg of dry pellet into 600 – 800 μL of NMR buffer (50 mM phosphate buffer, pH 6.5 containing 1 mM EDTA, 0.01% NaN_3) containing 100 mM deuterated DPC micelles. This yielded an NMR sample having concentration of 400 – 750 μM. Approximately 400 μL of NMR sample was transferred into a Shigemi tube for the NMR data collection. All NMR data used for backbone and side chain assignments of Ost4V23D

protein were collected using a Bruker 800 MHz (Avance II) spectrometer equipped with a 5 mm triple resonance pulsed-field gradient TCI cryoprobe at the National High Field Magnetic Laboratory, Tallahassee, Florida. Some of the data such as 3D H(CCO)NH-TOCSY and ^{13}C-edited HCCH-TOCSY were acquired using either a Varian Inova 600 MHz or 900 MHz spectrometer equipped with cold probes at the Department of Pharmacology, University of Colorado School of Medicine, Colorado, or a Bruker 600MHz (Avance III). Similar to the NMR experiments acquired for Ost4 protein, 2D [^{1}H,^{15}N]- HSQC, 2D [^{1}H, ^{13}C]-HSQC, 3D HNCACB, 3D CBCACONH, 3D HNCA, 3D ^{13}C-edited HCCH-TOCSY, 3D H(CCO)NH-TOCSY, and 3D ^{15}N-edited HSQC-TOCSY experiments were performed for backbone and side chain assignment of Ost4V23D at 35 ºC. Additionally, 3D ^{15}N-edited HSQC-NOESY experiments were acquired for the structure calculation of this protein. The list of experiments with the important parameters are listed in Table 3.2.

Table 3.2: Summary of experiments with important parameters used for the data collection for Ost4V23D protein

Experiment	Dimension	Isotope	Carrier Frequency (ppm)	Spectral width (ppm)	Number of data points	Number of scans
2D [^{1}H,^{15}N]-HSQC	1	1H	4.70	15.00	2048	32
	2	15N	117.00	36.00	512	
2D [^{1}H, ^{13}C]-HSQC	1	1H	4.68	16.08	1314	144
	2	13C	75.32	48.35	512	
3D HNCACB	1	1H	4.68	12.00	1024	64
	2	15N	117.81	36.20	64	

	3	13C	47.70	70.00	160	
3D CBCACONH	1	1H	4.68	15.00	1534	88
	2	15N	117.81	36.23	64	
	3	13C	47.70	70.00	132	
3D HNCA	1	1H	4.68	15.00	1536	32
	2	15N	117.81	36.23	64	
	3	13C	51.32	75.00	80	
3D HN(CO)CA	1	1H	4.68	15.00	1534	32
	2	15N	117.81	36.23	100	
	3	13C	47.70	70.00	60	
3D ^{13}C-edited HCCH-TOCSY	1	1H	4.68	13.62	1786	16
	2	1H	4.68	7.34	4	
	3	13C	35.02	80.02	1217	
3D H(CCO)NH-TOCSY	1	1H	4.68	15.61	2048	24
	2	15N	117.02	25.23	2	
	3	1H	4.68	11.56	2448	
3D ^{15}N-edited HSQC-TOCSY	1	1H	4.70	13.02	2048	128
	2	15N	117.49	34.00	32	
85 ms	3	1H	4.70	13.02	128	
15N – edited HSQC – NOESY	1	1H	4.68	15.61	2048	24
	2	15N	117.02	25.23	1400	
90 ms	3	1H	4.68	11.48	4	
	1	1H	4.68	15.61	2048	

15N – edited HSQC – NOESY	2	15N	117.03	25.23	1678	24
120 ms	3	1H	4.68	11.48	4	

3.3.6 Data processing

The data processing was performed by using the same procedures as described for Ost4 protein.

3.3.7 Results and discussions

3.3.7.1 Backbone assignment of Ost4V23D protein

Similar to that of Ost4, the backbone resonances of the protein were assigned sequentially by using HNCAB and CBCA(CO)NH 3D NMR experiments. The HNCACB 3D experiment provided the connectivity among the Cα and Cβ atoms of both the (i) and (i-1) residues of the peptide chain (Figure 3.12). By using this pair of 3D NMR experiments, it was possible to walk through the backbone sequence of the protein. In addition, a 3D HNCA experiment was used to resolve any ambiguity that arises from overlap in the Cα of the (i) and (i-1) residues. This experiment provides connectivity only between Cα atoms of both the (i) and (i-1) residues. The CBCA(CO)NH experiment provides only Cα and Cβ atoms of the (i-1) residues. This helped to reconfirm the assignments from the HNCACB experiment. The peak intensity of the (i-1) residue of the HNCACB spectrum is almost the same as that of the (i-1) residue of the CBCA(CO)NH spectrum.

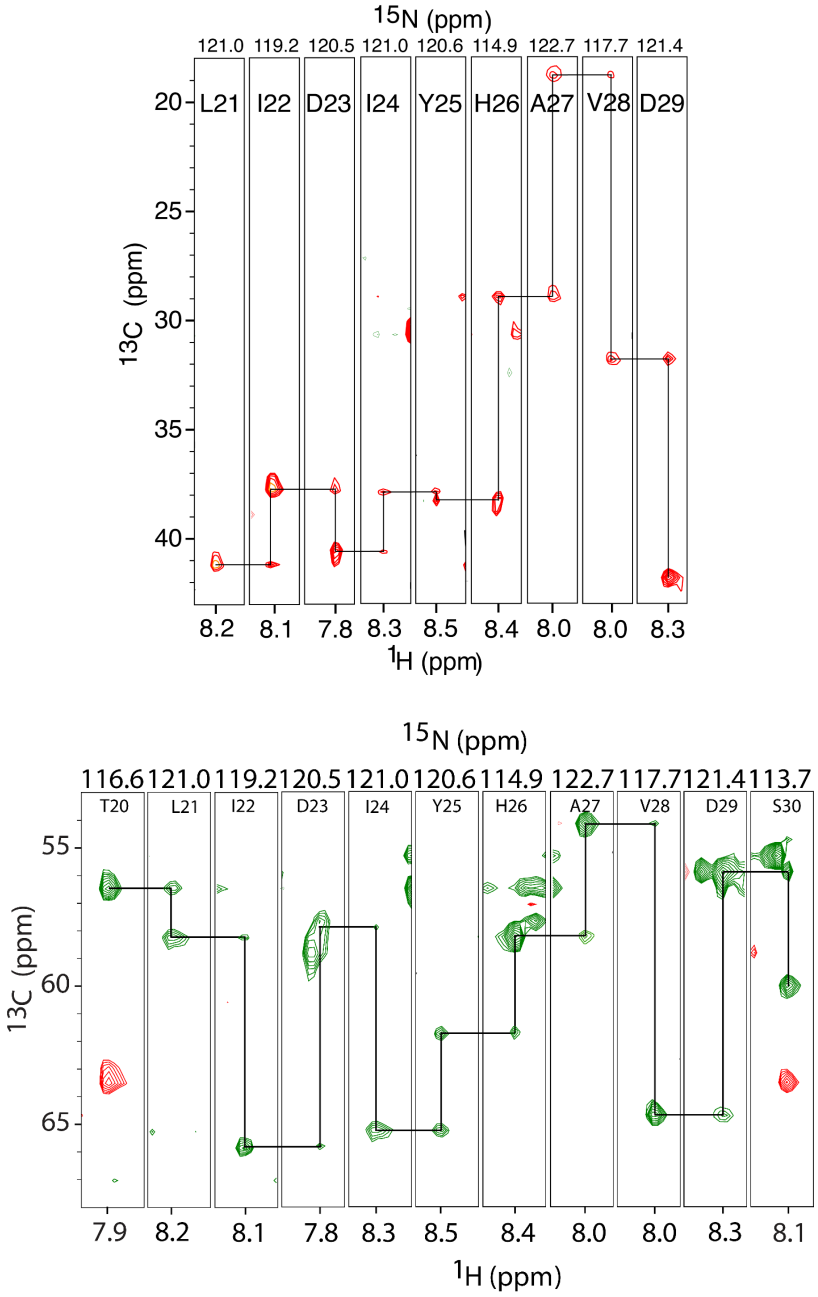


Figure 3.12: Strip plot of the HNCACB for sequential backbone assignment for residues Thr^{20} to Ser^{30}. The Cα (bottom) and Cβ (top) of the residues are connected with black lines to show the connectivity. Positive signals are shown in green and negative signals

are shown in red. The Cα was made positive and Cβ was made negative during data processing.

All these assignments were accomplished by continuously referring to the table of statistics calculated for all chemical shifts of atoms from the 20 common amino acids (biological magnetic resonance data bank, BMRB, http://www.bmrb.wisc.edu/) which is always updated.

Once the sequential assignment was completed, the [^{1}H, ^{15}N]-HSQC spectrum was assigned easily (Figure 3.13). Well-resolved resonances in 2D [^{1}H, ^{15}N]- HSQC of the protein indicate that the proteins are well-folded in 100 mM DPC micelles (Figure 3.13). The ^{1}H, ^{15}N and ^{13}C resonance assignments for Ost4 has been deposited into BMRB (http://www.bmrb.wisc.edu) with accession number 50160. The protein is composed of 45 residues including nine additional residues in the C-terminus. Besides the 6xHis- tag (H40-H45), both the proteins contain three additional residues: R37, L38, and E39.

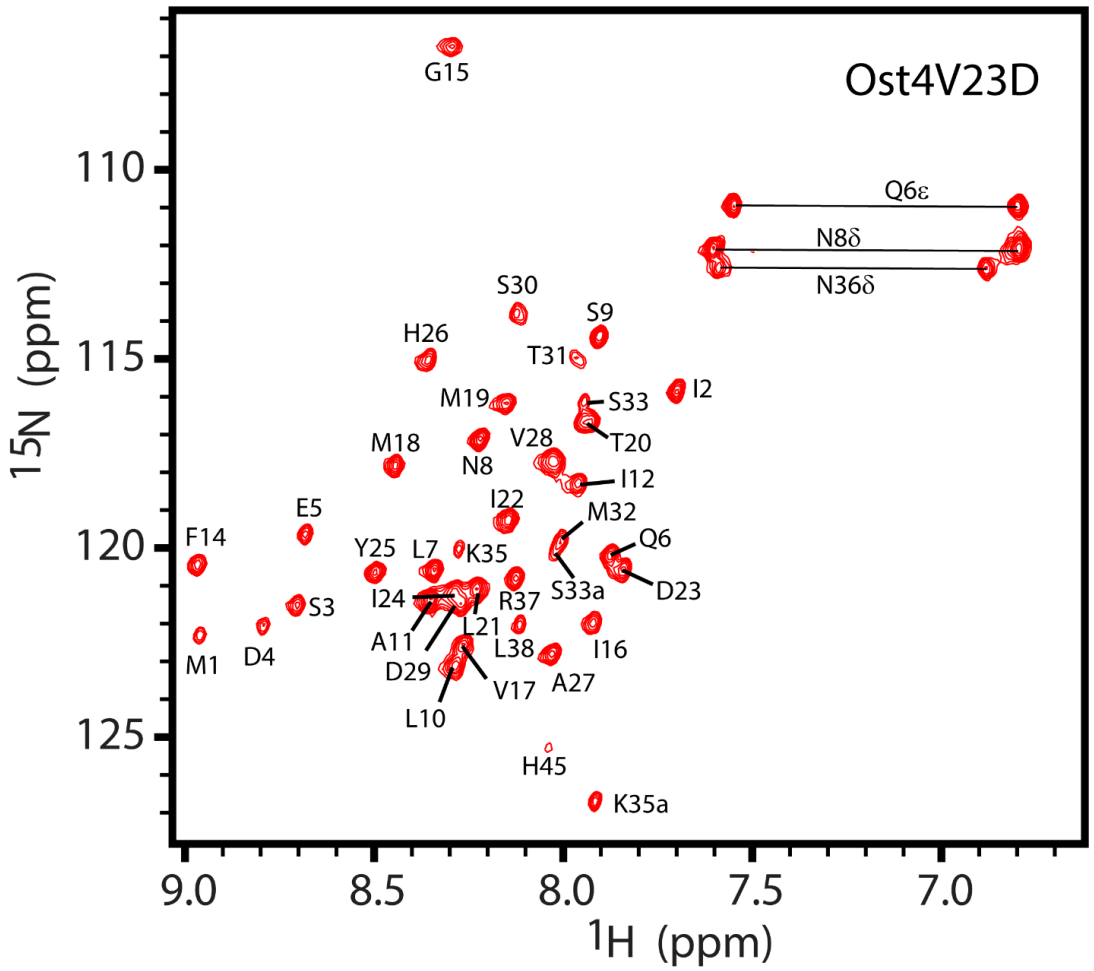


Figure 3.13: 2 D [^{1}H, ^{15}N]-HSQC spectra of Ost4V23D at pH 6.5 at 308 K. The proteins contain 45 amino acid residues including nine additional residues in the C-terminal 6X-His-tag. Only the terminal histidine (H45) of the hexahistidine tag was assigned. The residues from Met^{1} - Asn^{36} belong to Ost4 and Ost4V23D. The residues that do not belong to the proteins are R37, L38, E39, and His^{40}-His^{45} in the C-terminus.

In Ost4V23D, backbone resonance (^{1}H, ^{15}N, $^{13}C_{\alpha}$, and C_{β}) assignments were completed for all residues except ^{1}H, ^{15}N, $^{13}C_{\alpha}$ of T13 in the primary sequence (Met^{1}-Asn^{36}). The only proline residue at position 34 is present in its cis and trans forms. Thus, two sets of

resonances for residues S33 and K35 were observed for Ost4V23D in the 2D [^{1}H, ^{15}N]-HSQC spectra (Figure 3.13).

3.3.7.2 Secondary structure determination of OstV23D protein

The secondary structures of Ost4V23D were calculated using backbone chemical shifts by using TALOS+ (*13*) and SSP programs (*16*). The deviation of the ^{13}Cα and ^{13}Cβ chemical shifts from their mean random coil values was evaluated and the secondary structure was determined by using the secondary structure propensity (SSP) program (*16*). As shown in Figure 3.13, $\Delta\delta C_\alpha$ and $\Delta\delta C_\alpha$-$\Delta\delta C_\beta$ were plotted against the protein sequence. The secondary chemical shift values were derived by using the SSP program. Both programs predict a single helix encompassing residues Asp^{4}-Met^{32} for Ost4V23D (Figure 3.14).

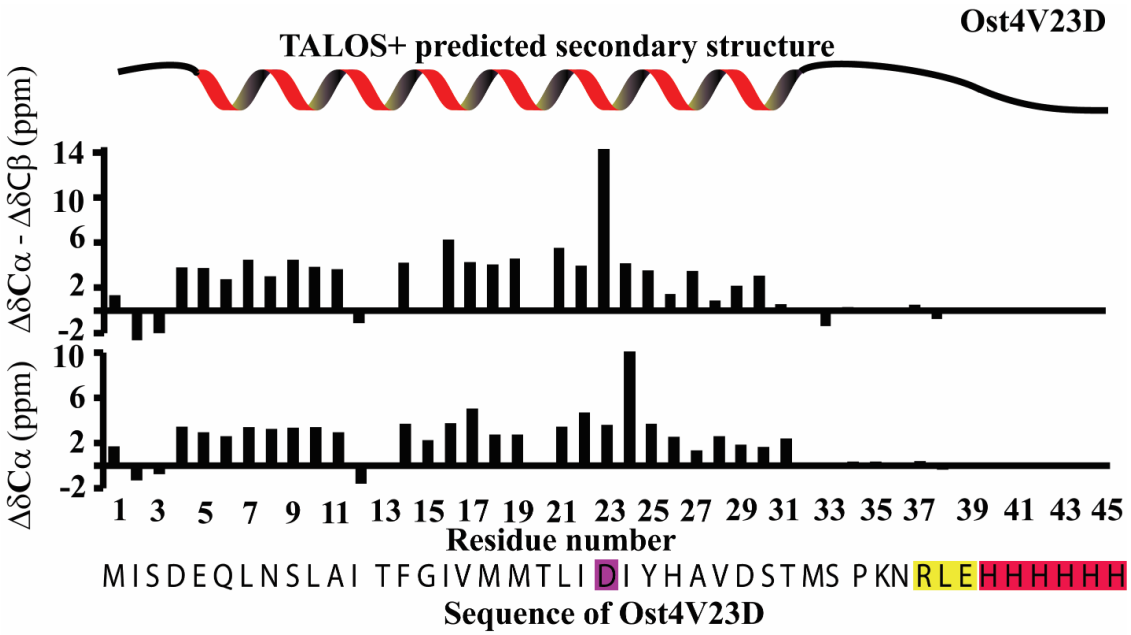


Figure 3.14: TALOS+ predicted secondary structure, $\Delta\delta C_\alpha$–$\Delta\delta C_\beta$, and $\Delta\delta C_\alpha$ secondary chemical shifts of Ost4. The TALOS+ predicted secondary structure of the protein are

shown on the top. Positive values in secondary chemical shifts indicate α–structure propensity and negative values indicate β–structure or random coil propensity. The secondary chemical shifts were derived by using the SSP program (*16*). The residues R37, L38, and E39 (highlighted in yellow) along with the C-terminal 6X-His-tag (highlighted in red) do not belong to Ost4V23D. The mutated residue, D23, is highlighted in a purple background. The figure was adapted from reference (*20*).

3.3.7.3 Side chain assignment of the Ost4V23D

Along with the backbone chemical shift, we have reported the side chain chemical shift assignment of Ost4V23D protein (*20*). The side chain chemical shift assignment of proteins is a bit easier and more straight-forward if the backbone chemical shifts are successfully assigned for a protein. Various TOCSY experiments such as ^{15}N HSQC TOCSY, HCCH TOCSY, HC(CO)NH experiments are used to assign the side chain chemical shifts.

The side chain protons of the (i) residue were assigned by using the ^{15}N HSQC TOCSY experiment. In these experiments, the side chain protons of the (i-1) residue could appear with the negative signal. HC(CO)NH TOCSY experiment provides a side chain proton of the (i-1) residue. The strips of residues Ile^{2}-Leu^{10} are providing resonances of their preceding residues are shown in Figure 3.15.

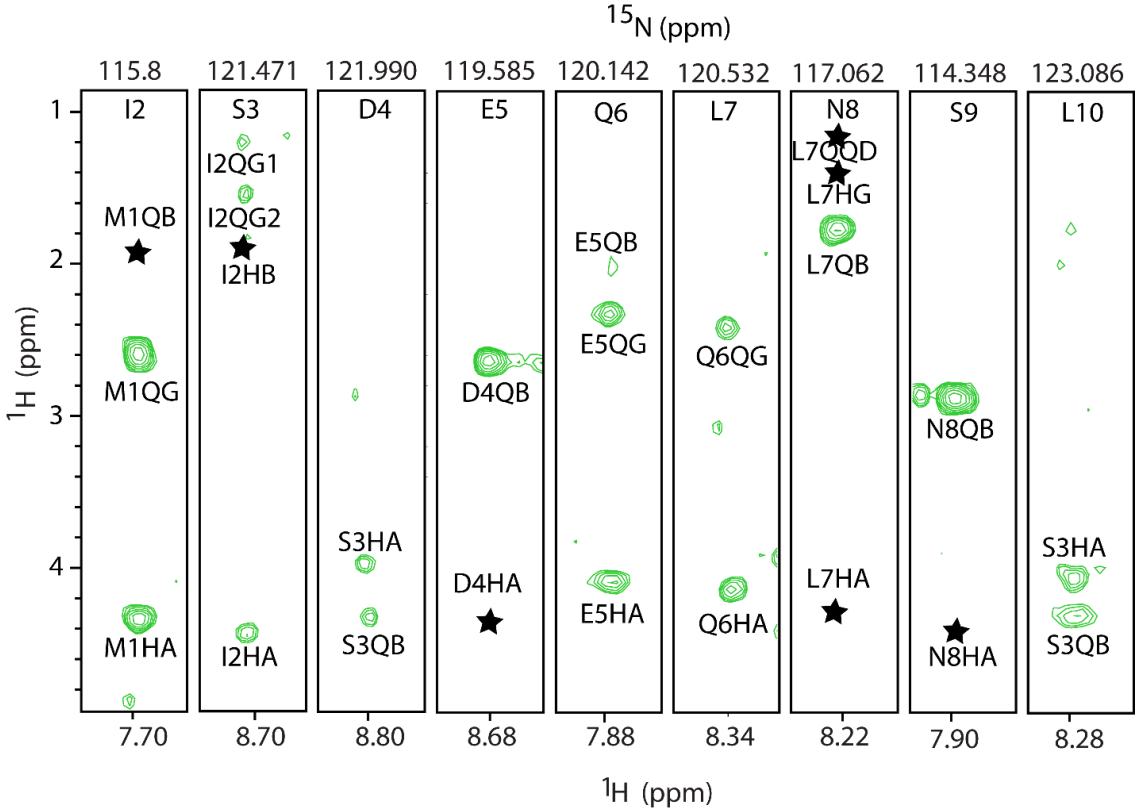


Figure 3.15: Strips of residues Ile[2]-Leu[10] of HC(CO)NH TOCSY spectrum. This spectrum provides the ^{1}H side chain resonances of the (i-1) residues. The side chain resonances that did not appear in the spectrum are indicated by *.

In the spectrum of HC(CO)NH TOCSY, some of the resonances may not appear (indicated by * in Figure 3.15). The missing resonances were assigned successfully by using the HCCH TOCSY experiment (Figure 3.16).

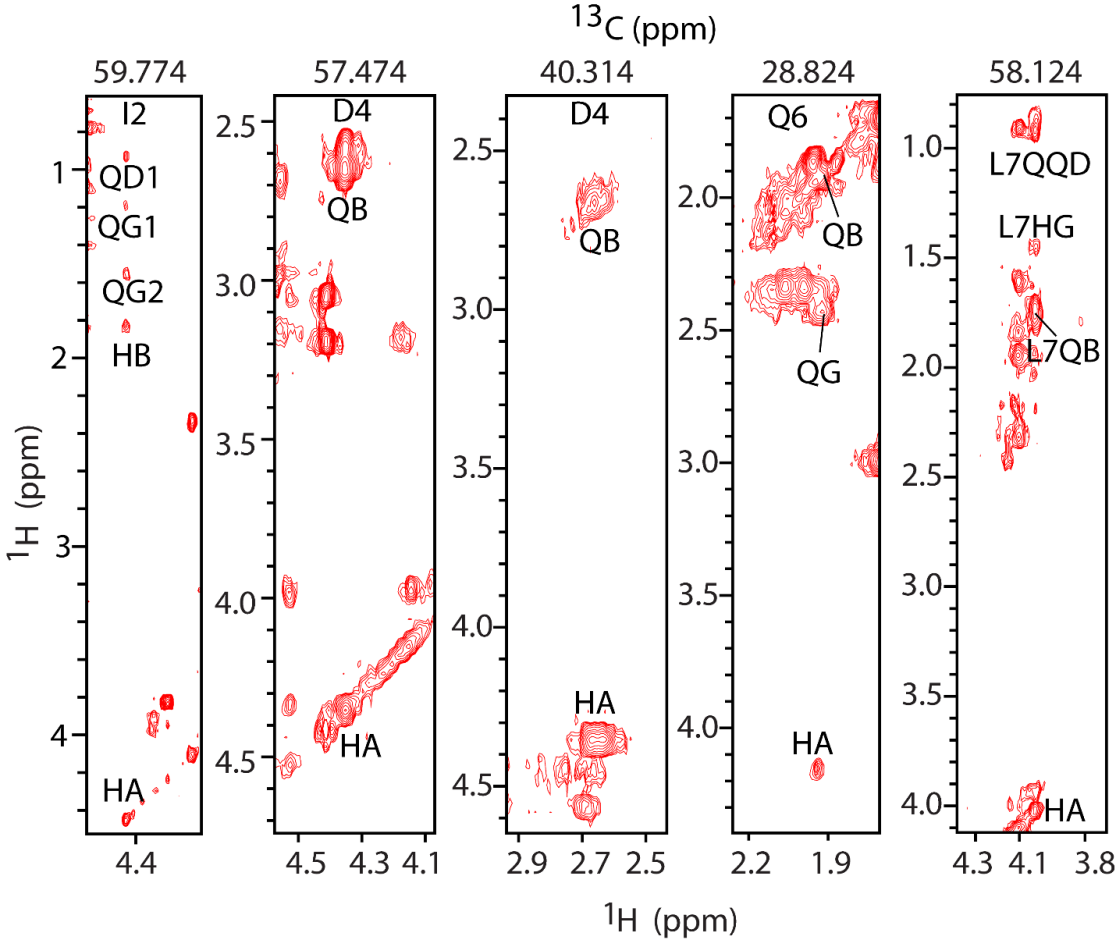


Figure 3.16: Strips of residues I2, D4, Q6, and L7 from 3D HCCH TOCSY experiment. This experiment provides the side chain ^{1}H resonances of the (i) residue.

3.3.7.4 Comparison of backbone resonances and secondary structures of Ost4 and Ost4V23D

The backbone chemical shifts of the Ost4 protein, derived from various 2D and 3D data sets, changed significantly upon V23D mutation in the protein (Figure 3.17 a, b, and c). This chemical shift deviation was throughout the protein sequence with large deviations at the proximity of V23 residue. This indicates that the two proteins, the wild type and the V23D mutant, either have a structural difference or are in a different environmental setting.

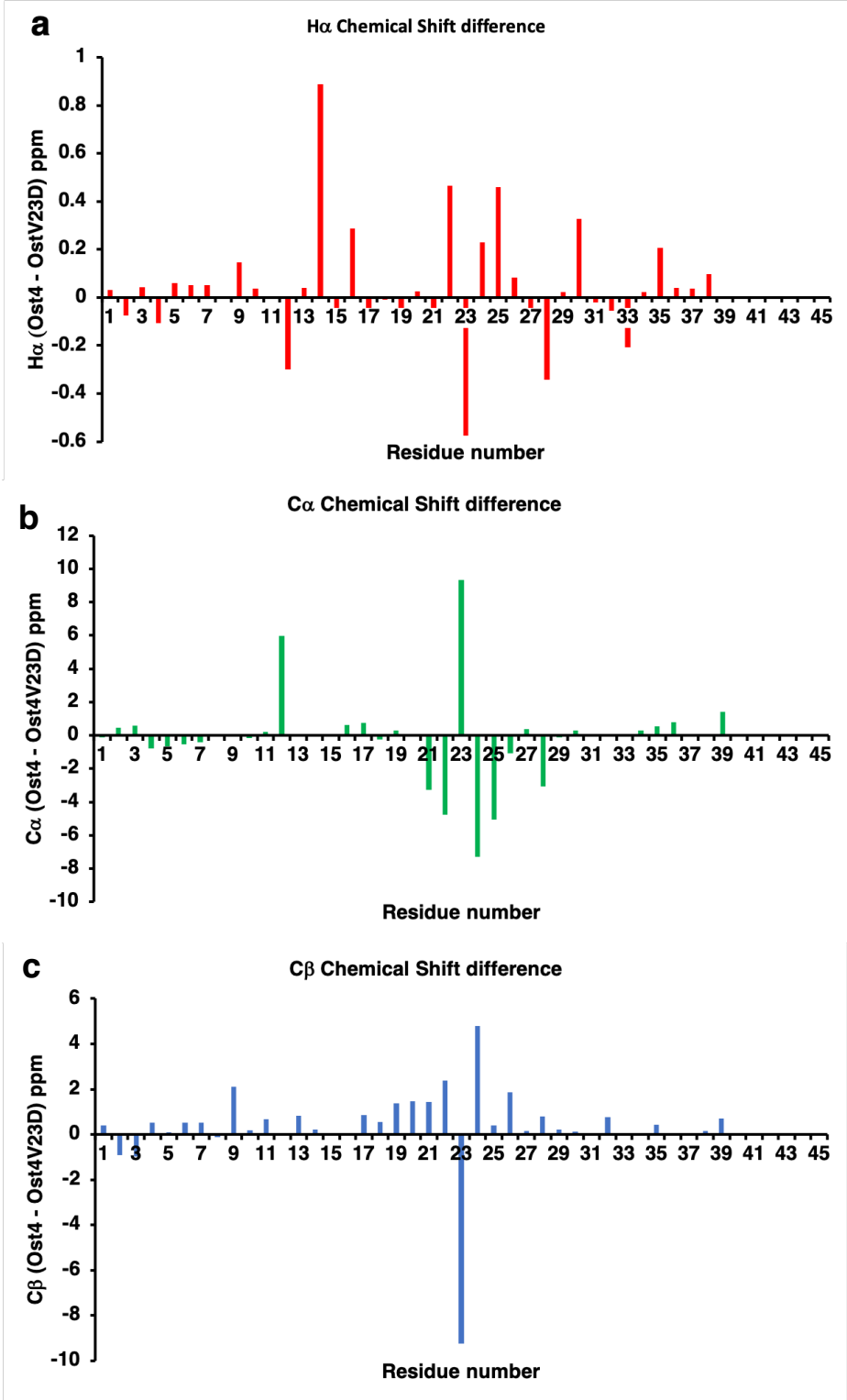


Figure 3.17: Comparison of solution NMR (a) Hα, (b) Cα, and (c) Cβ chemical shift differences upon mutation of the V23 residue to D in Ost4 protein.

Figure 3.18 shows the overlay of [^{1}H, ^{15}N]-HSQC spectra of Ost4p and Ost4pV23D with resonance assignments. Well-resolved resonances in the [^{1}H, ^{15}N]-HSQC of both the proteins indicate that the proteins are well-folded in 100 mM DPC micelles

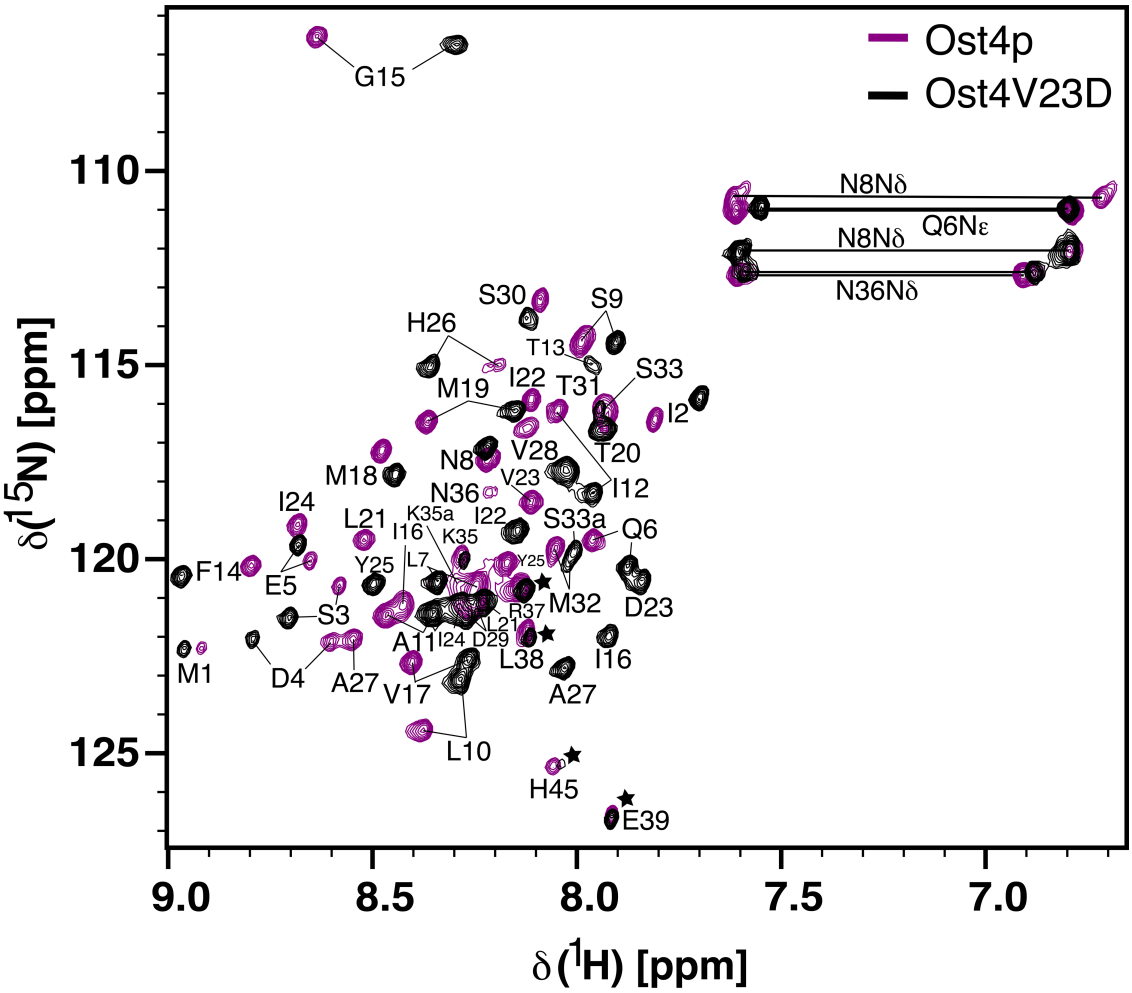


Figure 3.18: Overlaid 2D [^{1}H, ^{15}N] -HSQC spectra of Ost4p (green) with Ost4V23D (red) at pH 6.5 at 35 ºC. The protein contains 45 amino acid residues including six histidine residues. Only the terminal histidine (H45) of the 6X-His-tag was assigned. The residues from Met^{1}-Asn^{36} belong to the actual protein. The residues that do not belong to the protein, R37, L38, E39, and six histidine residues, are part of the C-terminal tag. These residues are indicated by * in the spectrum. An expanded view of the central region of the HSQC spectrum is shown on the lower right corner of the spectrum.

The backbone and side chain resonance assignments were completed for 95% and 91% respectively for Ost4 as well as 98 % and 87.6% for Ost4V23D. The lone proline residue at position 34 is present in the cis and trans forms for both of the proteins. Thus, two sets of resonances for residues S33 and K35 were observed for Ost4 and Ost4V23D in the 2D [^{1}H, ^{15}N]-HSQC spectra.

The secondary structures of Ost4 and Ost4V23D were determined independently using chemical shifts in TALOS+ (*13*) and SSP (*16*). As shown in Figures 3.8 and 3.14, both programs predict a single helix encompassing residues Asp^{4}-Met^{32} for Ost4 and Ost4V23D. The predictions from both the programs were supported by Hα secondary chemical shifts (ΔδHα) of both the proteins (Figure 4.19). In addition, single residue-specific secondary structure propensity (SSP) scores were calculated for both the proteins separately by using their $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts (Figures 3.8 and 3.14). The SSP scores shown in Figure 3.20 illustrate the difference in secondary structure propensities between Ost4 and Ost4V23D proteins more prominently than the $\Delta\delta C_{\alpha}$ and $\Delta\delta C_{\alpha}$-$\Delta\delta C_{\beta}$ plots shown in Figures 3.7 and 3.14. On the basis of the SSP scores, the overall propensities for α- structure of Ost4 and Ost4V23D are ~ 47% and ~ 60% respectively. This observation is consistent with what we reported previously (30% and 62% helicity for Ost4 and Ost4V23D respectively) using circular dichroism (*21*). Since Ost4 and Ost4V23D have significantly different α-structure propensity, their 2D [^{1}H, ^{15}N]-HSQC spectra are distinct (Figure 3.18).

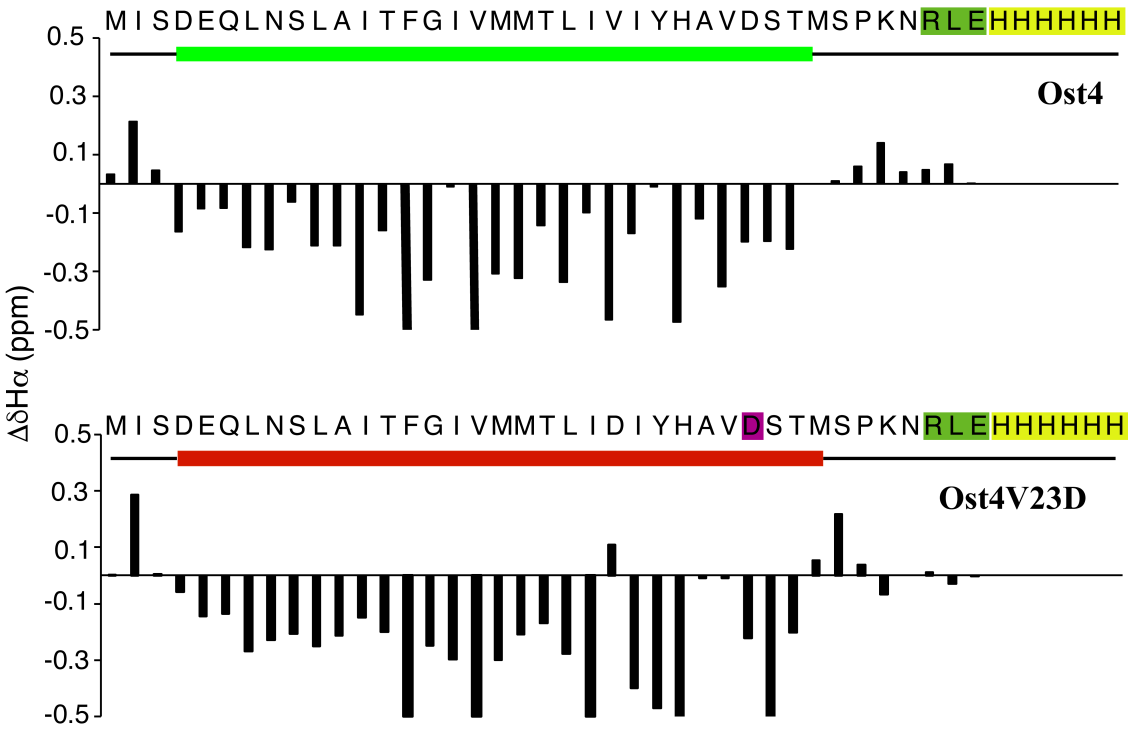


Figure 3.19: ΔδHα secondary chemical shift of Ost4 (top) and Ost4V23D (bottom). The TALOS+ predicted secondary structures of each protein are shown on the top. Negative values in secondary chemical shifts indicate α–structure propensity and positive values indicate β–structure or random coil propensity. The secondary chemical shifts were derived by using the SSP program (*16*). Residues R37, L38, and E39 (highlighted in green) along with the C-terminal 6xHis-tag tag (highlighted in yellow) do not belong to Ost4V23D. The mutated residue, D23, is highlighted in a purple background.

The distinct 2D [^{1}H, ^{15}N]-HSQC spectra of Ost4 and Ost4V23D along with significant differences in their helical structure suggests that mutation of V^{23} to D changes the environment around the protein, thus affecting the stability and function of the OST

enzyme. Structure-function studies of OST enzyme complex are critical in elucidating the mechanism of *N*-linked glycosylation.

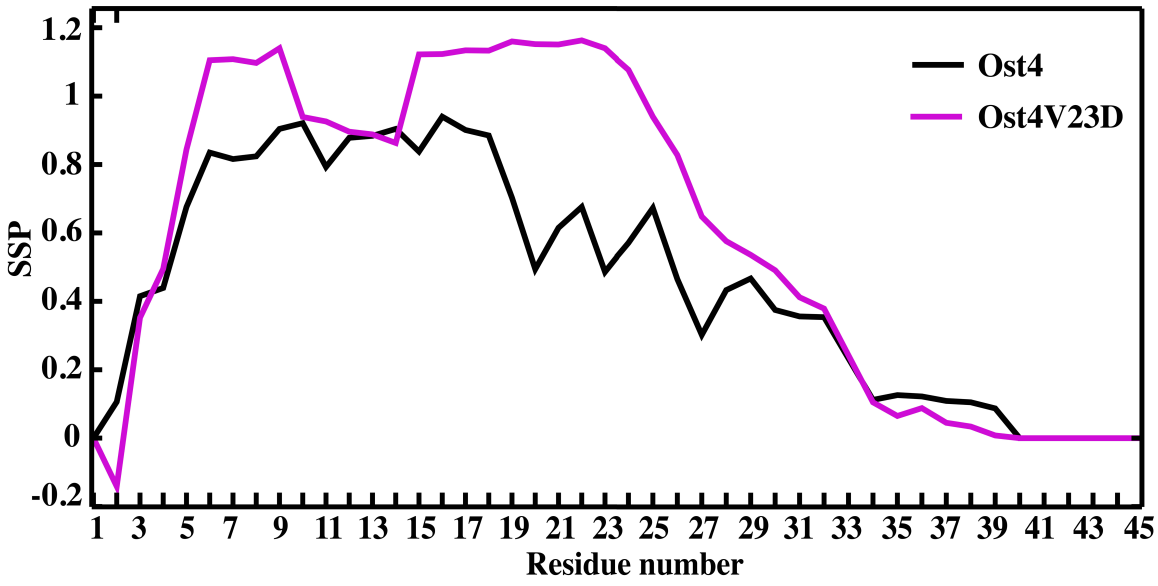


Figure 3.20: Secondary structure propensities (SSP) for wildtype Ost4 (green line) and mutant Ost4V23D (red line). The SSP for each protein was calculated using the $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts with an SSP limit of 1.2. Ost4V23D mutant protein contains a higher overall α–structural propensity than that of wildtype Ost4.

3.4 Conclusion

Backbone and side chain resonances of Ost4 and Ost4V23D protein were assigned successfully by using various 2D and 3D solution NMR data sets. A significant deviation of backbone chemical shifts, specifically near the V23 residue of Ost4, upon V23D mutation indicated that the two proteins may either have a structural or an environmental difference. The shifting of amide (NH) resonances in HSQC spectra upon V23D mutation

is another indication of changes in the chemical environment. Secondary structures of both the proteins were assessed by using assigned backbone chemical shifts. Although the secondary structure of both the proteins was similar, their SSP was significantly different suggesting the difference in their α-helical propensity. Therefore, the V23D mutation might affect the local environment of the protein thereby affecting the OST activity and function. The complete backbone and side chain assignment opened the door for NOE assignment and 3D structure calculation to get insight into the structure and function of Ost4 in the OST enzyme.

3.5 References

1. M. Forstner, L. Leder, L. M. Mayr, Optimization of protein expression systems for modern drug discovery. *Expert Rev Proteomics* **4**, 67-78 (2007).
2. K. Ito *et al.*, Advanced method for high-throughput expression of mutated eukaryotic membrane proteins in Saccharomyces cerevisiae. *Biochem Biophys Res Commun* **371**, 841-845 (2008).
3. E. P. Carpenter, K. Beis, A. D. Cameron, S. Iwata, Overcoming the challenges of membrane protein crystallography. *Current opinion in structural biology* **18**, 581-586 (2008).
4. K. Wüthrich, *NMR of proteins and Nucleic Acids*. (Wiley, New York, 1986).
5. J. L. Baneres *et al.*, Structure-based analysis of GPCR function: conformational adaptation of both agonist and receptor upon leukotriene B4 binding to recombinant BLT1. *J Mol Biol* **329**, 801-814 (2003).

6. D. S. Wishart, B. D. Sykes, The 13C Chemical-Shift Index: A simple method for the identification of protein secondary structure using 13C chemical-shift data. *Journal of Biomolecular NMR* **4**, 171-180 (1994).

7. D. S. Wishart, B. D. Sykes, F. M. Richards, The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **31**, 1647-1651 (1992).

8. D. A. Case, Calibration of ring-current effects in proteins and nucleic acids. *Journal of Biomolecular NMR* **6**, 341-346 (1995).

9. G. Cornilescu, F. Delaglio, A. Bax, Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *Journal of Biomolecular NMR* **13**, 289-302 (1999).

10. Y. Shen, A. Bax, Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. *J Biomol NMR* **38**, 289-302 (2007).

11. G. Wagner, A. Pardi, K. Wuethrich, Hydrogen bond length and proton NMR chemical shifts in proteins. *Journal of the American Chemical Society* **105**, 5948-5949 (1983).

12. M. P. Williamson, T. Asakura, Empirical Comparisons of Models for Chemical-Shift Calculation in Proteins. *Journal of Magnetic Resonance, Series B* **101**, 63-71 (1993).

13. Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* **44**, 213-223 (2009).

14. L.-H. Hung, R. Samudrala, Accurate and automated classification of protein secondary structure with PsiCSI. *Protein Sci* **12**, 288-295 (2003).

15. L. H. Hung, R. Samudrala, PROTINFO: Secondary and tertiary protein structure prediction. *Nucleic Acids Res* **31**, 3296-3299 (2003).

16. J. A. Marsh, V. K. Singh, Z. Jia, J. D. Forman-Kay, Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Protein Sci* **15**, 2795-2804 (2006).

17. A. Kumar, P. Ward, U. V. Katre, S. Mohanty, A novel and simple method of production and biophysical characterization of a mini-membrane protein, Ost4p: a subunit of yeast oligosaccharyl transferase. *Biopolymers* **97**, 499-507 (2012).

18. F. Delaglio *et al.*, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995).

19. W. Lee, M. Tonelli, J. L. Markley, NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325-1327 (2015).

20. B. P. Chaudhary, D. Zoetewey, S. Mohanty, 1H, 13C, 15N resonance assignments and secondary structure of yeast oligosaccharyltransferase subunit Ost4 and its functionally important mutant Ost4V23D. *Biomolecular NMR Assignments*, (2020).

21. B. Chaudhary, S. Mazumder, S. Mohanty, Production and biophysical characterization of a mini-membrane protein, Ost4V23D: A functionally important mutant of yeast oligosaccharyltransferase subunit Ost4p. *Protein Expr Purif* **139**, 43-48 (2017).

CHAPTER 4

SOLUTION NMR AND MOLECULAR DYNAMICS STUDIES OF YEAST OLIGOSACCHARYLTRANSFERASE SUBUNITS Ost4 AND ITS CRITICAL MUTANT, Ost4V23D

4.1 Introduction

The NMR structure of protease inhibitor IIA was the first structure determined by using solution NMR structure and was reported in 1985 (*1*). Since then the recent advancements in the NMR field have made substantial improvements to provide atomic resolution structures of macromolecules including protein and protein complexes. In addition to providing structural information, NMR is able to provide information about the molecular motions, or dynamics, which could influence the protein function. Solution NMR is the only technique that allows 3D structure determination of biomolecules in solution, under physiological condition and without affecting the native protein conformation. Although the environment of the protein in solution NMR may not be identical to its native *in vivo* environment, but the sample condition for NMR can often be made in near-to-native environment conditions (*2*-*5*).

The solution NMR structure of a small/macromolecule is calculated by using distance restraints derived from proton-proton cross-relaxation called the nuclear Overhauser effect (or NOE). NOE distance restraints are generated from the assigned cross-peaks in a nuclear Overhauser effect (NOE) spectrum. The overall outline of the steps for 3D structure

determination of a membrane protein by using solution NMR is represented by the flowchart provided below (Figure 4.1). Generally, the NOESY cross-peaks are assigned on the basis of backbone and side-chain chemical shift assignments. The backbone and side-chain assignments have been previously discussed in Chapter 3.

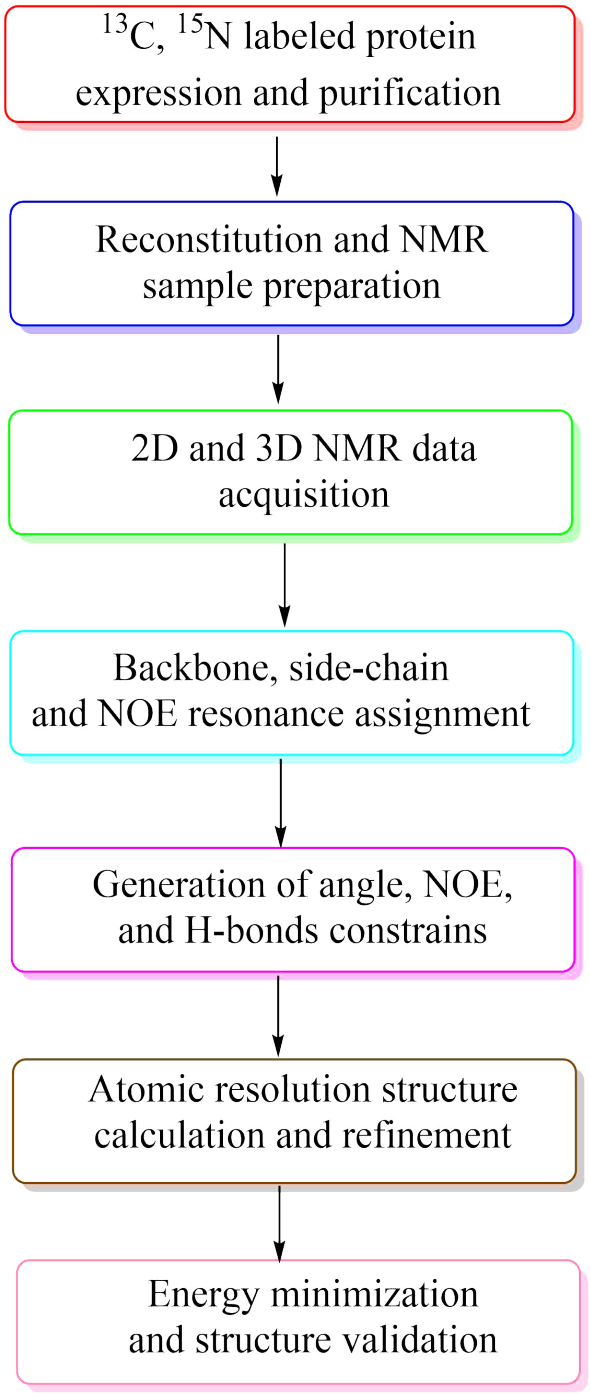


Figure 4.1: A schematic representation of 3D solution NMR structure determination.

Proteins are nanomolecular machines having various types of local and global motions. Knowledge about only the three-dimensional structure of the protein is not sufficient to understand various protein motions and their roles in protein folding and stability. NMR is the best analytical tool to study molecules and macromolecules such as proteins in their native-like environment. Additionally, the applied magnetic field in NMR is too weak to affect the molecular structure and motions. NMR is the most suited technique to study the motions of proteins (*6*). The measurement of ^{15}N relaxation rates is very important for gaining protein dynamic information since the nuclei relaxation is caused predominantly by the dipolar interaction with the connected protons. Spin – lattice relaxation time T_1, spin-spin relaxation time T_2, and $^{1}H - ^{15}N$ heteronuclear NOEs can provide information about protein dynamic events occurring over a wide range of NMR time scales.

In this chapter, the NOE assignments, 3D structure calculation, backbone dynamics, and molecular dynamics studies of Ost4 and Ost4V23D proteins in DPC micelles, a membrane mimicking system will be discussed.

4.2 Materials and methods

4.2.1 GB1 – Ost4 and GB1 – Ost4V23D protein expression and purification

GB1 – Ost4 and GB1 – Ost4V23D proteins were expressed and purified by following the procedure as described previously (*7, 8*) and have been explained in Chapter 2.

4.2.2 NMR data acquisition

The NMR data for this work were recorded using either a Bruker Avance 800 MHz spectrometer with a triple resonance $^{1}H/^{13}C/^{15}N$ TCI cryoprobe equipped with z-axis

pulsed-field gradients at the National High Field Magnetic Laboratory, Tallahassee, Florida, a Bruker Neo 700 MHz or 800 MHz spectrometers at Bruker BioSpin Corporation, Switzerland, or Varian Inova 600 MHz or 900 MHz spectrometer equipped with cold probes at the Department of Pharmacology, University of Colorado School of Medicine, Colorado. Some of the NMR data were collected on a Bruker Avance 600 MHz or a Varian Inova 900 MHz spectrometers with cryoprobes at the University of Minnesota NMR center. The samples between 300 μM and 1 mM of uniformly $^{15}N/^{13}C$ – labeled Ost4p or Ost4V23D in 50 mM phosphate buffer at pH 6.5 containing 1 mM EDTA, 0.01% NaN_3, 100 mM DPC and 5% D_2O were prepared for the backbone and side chain chemical shift assignment and structure determination. The backbone and side-chain assignments of Ost4 and Ost4V23D were performed using the NMR data sets as mentioned in Chapter 3. The raw data size for 2D ^{1}H, ^{15}N – HSQC, and ^{1}H, ^{13}C – HSQC was 2048 x 256 in the ^{1}H and $^{15}N/^{13}C$ dimensions. NOE distance restraints were collected from 3D ^{15}N – edited HSQC – NOESY (*9, 10*) with mixing times of 90 ms and 120 ms. The raw data size for ^{15}N – edited HSQC – NOESY was 2048 x 64 x 180 in ^{1}H (f3), ^{15}N (f2), and ^{1}H (f1) dimensions.

4.2.3 NMR data processing

All data were processed using either Topspin or NmrPipe (*11*) by following the procedure described in Chapter 2. The data were analyzed using NMRFAM – SPARKY (*12*) or NMRviewJ (*13*).

4.2.4 Resonance assignment of Ost4 and Ost4V23D

The backbone and side chain resonance assignment of Ost4 and Ost4V23D have been reported and discussed in Chapter 3 (*14*). The NOE cross-peaks in 3D ^{15}N-edited NOESY spectra were assigned successfully by making use of previously assigned backbone and side chain chemical shift values. A total of 217 and 254 NOE cross-peaks for Ost4 and Ost4V23D, respectively, were manually assigned by using NMRFAM – SPARKY (*12*).

4.2.5 3D Structure calculation of Ost4 and Ost4V23D

To calculate the 3D structure of Ost4 and Ost4V23D proteins, the CYANA 3.98beta program (*15*) was used. The initial step of the 3D structure calculation of proteins from NMR data is to generate the input files. The required input file for structure calculation is the primary sequence of the protein (*.seq), chemical shift list file (*.prot), dihedral angle restraint file (*.aco), and distance restraint file (*.peaks). The sequence file and chemical shift list file were used as input to CYANA to generate *.tab file to be used by TALOS+ to produce dihedral restraints file (*.aco) (Figure 4.2).

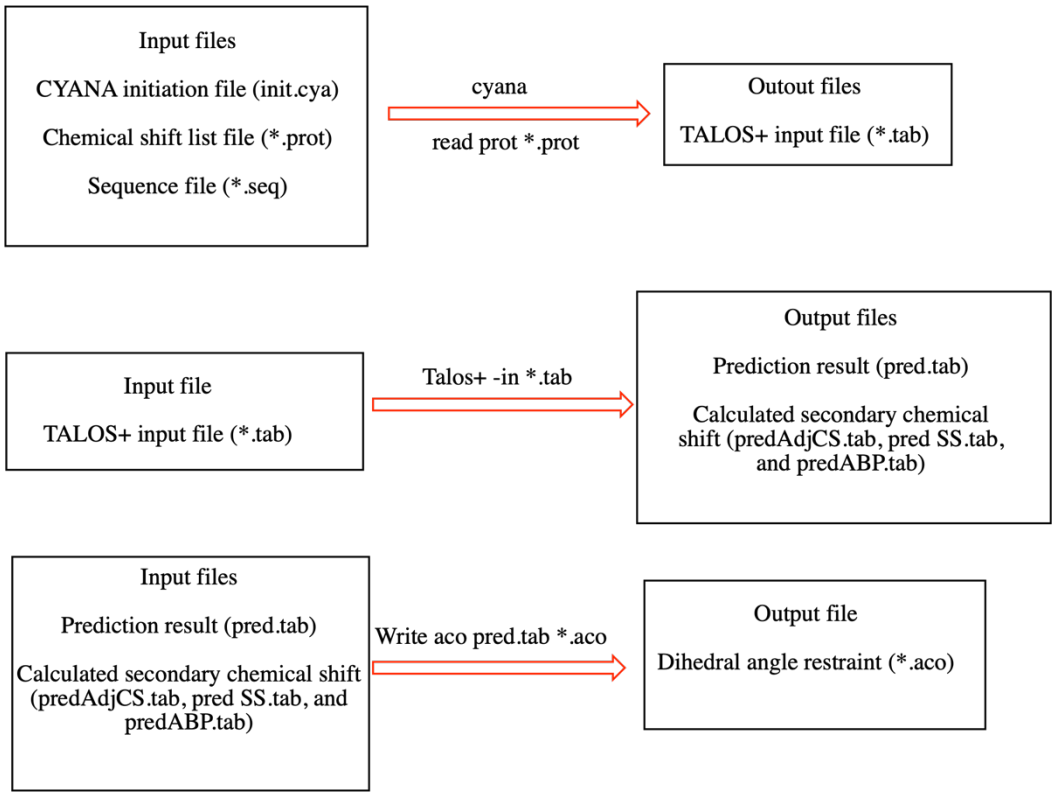


Figure 4.2: The steps of procuring dihedral angle restraints file by using CYANA 3.98beta and TALOS+ programs. The input files are listed on the left panel and the output files are listed on the right panel. * indicates possible preceding or following letters.

The distance restraint files (*.peaks) were generated from assigned NOE cross-peaks in XEASY format. The assignments were confirmed and/or corrected with the NOEASSIGN module of CYANA 3.98beta (*15*) using the standard protocol of eight iterative cycles of NOE assignment and structure calculation (Figure 4.3). During the iterative assignment process, 28 and 59 NOE peaks were removed due to overlap, ambiguity, or redundancy yielding a total of 182 and 195 experimental upper distance restraints, for Ost4 and Ost4V23D, respectively. The experimental upper distance limits were derived from the intensities of the assigned NOE cross-peaks using CYANA 3.98beta. A total of 68 dihedral

angle restraints for Ost4 and 62 dihedral angle restraints for Ost4V23D were derived from the assigned chemical shifts and primary sequences by using TALOS+ (*16*). In addition, 27 hydrogen bond restraints (two restrains per bond) for Ost4p and Ost4V23D were separately generated from the CSI by TALOS+. The prepared files were used as input in the program CYANA 3.98beta and eight iterative cycles of NOE assignment and structure calculation were run using the NOEASSIGN module. After each CYANA run, the generated output file was checked for further improvement of structure in the next run. The list of violations in *.ovw files was checked and the violations were fixed either by changing the assignment of by increasing the distance between the resonances (Figure 4.3). The violations along with energy functions (indicated by target function) were checked followed by the removal of suspicious upper limits of distance constraints. This process provided lesser violations, lower energy functions of the calculated structures. After several runs of the iterative process by checking and correcting the assignments, a final *.upl file containing correct distance restraints was used which generated *.pdb files of the structures having no violations and lowest energy function (Figure 4.3).

A total of 100 random structures were calculated and 20 structures with the lowest target functions were selected for solvent refinement using CNS (*17*). The 20 structures having the lowest energy and the best Ramachandran statistics were assessed by PROCHECK (*18*) and selected to represent the 3D structures of Ost4p and Ost4V23D in DPC micelles. The statistics of NMR structures of Ost4p and Ost4V23D proteins are shown in Table 4.1. The structures were visualized with Chimera (*19*), VMD (*20*), and Pymol (*21*). Figures were prepared either with Molmol (*22*) or Chimera.

Table 4.1: Structural statistics of the solution NMR structure of Ost4 and Ost4V23D. The table was adapted from reference (*23*).

Distance restraints	Ost4	Ost4V23D
Unambiguous	189	195
Intra residue, |i-j|=0	54	46
Sequential, |i-j|<=1	73	74
Medium range, 1<|i-j|<5	62	75
Long range, |i-j|>=5	0	0
Hydrogen bond restraints	27	27
Dihedral angle restraints	68	62
Interresidue distance restraints violations		
Violations	0	0
RMSD to average structures, Å		
Backbone (residues 4-32)	0.24	0.20
Heavy atoms (residues 4-32)	0.89	0.90
Ramachandran plot outliers, %		
Residues in most favored regions	97.1	91.4
Residues in additionally allowed regions	2.9	8.6
Residues in generously allowed regions	0	0
Residues in disallowed regions	0	0

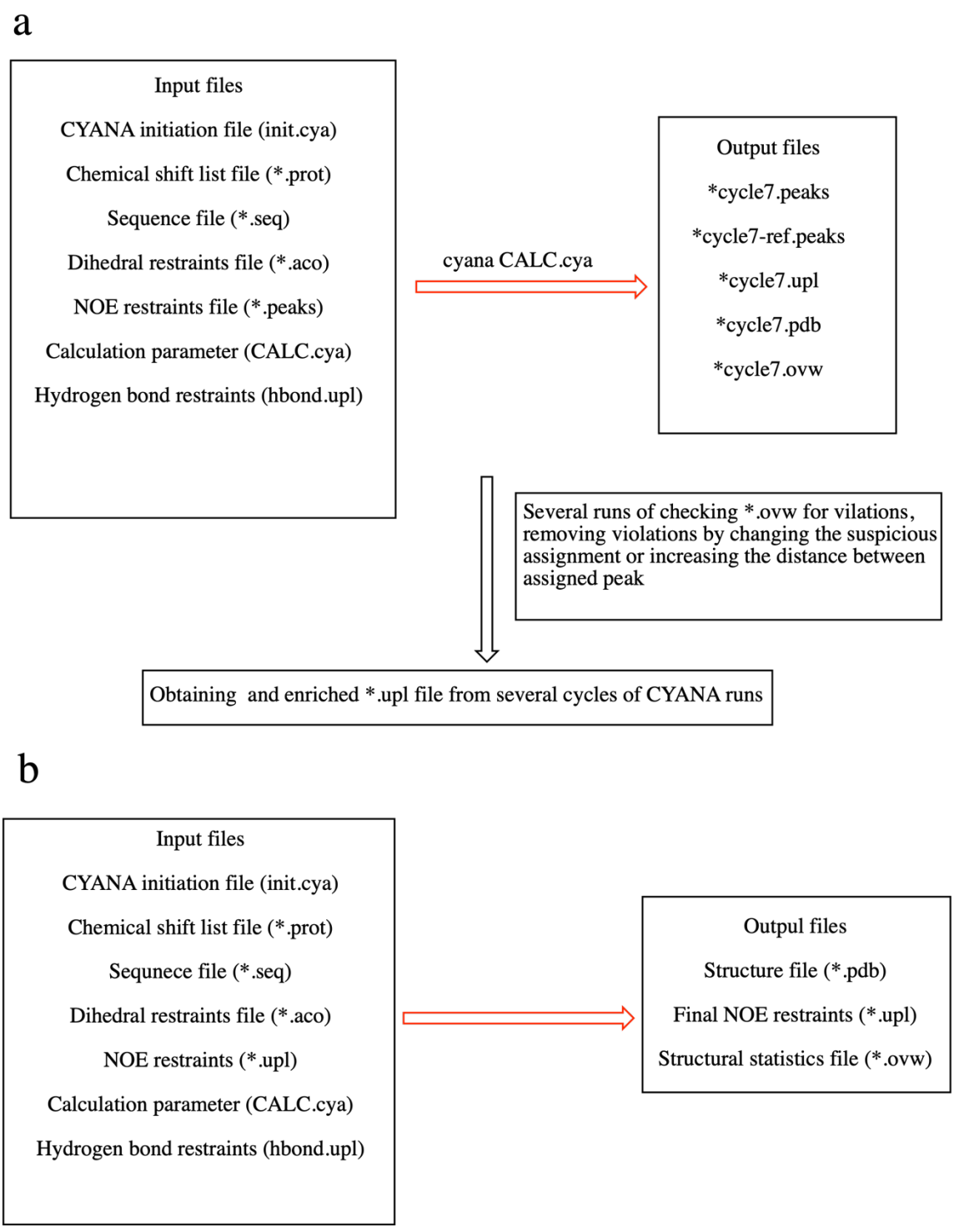


Figure 4.3: Iterative cycle for the procurement of an enriched upl file for the final step of structure calculation (a), and input and output files at the end of the structure calculation process (b). * denotes possible preceding or following letters.

4.2.6 Relaxation experiments

^{15}N T_1, T_2, and heteronuclear NOEs spectra were recorded using a Bruker 800 MHz (Avance Neo) spectrometer equipped with CP-TCI $^{1}H/^{13}C/^{15}N$ 5 mm inverse triple resonance high-resolution cryoprobe at Department of Chemistry, Oklahoma State University, Stillwater, OK. ^{15}N T_1, T_2, and heteronuclear NOE measurement data were recorded using pseudo-three-dimensional pulse programs provided by Dr Youlin Xia, Director of Structural Biology Department at St. Jude Children's Research Hospital, Memphis, TN. In all the relaxation measurement experiments, 128 x 2048 real data matrices in the t_1 x t_2 time domain were employed with spectral widths of 2919.86 and 12500 Hz, respectively. ^{15}N T_1 and T_2 measurement experiments were recorded with 8 scans and NOE measurement experiments were recorded with 32 scans per t_1 increment. T_1 values were acquired using eight relaxation delays of 20, 100, 200, 300, 400, 500, 600, and 800 ms. T_2 values were recorded using eight relaxation delays of 16.9, 33.8, 50.7, 67.6, 101.4, 135.2, 169, and 202.8 ms. Steady-state $^{1}H - ^{15}N$ NOE values were determined from the ratio of peak height in a pair (I/I_0) of NMR spectra recorded in the presence (NOE) and absence (NONOE) of a proton saturation period of 5 s. All the data were processed using the NmrPipe software package (*11*). The relaxation data analysis was performed using NMRFAM Sparky (*12*). Overall correlation time (τ_c) for Ost4 and Ost4V23D proteins were calculated from T_1, and T_2 data using the Tensor2 software package (*24*).

4.2.7 Molecular dynamics simulation

4.2.7.1 Molecular dynamics simulation of Ost4 and Ost4V23D in DPC micelles

All molecular dynamics (MD) simulations were performed using the GROMACS simulation package version 5.1.5 (*25*). Initial protein micelle systems for Ost4 and Ost4V23D were derived by using the web version membrane builder module of the CHARMM-GUI (*26, 27*). The protein-micelles systems were built using 65 DPC detergent molecules. The systems were solvated with the 6729 molecules of the TIP3P water model and neutralized by adding Na^+ ions yielding a cubic box system of dimension 9.484 Å x 9.484 Å x 9.484 Å. All bonds were constrained using the LINCS algorithm (*28*). An integration time step of 2 fs was used during the MD simulation. The temperature was kept constant at 35 ºC using Nose-Hoover coupling (*29*) with a 1 ps coupling time constant. The pressure was maintained constant at 1 bar using Parrinello-Rahman coupling (*30, 31*) with a 5 ps coupling time constant. A cutoff of 1.2 nm was applied for van der Waals interactions. Particle mesh Ewald (PME) (*32*) was applied to treat the electrostatic interactions. The systems were subjected to <1000 steps of steepest descent energy minimization using the CHARMM36 (*33*) force field. For protein-micelles systems, 1 ns of NPT (constant number of particles, pressure, and temperature) equilibration was performed after NVT (constant number of particles, volume, and temperature) equilibration. Finally, MD simulations of 150 ns for protein-micelle systems were carried out under similar conditions of NPT equilibration. The last 40 ns were used for analyses. Analyses were performed using GROMACS inbuilt tools. The number of DPC-tail to protein contacts and backbone root mean squared deviation (RMSD) were assessed using GROMACS inbuilt tools (distance and rms). Some of the computing for this project was performed at the OSU High-Performance Computing Center at Oklahoma State University.

4.2.7.2 Molecular dynamics simulation of Ost4 and Ost4V23D in bilayers

To understand the impact of mutations in a cell, we carried out molecular dynamic simulation of both Ost4 and Ost4V23D in the OST complex in a bilayer. The simulation was carried out using the OSU High-Performance Computing Center at Oklahoma State University by Dr. Martin McCullagh (*23*).

Membrane-bound systems of the WT and Ost4V23D system consisted of Ost4, Stt3, Ost1, Ost3, and Ost5. The systems were generated from the initial cryoEM structure (*34*) using the CHARMM GUI. The bilayer of dipalmitoylphosphatidylcholine (DPPC) lipids containing 307 lipid molecules in the upper leaflet and 320 lipid molecules in the lower leaflet was used. TIP3P water with a buffer size of 2.0 nm was added along with 0.1 M NaCl yielding a box of initial dimension 15 x 15 x 12.4 nm. The systems were simulated using the GPU accelerated Amber 18 software (*35*). The simulations were carried out in the NPT ensemble with anisotropic pressure coupling using the constant surface area. Combined, this system was 240K atoms and the simulation of each system exceeded 600 ns.

Solvent accessible surface area (SASA) calculations were performed using the VMD tool. The fraction of hydrophobic and hydrophilic surface areas in the peptide that is exposed to water was calculated. The selection of hydrophobic residues consists of residues in the transmembrane domain (residues Leu^{10}-Val^{28}). Hydrophilic residues such as Thr^{13}, Thr^{20}, Asp^{23}, His^{26} in the transmembrane range were chosen for the hydrophilic SASA calculation.

4.3 Results

4.3.1 Chemical shift perturbation upon mutation

By using far-UV circular dichroism (CD) and [^{1}H, ^{15}N] HSQC spectroscopy, we have previously reported that Ost4 and Ost4V23D in DPC micelles are well folded and it is feasible to determine the 3D structures of these proteins in DPC micelles (*7*). The recombinant Ost4 and Ost4V23D proteins were expressed, purified and reconstituted in 100 mM deuterated DPC micelles for NMR data collection. Although Ost4 is a very small (~5 kDa) membrane protein, it behaves like a ~21 kDa protein in DPC micelles. Therefore, 3D NMR experiments were used for complete resonance assignments. We have previously reported the backbone and side chain resonance assignments of Ost4 and Ost4V23D proteins (*36*). The ^{1}H, ^{13}C, and ^{15}N resonances in Ost4V23D protein were reassigned by using various 3D NMR data sets. The reassignment was essential because nearly all backbone resonances in Ost4V23D shifted either slightly or significantly as shown by the chemical shift perturbation in Figure 4.4a and Figure 4.4b. In Ost4V23D, the amide backbone resonance of L21, I22, D23, Y25 and A27 residues shifted dramatically (Figure 4.4a and Figure 4.4b). The changes in chemical shifts were quantified as $\sqrt{0.5[\delta_H^2+(0.3\delta_N^2)]}$, where δ_H and δ_N are observed chemical shift changes for ^{1}H and ^{15}N shifts in ppm, respectively. The change in chemical shifts of these residues was observed to be > 0.5 ppm (Figure 4.4b).

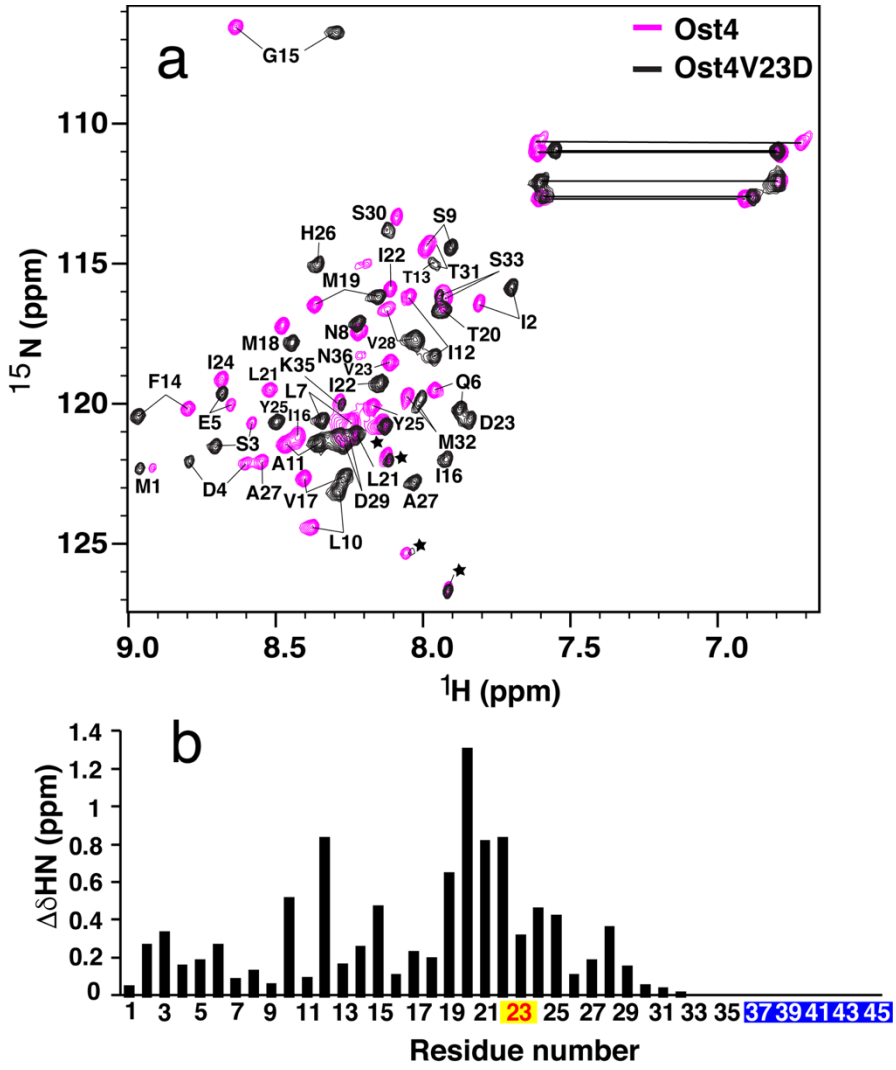
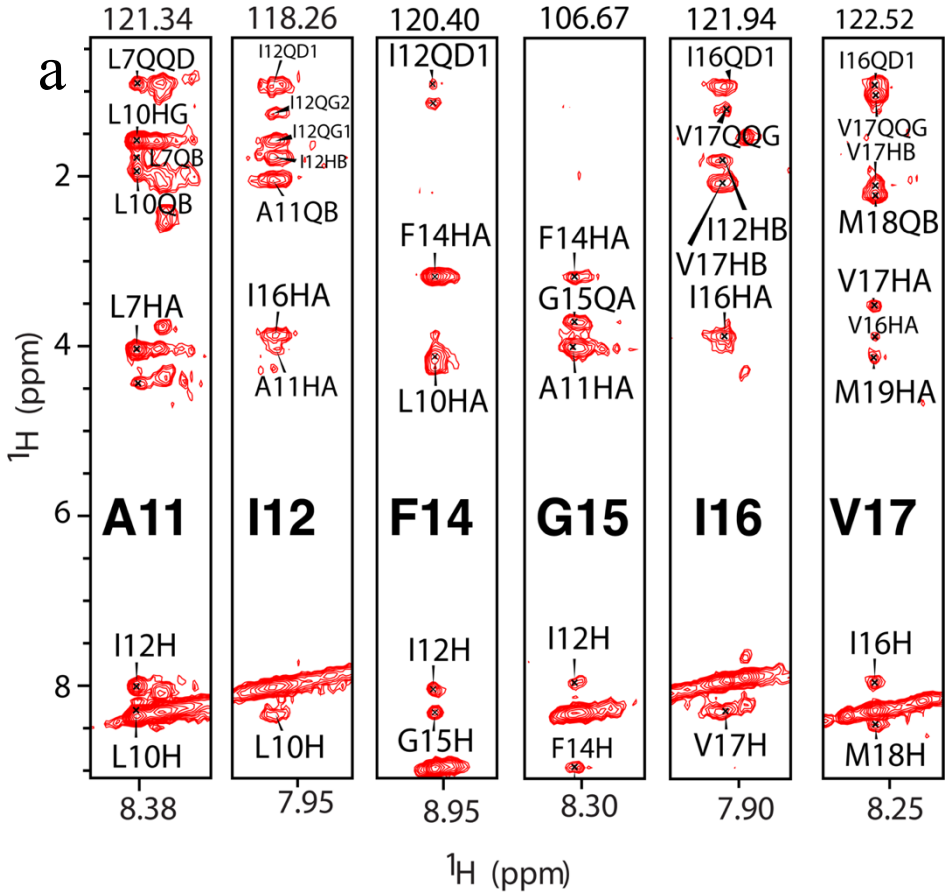


Figure 4.4: (a) Overlay of 2D [^{1}H, ^{15}N] HSQC of Ost4 (purple) with Ost4V23D (black). The C-terminal tag residues that do not belong to Ost4 or Ost4V23D protein are indicated by * in the spectrum. (b) The plot of the chemical shift perturbations (CSP) upon mutation of valine at position 23 to aspartate. The resonances close to the mutation site for residues L21, I22, D23, and A27 are perturbed significantly showing CSP of >0.5 ppm. Residue number Met^{1}-Asn^{36} belongs to Ost4 and Ost4V23D. The mutated residue D23 is highlighted red in yellow background. The C-terminal tag residues Arg^{37}-His^{45} (highlighted white in blue background) do not belong to Ost4 or Ost4V23D.

4.3.2 NOE assignment and structure calculation of Ost4 and Ost4V23D proteins

The final and the most important stage of the resonance assignment is the NOE assignment because the NOE constraints are utilized to generate inter proton distance restraints that are required for structure determination.

The secondary structures of Ost4 and Ost4V23D determined by using the TALOS+ and SSP program indicated that both the proteins have a single α-helix encompassing residues Asp^{4}-Met^{32} (*36*). Once the backbone and side chain resonances are assigned, the very next step before structure calculation is the NOE cross-peak assignment. The NOE peaks provide the necessary distance constraints required for the structure determination. For both the proteins, NOE cross-peaks in ^{15}N-edited HSQC spectra were assigned for the generation of distance constraints. Figure 4.5a and 4.5b demonstrate strips of NOE assignments for Ost4V23D protein residues A11, I12, F14, G15, I16, V17, and L21 to H26. Although the NOE assignment appears to be straight-forward to assign the proton cross-peaks in the NOE spectra, the assignment is quite challenging because of either overlap between peaks or absence of expected peaks. During the process of NOE assignments, d_{NN} connectivity were used to resolve any ambiguity in the assignment (Figure 4.6). Additionally, extra care needs to be taken during NOE assignment to avoid picking up noise and not omitting useful NOE peaks.


a
121.34
118.26
120.40
106.67
121.94
122.52
L7QQD
L10HG
L7QB
L10QB
L7HA
I12QD1
I12QG2
I12QG1
I12HB
A11QB
I16HA
A11HA
I12QD1
F14HA
L10HA
F14HA
G15QA
A11HA
I16QD1
V17QQG
I12HB
V17HB
I16HA
I16QD1
V17QQG
V17HB
M18QB
V17HA
V16HA
M19HA
A11
I12
F14
G15
I16
V17
I12H
L10H
I12H
L10H
I12H
G15H
I12H
F14H
I16H
V17H
I16H
M18H
2
4
6
8
1H (ppm)
8.38
7.95
8.95
8.30
7.90
8.25
1H (ppm)

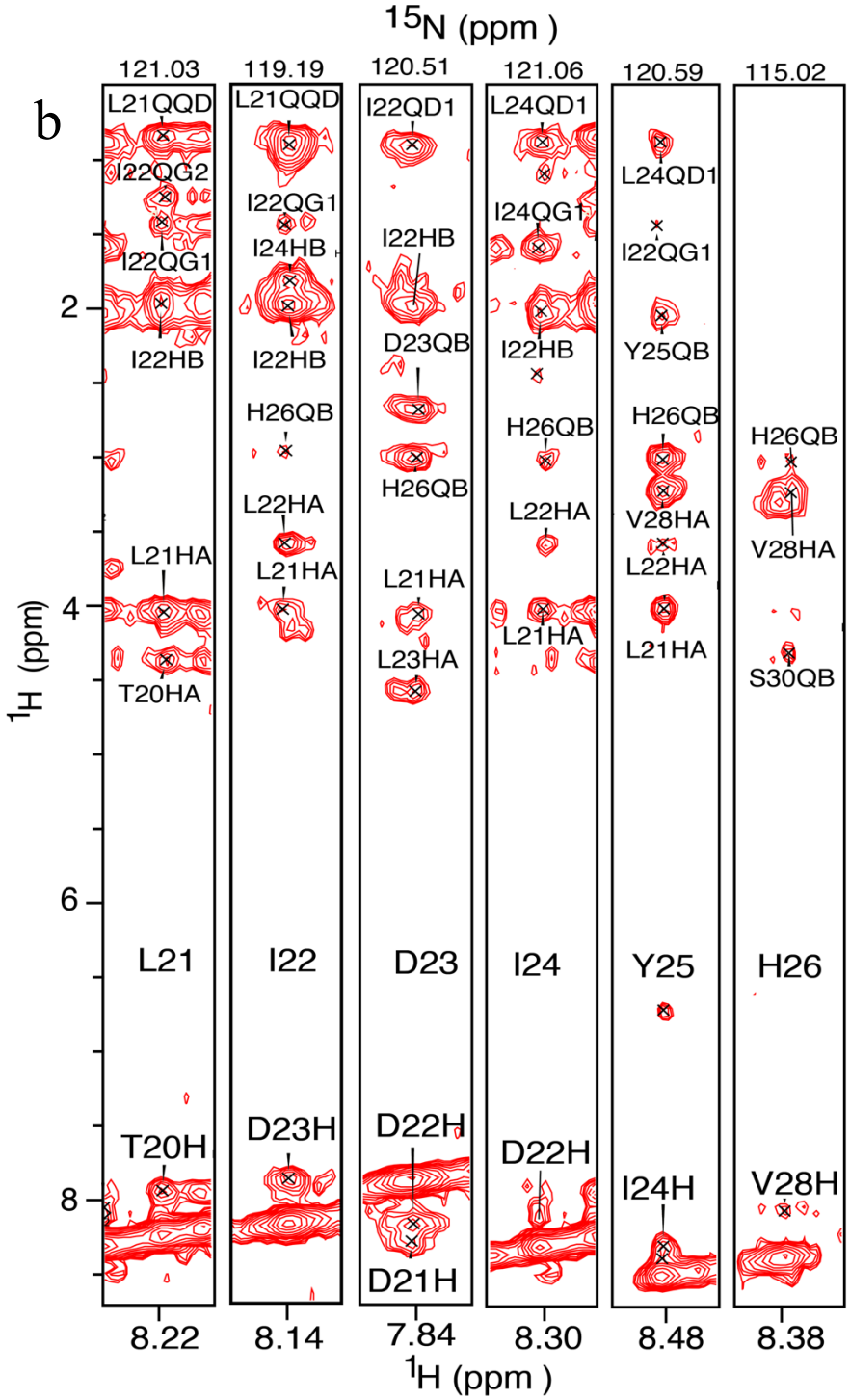


Figure 4.5: Strips of NOE assignment of Ost4V23D protein (a) for residues Ala[11], Ile[12], Phe[14], Gly[15], Ile[16], and Val[17]and (b) for residues Leu[21]-His[26].

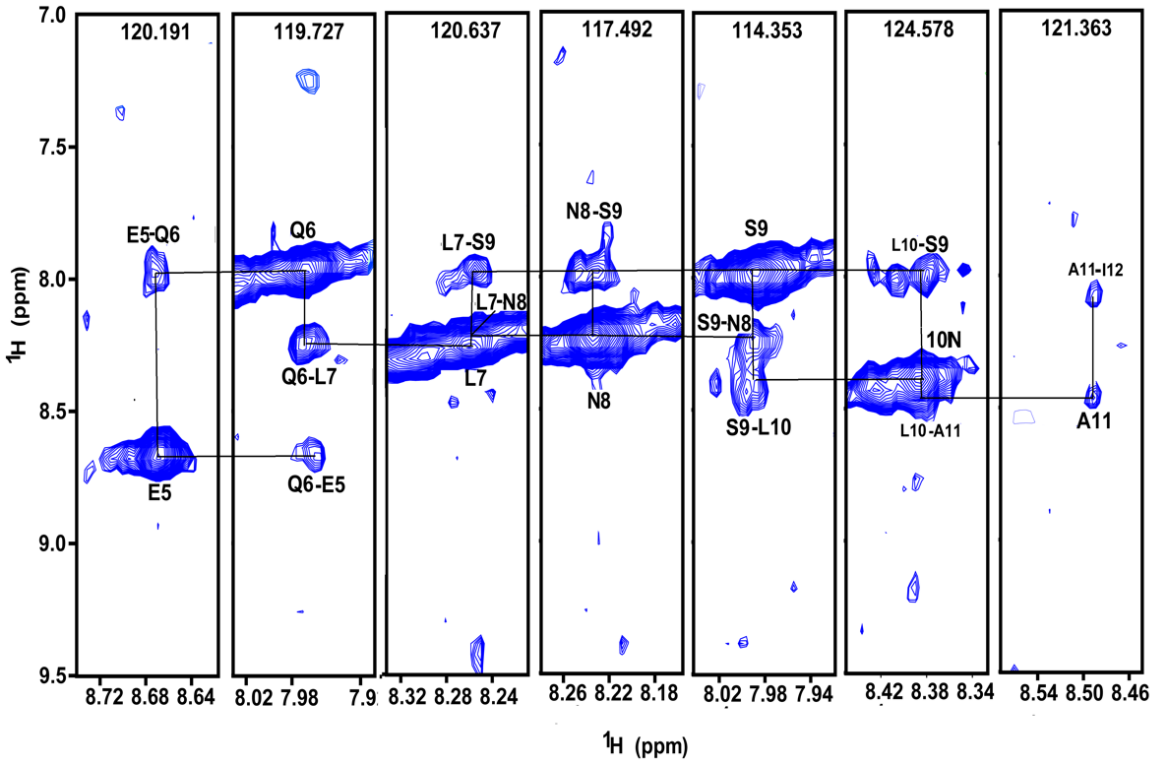


Figure 4.6: Strips showing sequential connectivity d_{NN} (i,i+1) and d_{NN} (i,i-1) in the Ost4 protein. The d_{NN} (i,i+1) and d_{NN} (i,i-1) for residues Leu^{21}-His^{26} can be observed in Figure 4.5.

A total of 217 peaks for Ost4 and 254 peaks for Ost4V23D were assigned manually using NMRFAM SPARKY. The assigned and unassigned NOEs were utilized by the NOE assign module of CYANA to assign the NOEs, correct the incorrectly assigned NOEs, and generate the upper distance limit necessary for structure calculation. Finally, 189 and 195 distance restraints were generated and used for structure determination of Ost4 and Ost4V23D, respectively. A summary of the NOE connectivity of Ost4 and Ost4V23D proteins, along with their secondary structural elements, is illustrated in Figure 4.7. Identification of the secondary structural elements was based on the secondary chemical shift $\Delta\delta^{13}C\alpha$, $\Delta\delta^{13}C\alpha$-$\Delta\delta^{13}C\beta$, and TALOS+ (*14*).

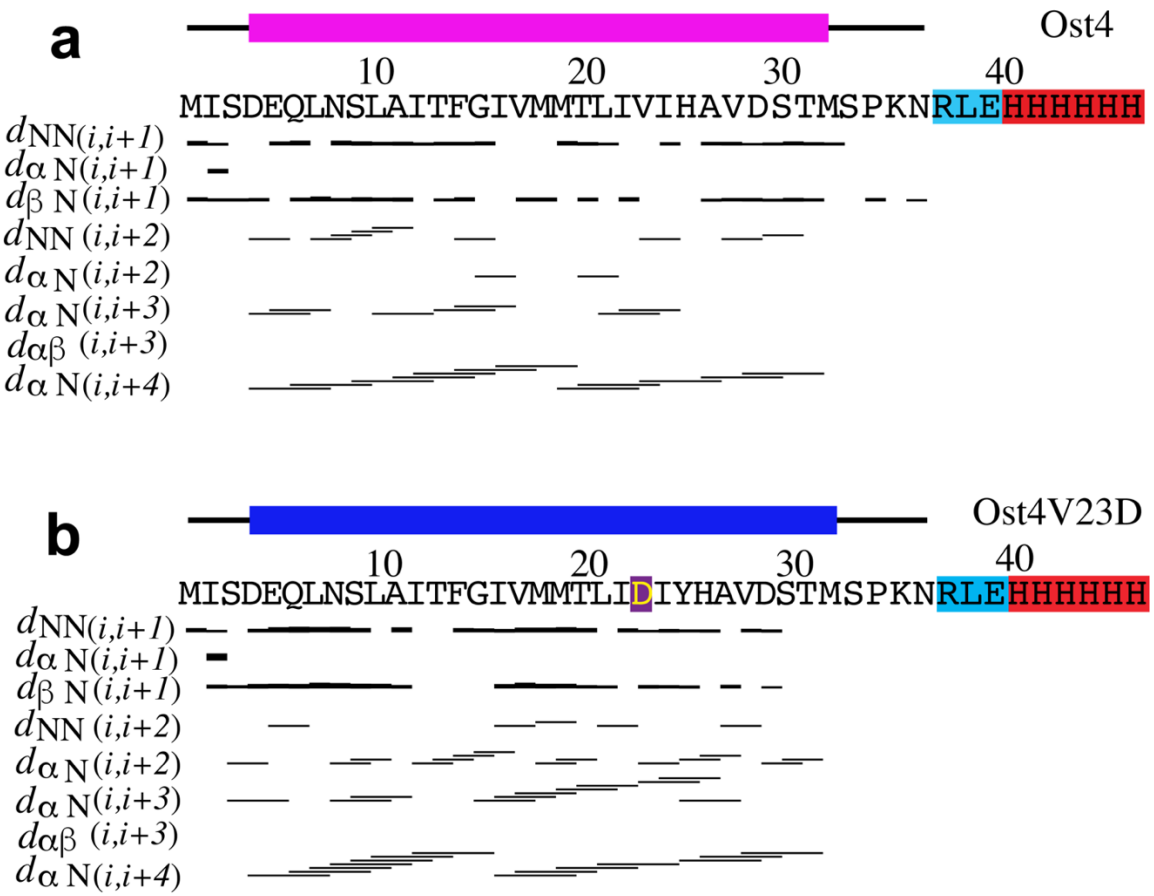


Figure 4.7: Predicted secondary structure, sequence, and summary of sequential NOE contacts of Ost4 (a) and Ost4V23D (b). The secondary structural elements were based on the secondary chemical shift $\Delta\delta^{13}C\alpha$, $\Delta\delta^{13}C\alpha$-$\Delta\delta^{13}C\beta$, and TALOS+ (*14*). The residue in yellow highlighted in purple is the mutated residue in Ost4V23D. The residues Arg^{37}-Glu^{39} (highlighted in cyan) and C-terminal 6X-His-tag (highlighted in red) do not belong to Ost4 or Ost4V23D.

For structure calculation, dihedral angles were derived from backbone chemical shifts using the TALOS+ program incorporated in NmrPipe. The energetically allowed regions for backbone dihedral angles ψ against Φ the Ost4 (Figure 4.8) and Ost4V23D (Figure 4.9) residues were

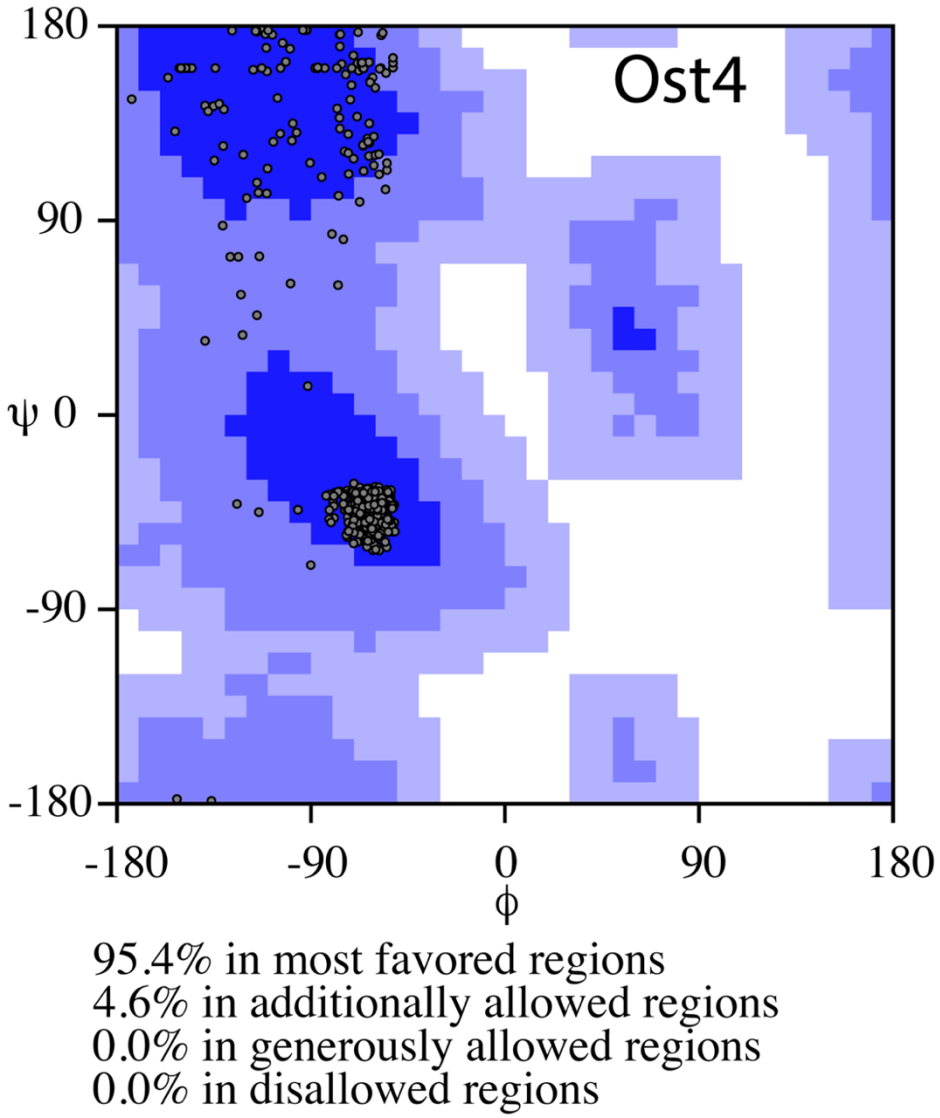


Figure 4.8: Ramachandran plots of Ost4. This plot validates the energetically allowed regions of dihedral angles of protein residues in a protein.

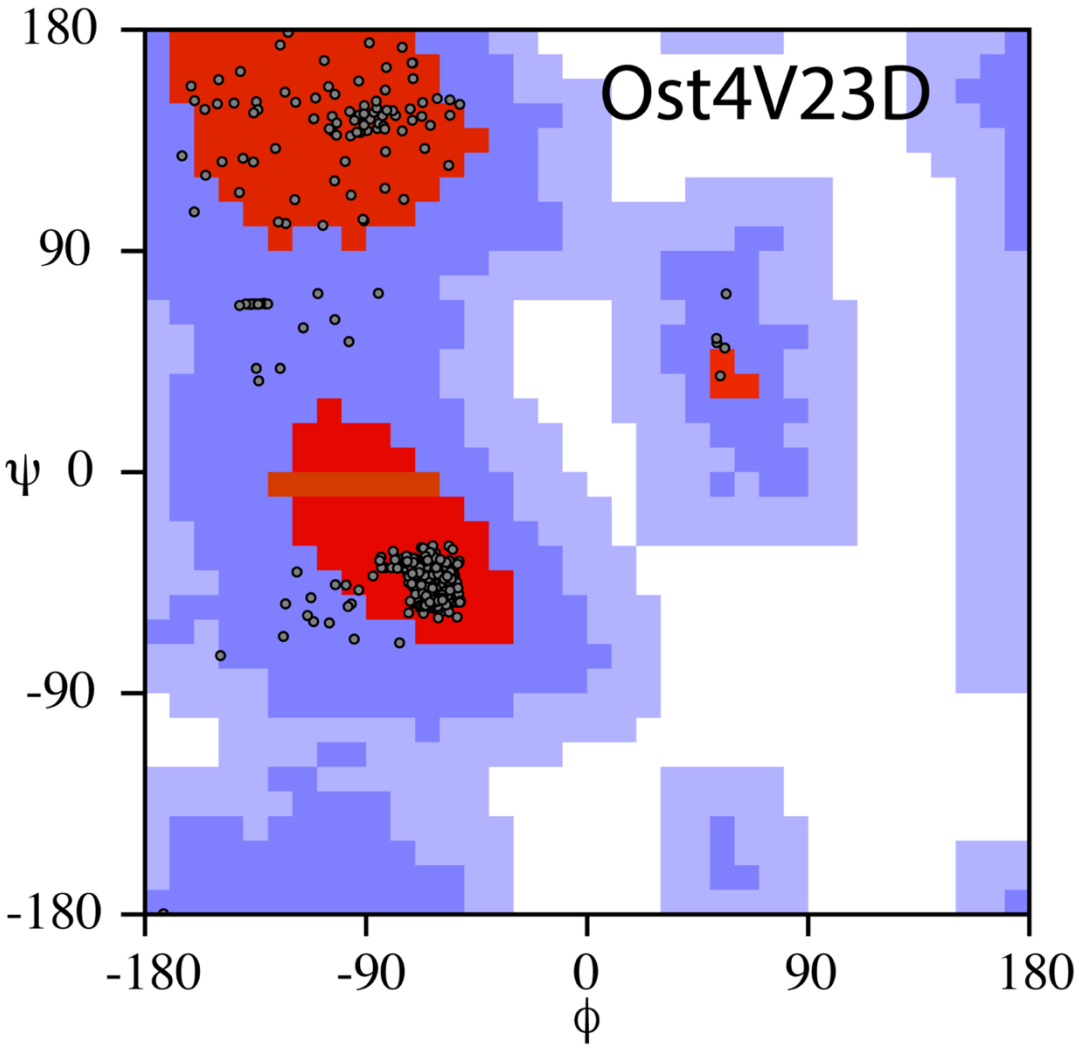


Figure 4.9: Ramachandran plot of Ost4V23D. This plot validates the energetically allowed regions of dihedral angles of protein residues in a protein.

visualized and validated by using the Ramachandran plot. The prediction results from the TALOS+ (Figure 4.2) was opened by invoking the Ramachandran plot visualization script rama+ incorporated into an interactive display and refinement of predictions software RAMA+. The display window of the Ramachandran plot, random coil index (RCI) and secondary structure, residue-wise dihedral angle result, and secondary shit distribution was opened by using rama+ command (Figure 4.10). The residue-wise dihedral angles were visualized and validated by selecting the residues using residue window (Figure 4.10b). The residues containing good, ambiguous, dynamic, and bad dihedral angles were classified by using this window (Figure 4.10a to 4.10e). The dihedral angles that were in the allowed region of the plot were selected for the 3D structure calculation. All the distance restraints used for structure calculation were derived from a combination of different NOE experiments such as 3D ^{15}N – edited HSQC – NOESY with 90 ms mixing time and 3D ^{15}N – edited HSQC – NOESY with 120 ms mixing time.

All these NOE restraints along with the dihedral angle restraints enabled us to calculate the 3D structures of Ost4 and Ost4V23D. The ensemble of the 20 lowest energy structures of Ost4 (Figure 4.11) and Ost4V23D (Figure 4.12) were very tight with a backbone RMSD of 0.24 Å and 0.20 Å, respectively. The structures were independently calculated using different NOE restraint sets for each protein. Surprisingly, 3D structures of both Ost4 (Figure 4.13) and its critical mutant Ost4V23D (Figure 4.14) were found to be nearly identical in DPC micelles with a backbone RMSD of 0.75 (Figure 4.15) for all the residues. Each of them had a single α-helix spanning residues Asp^{4}-Met^{32}. The N-terminal residues Met^{1}-Ser^{3} and the C-terminal residues Ser^{33}-Asn^{36} were random coils.

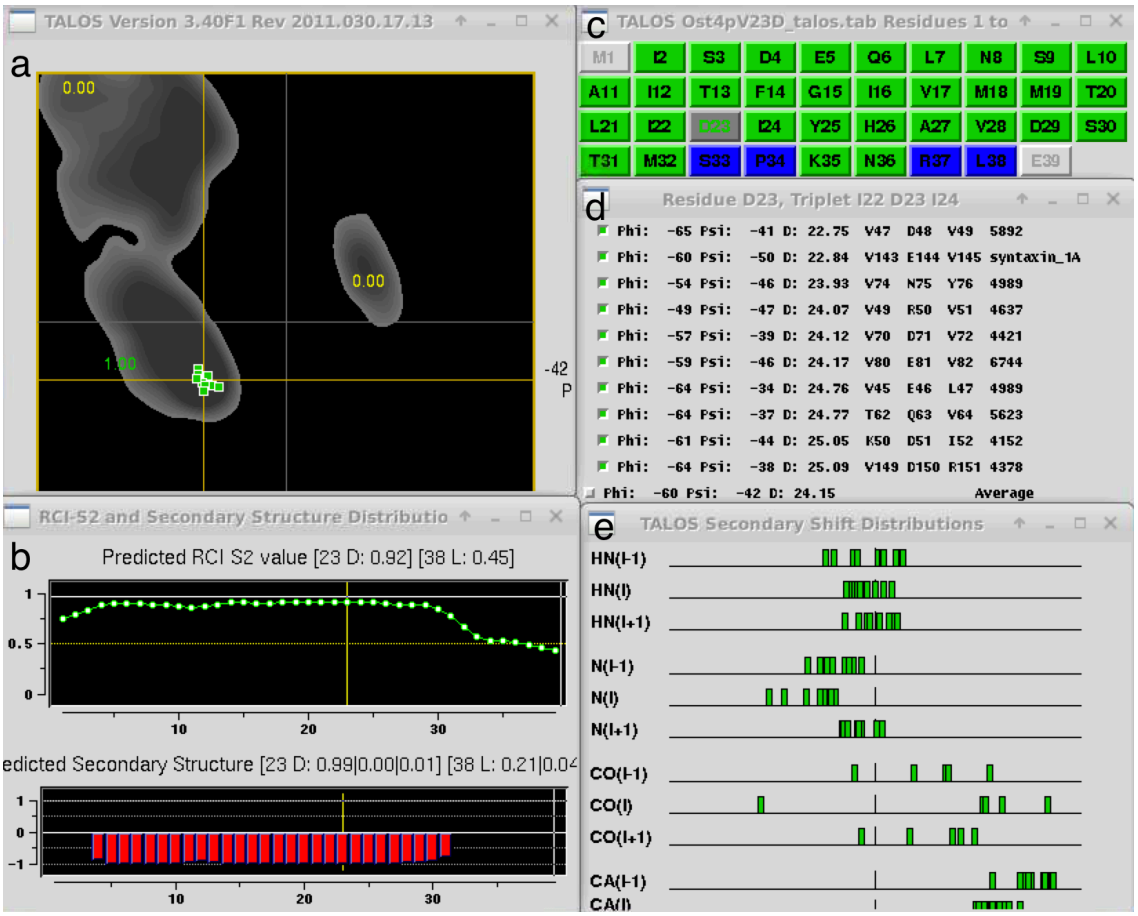


Figure 4.10: The interactive display and prediction window of the RAMA+ displaying the Ramachandran plot (a), predicted RCI and secondary structure (b) residue wise selection window (c), predicted dihedral angle results for the triplet (d), and TALOS+ secondary shift distributions for the selected residue.

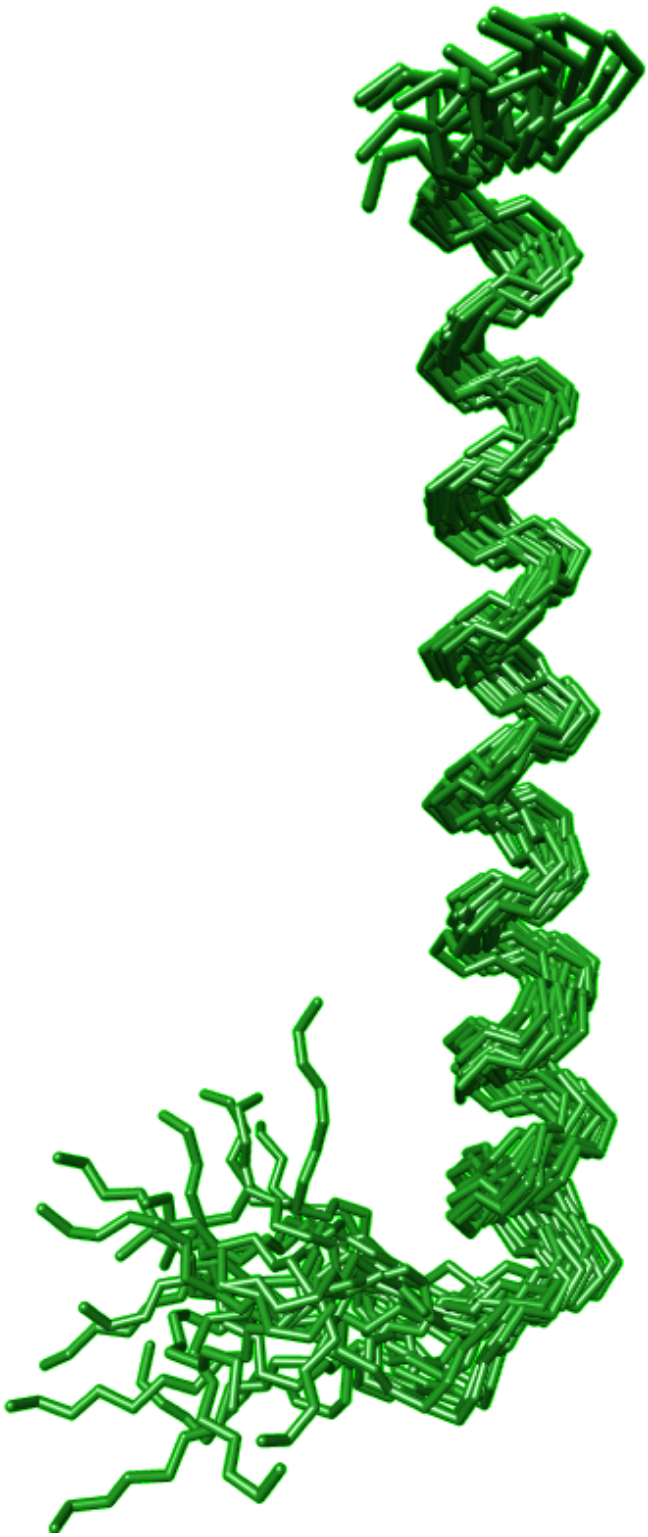


Figure 4.11: Ensembles of the 20 lowest energy NMR structures of Ost4 after solvent refinement.

Figure 4.12: Ensembles of the 20 lowest energy NMR structures of Ost4V23D after solvent refinement.

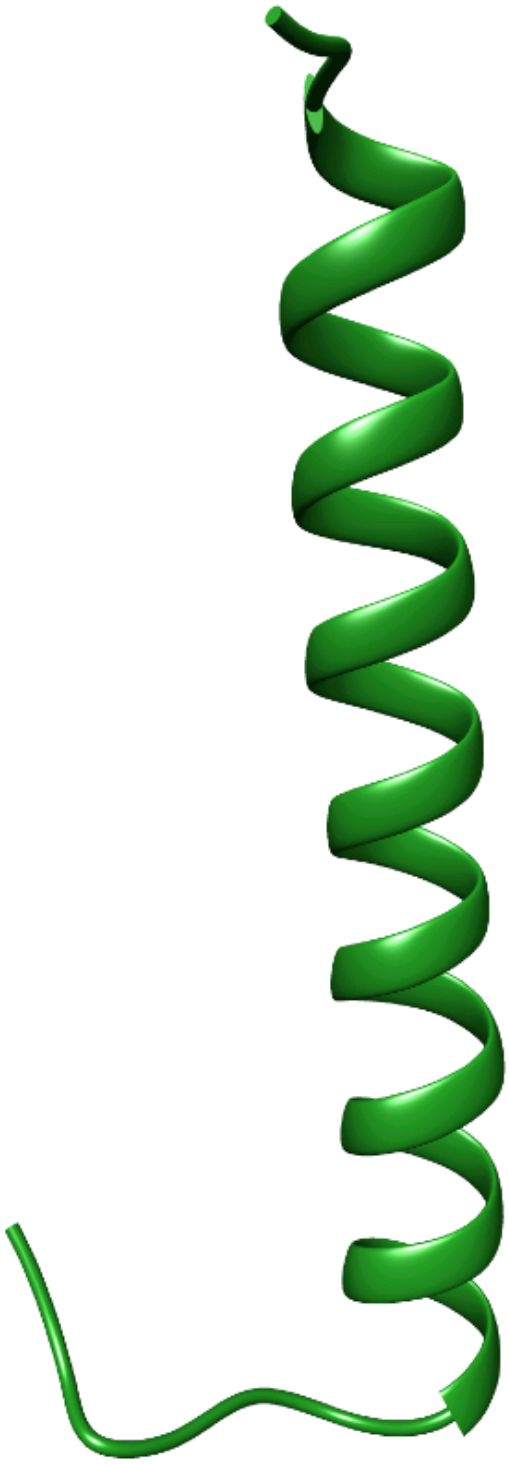


Figure 4.13: 3D structure of Ost4 protein in DPC micelles. It contains a straight single helix encompassing residues Asp^{4}-Met^{32}.

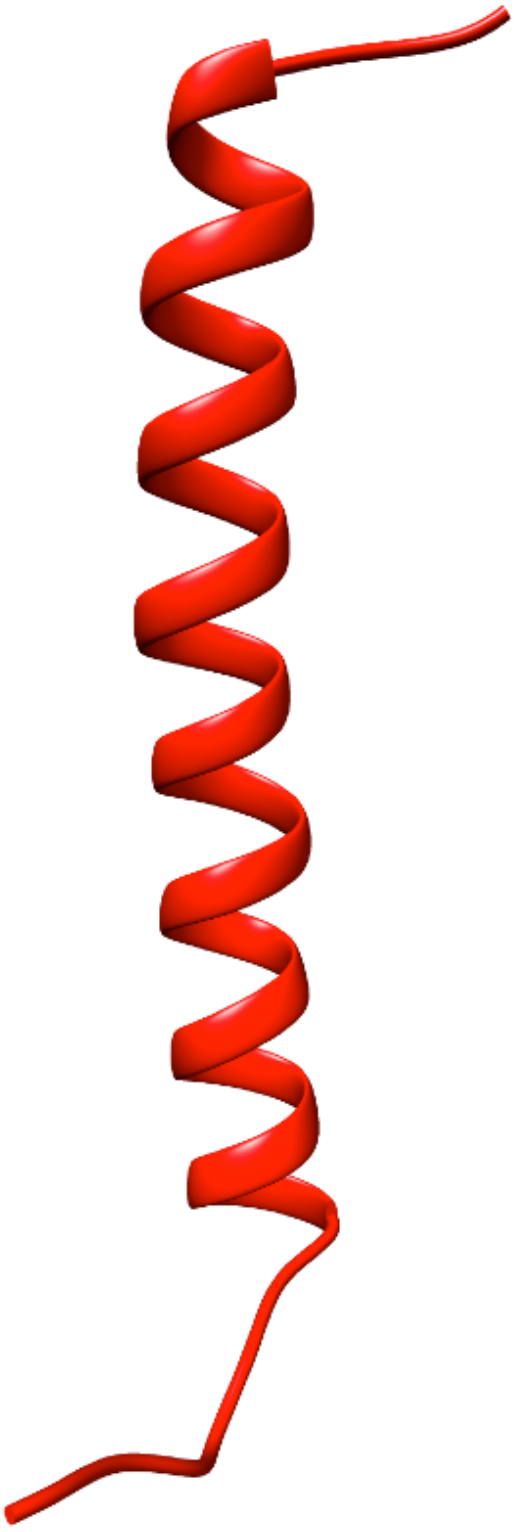


Figure 4.14: 3D structure of Ost4V23 protein in DPC micelles. It contains a straight single helix encompassing residues Asp^{4}-Met^{32}.

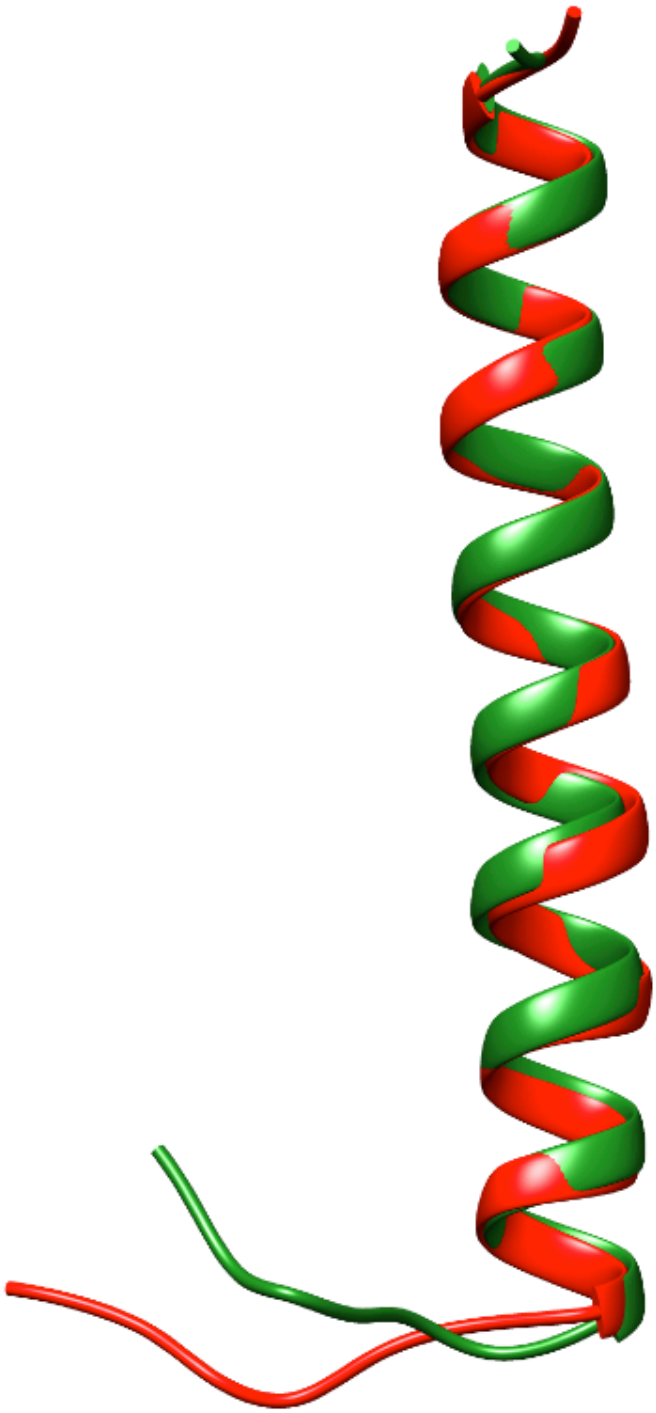


Figure 4.15: Overlay of the 3D structure of Ost4 (green) with that of Ost4V23D (red). Their 3D structures fit on each other with an RMSD of 0.75 Å for all residues.

4.3.3 Comparison of Ost4 structures determined in different membrane mimetic systems

The solution NMR structures of yeast and human Ost4 proteins, determined on chemically synthesized proteins in chloroform, methanol, and water solvent system, were reported to contain a kink in the transmembrane domain of the helix (Figure 4.16) (*37, 38*).

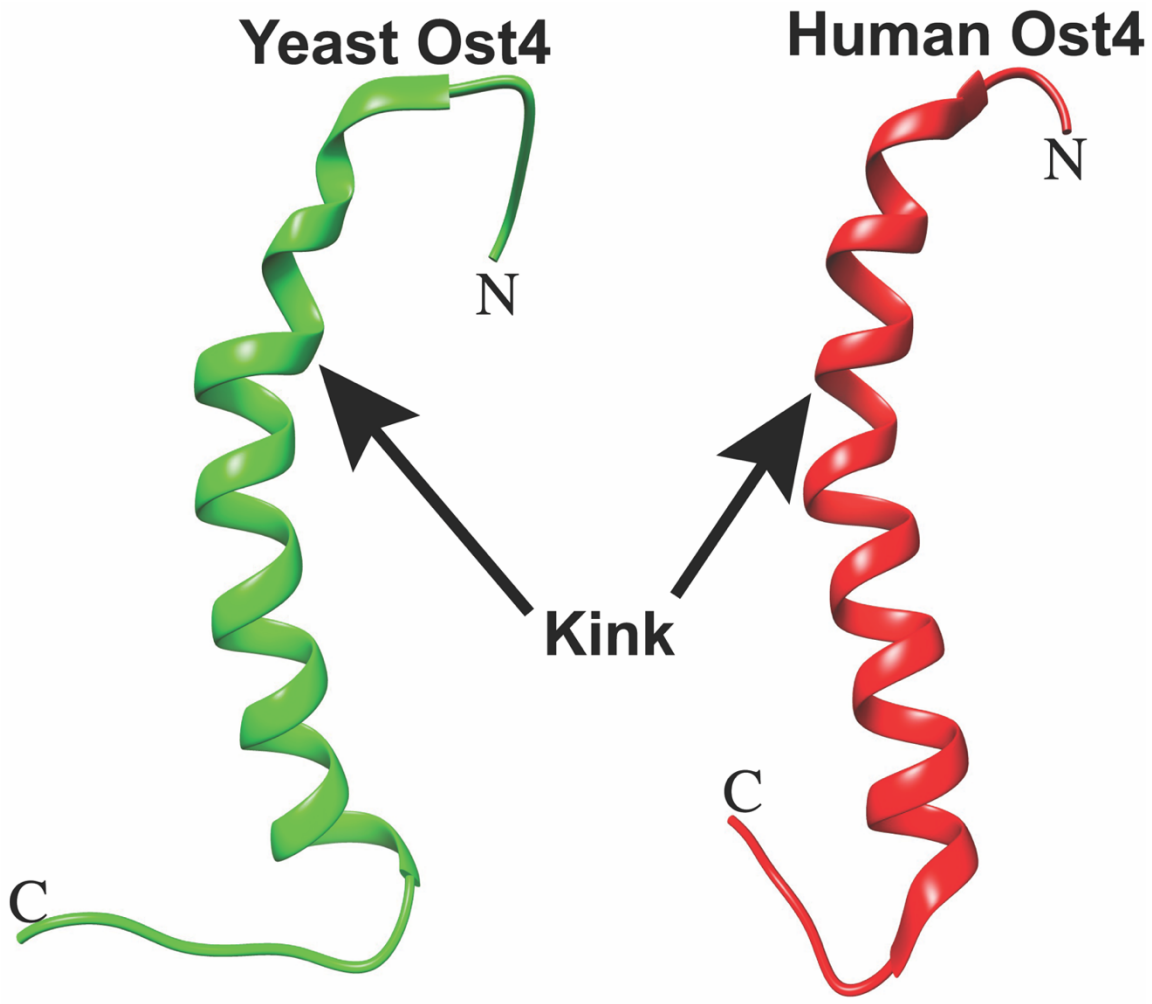


Figure 4.16: NMR structures of chemically synthesized yeast Ost4 (PDB ID 1RKL) and human Ost4 (PDB ID 2LAT) in mixed aqueous organic solvents. Both the structures have a kinked helix.

The NMR structure of the recombinant Ost4 determined in DPC micelles, reported here, has an RMSD (helical region) of 2.92 Å and 2.47 Å relative to the NMR structure of yeast Ost4 and human Ost4, respectively. The kinks in the structures of chemically synthesized yeast and human Ost4 proteins can be observed when compared to the NMR structure in DPC micelles reported in this dissertation (Figure 4.17 a and b).

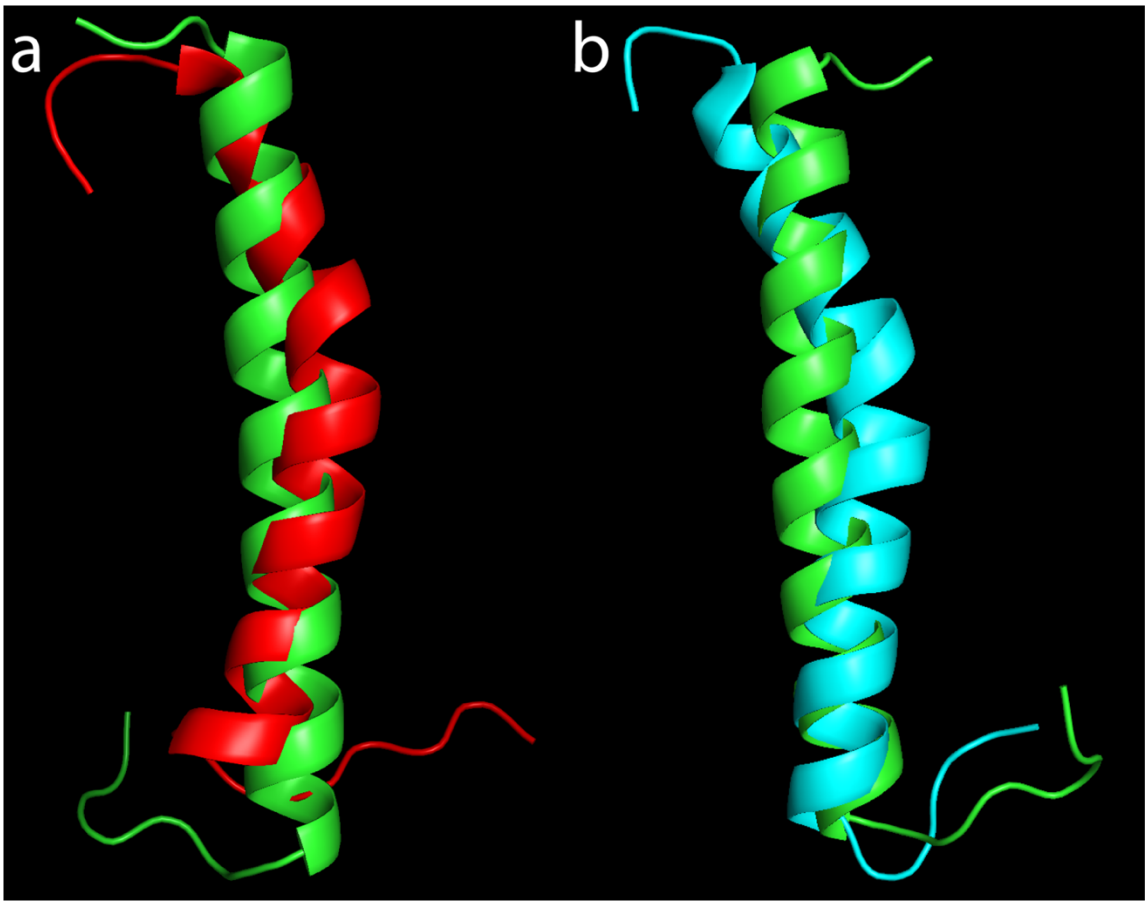


Figure 4.17: Overlay of NMR structure of Ost4 in DPC micelles (green) with solution NMR structure of yeast Ost4 shown in red (a) and human Ost4 shown in cyan (b) in a mixed aqueous organic solvent. NMR structure of yeast Ost4 and human Ost4 proteins show a kink in a mixed aqueous organic solvent.

The structure of the Ost4 subunit from the OST complex determined by the cryo-EM method in either nanodisc or digitonin (*34, 39*) (Figure 4.18) is depicted to be a single straight helix.

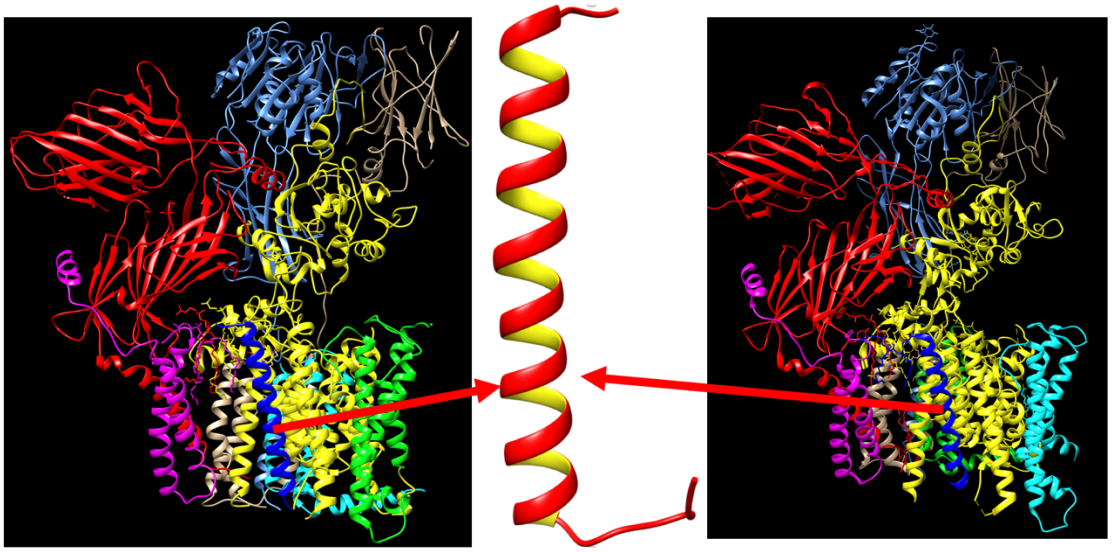


Figure 4.18: Cryo-EM structures of yeast OST in nanodisc (PDB ID 6EZN) (left panel) and yeast OST in digitonin (PDB ID 6C26) (right panel) depicting Ost4 (middle) as a single straight helix.

The Ost4 structure in both cases lacks the kink. The Ost4 structure in DPC micelles reported here, aligns very well to the Ost4 structure in yeast OST complex with an RMSD (helical region) of 0.68 and 1.04 Å determined in nanodisc and digitonin, respectively (Figure 4.19) (*23, 34, 39*). Recently high-resolution cryo-electron microscopy structures of human OST-A and OST-B complexes have been reported (*40*). The structures of the Ost4 subunit from both isoforms of human OST, OST-A and OST-B, showed a single helix (Figure 4.20). In both the isoforms, the Ost4 structures lack the kink. The helical region of

the Ost4 structure, in this report, aligns with the Ost4 subunit from the human OST-A complex and OST-B with an RMSD of 0.46 Å and 0.76 Å, respectively (Figure 4.21).

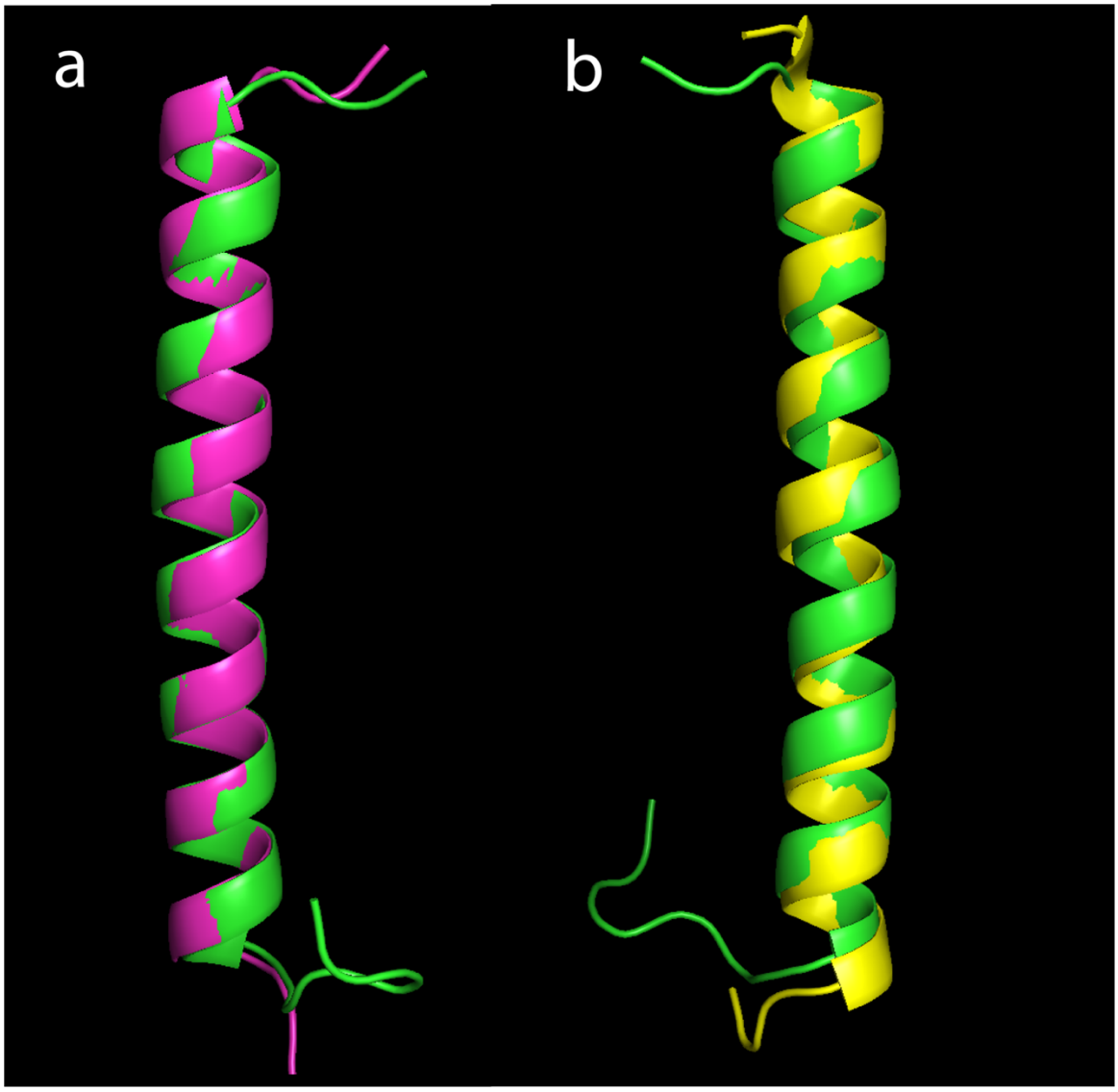


Figure 4.19: Overlay of the NMR structure of Ost4 in DPC micelles (green) with (a), the Ost4 structure (purple) from the cryo-EM structure of yeast OST determined in nanodisc and (b), the Ost4 structure (yellow) from the cryo-EM structure of yeast OST determined in digitonin. The Ost4 structures from yeast OST complex determined either in nanodisc or in digitonin fit well with the Ost4 structure reported in this dissertation.

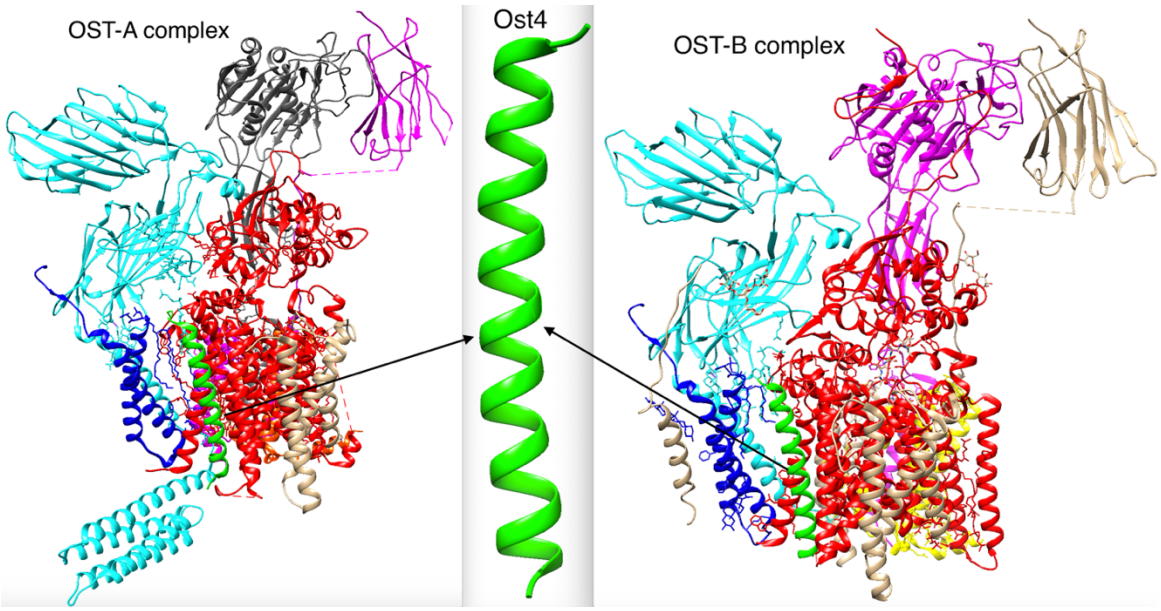


Figure 4.20: Cryo-EM structures of human OST-A (PDB ID 6S7O) and OST-B (PDB ID 6S7T) showing the Ost4 subunit as a single straight helix. These structures were generated by using chimera.

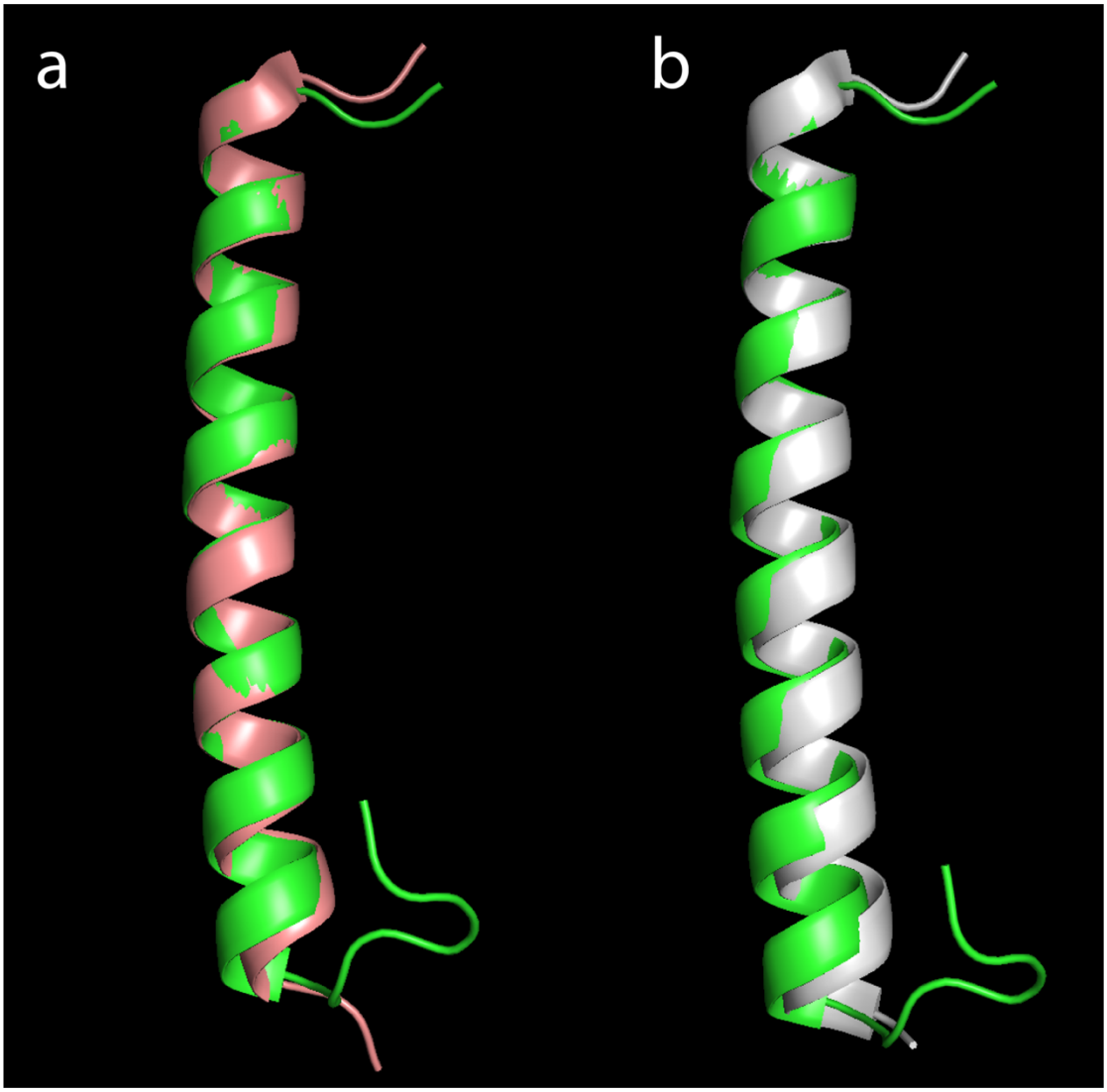


Figure 4.21: Overlay of the NMR structure of Ost4 in DPC micelles (green) with (a), the Ost4 structure (purple) from the cryo-EM structure of human OST-A complex and (b), the Ost4 structure (white) from the cryo-EM structure of human OST-B complex. The Ost4 structures from yeast OST complex determined either in nanodisc or in digitonin fit well with the Ost4 structure reported in this dissertation.

4.3.4 Analysis of Ost4 and Ost4V23D dynamics data

The information about the dynamics of residues in Ost4 and Ost4V23D proteins were obtained from the ^{15}N T_1, T_2, and steady-state ^{1}H – ^{15}N heteronuclear NOE relaxation experiments. In Ost4, the T_1 values and ^{1}H – ^{15}N heteronuclear NOE, for the majority of residues, are relatively higher and T_2 values are relatively lower than that of Ost4V23D protein (Figure 4.22 a, b and c), indicating that the WT protein is more rigid. In Ost4V23D, the motions of residues change revealing increased flexibility of the residues upon V23D mutation. The average T_1 and T_2 values were 1125.7 ms and 66.7 ms for Ost4 and 907.5 ms and 78.2 ms for Ost4V23D, respectively.

The average correlation times (τ_c) of Ost4 and Ost4V23D proteins were calculated from the NMR T_1/T_2 ratios by using the Tensor2 program. A decrease of overall τ_c from a value of 12.6 ns for Ost4 to a value of 10.5 ns for Ost4V23D was observed (Figure 4.23 a and b). This indicates a faster tumbling of Ost4V23D compared to Ost4 in DPC micelles. Based on τ_c values compiled for known monomeric Northeast Structural Genomics Consortium (NESG) targets (Table 4.2) (*41*), the molecular weights of Ost4 and Ost4V23D proteins in DPC micelles are approximately 21 kDa and 17 kDa, respectively. The observed decrease in molecular weight of the V23D mutated protein in DPC micelle indicates the protein could be in association with a smaller number of DPC molecules than in the WT protein.

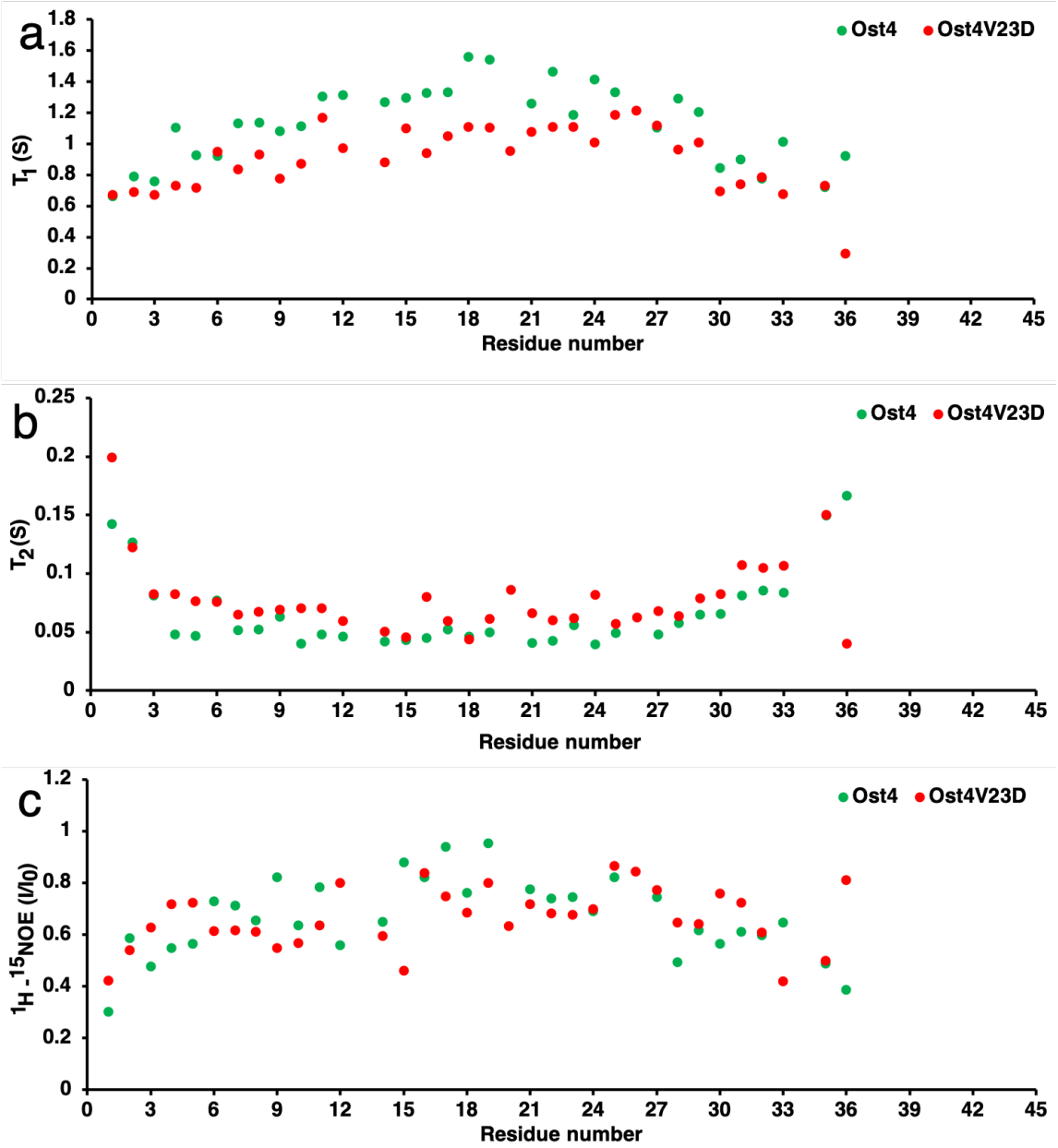


Figure 4.22: Comparison of (a) T_1 values, (b) T_2 values, and (c) $^{1}H - ^{15}N$ heteronuclear NOE of Ost4 (green) and Ost4V23D (red).

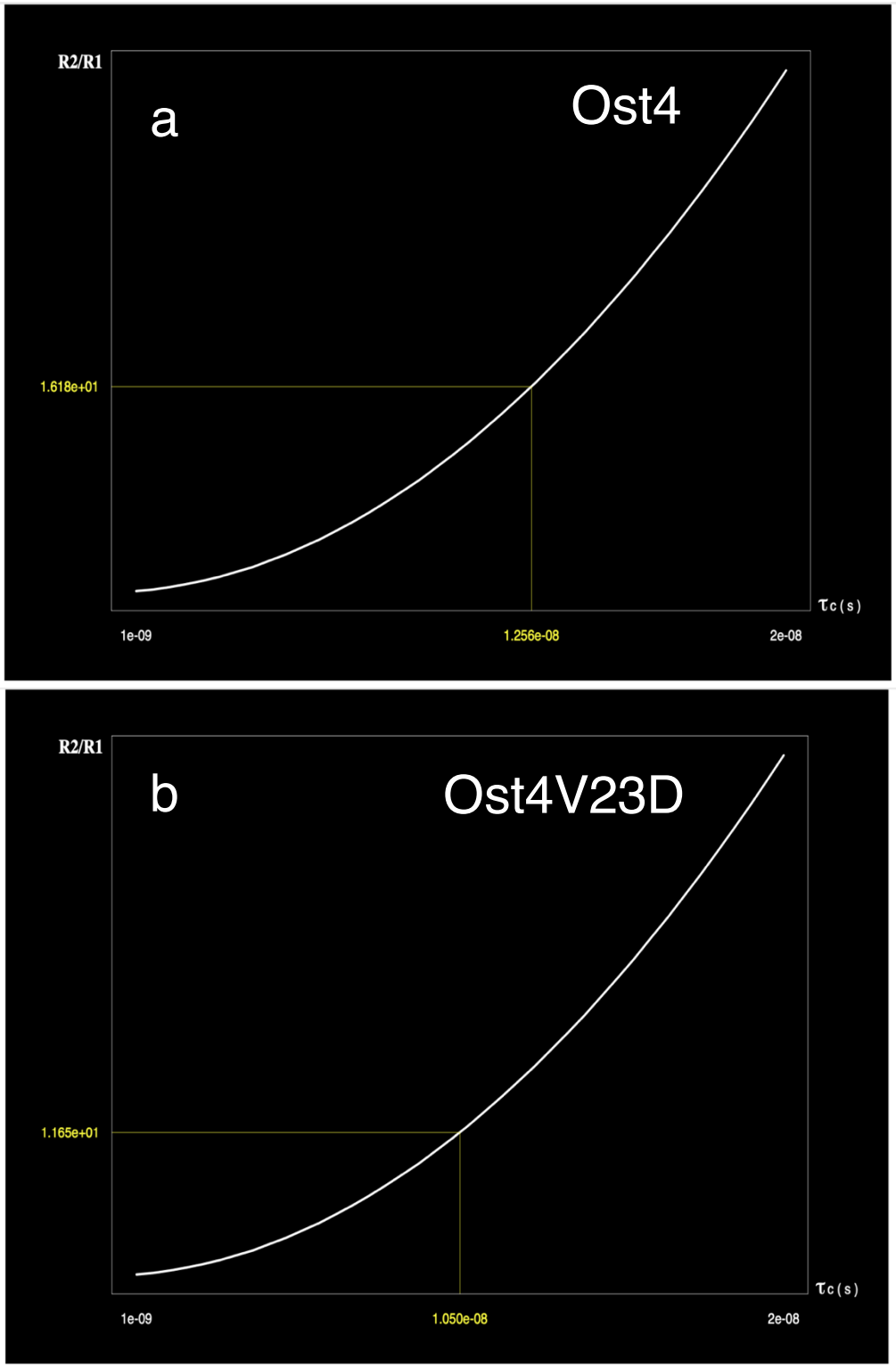


Figure 4.23: Plots of correlation time *vs* R_1/R_2 ratio of Ost4 (a) and Ost4V23D (b) obtained from Tensor2 program.

Table 4.2: Rotational correlation time values compiled for known monomeric NESG targets. The table was adapted from reference (*41*).

NESG target (isotope labeling)	**MW (kDa)**	**^{15}N T_1 (ms)**	**^{15}N T_2 (ms)**	**τ_c (ns)**
PsR76A (NC5)	7.2	478	128	5.1
VfR117 (NC)	11.2	605	119	6.3
SyR11 (NC5)	12.4	630	104	7.1
ER541-37-162 (NC5)	15.8	729	66.5	10
ER540 (NC5)	18.8	909	66.5	11.3
SoR190 (NC)	13.8	697.5	100.9	7.7
TR80 (NC5)	10.5	612.8	102.9	7
Ubiquitin (NC)	9	441.8	144.6	4.4
HR2873B (NC)	10.7	492	115	5.7
B-domain (NC)	7.2	423.5	153.3	4.05
BcR97A (NC)	13.1	705.8	80.6	8.8
PfR193A (NC)	13.6	733.9	80.9	9
MvR76 (NC)	20.2	1015	64.5	12.2
DvR115G (NC)	10.9	608.7	115.6	6.5
MrR110B (NC5)	11.8	707	99.2	7.8
VpR247 (NC5)	12.5	661.2	88.3	8.05
BcR147A (NC)	11.9	645	104	7.2
WR73 (NC5)	21.9	1261	41.3	13
NsR431C (NC5)	16.8	855.5	71.2	10.6
StR82 (NC)	9.2	537.3	100.4	6.6

4.3.5 Molecular dynamics simulation and behavior of proteins in the membrane

To understand the behavior of wildtype Ost4 and mutant Ost4V23D in the membrane, we performed MD simulation for both the proteins separately in DPC micelles. The MD simulations for both the proteins were carried out under identical conditions. The initial system for both contains protein partly inserted into preformed DPC micelles of 65 molecules and surrounded by 6729 molecules of TIP3P water (Figure 4.24). Both the systems were stable within the time scale of the simulation. Both the systems, Ost4-DPC and Ost4V23D-DPC, were energy minimized. The energy minimization was successful as indicated by negative potential energy in the order of 10^5-10^6 for both the systems (Figure 4.25). The energy minimization step was followed by equilibration under constant temperature (Figure 4.26) and constant pressure conditions (Figure 4.27). The average backbone RMSDs of the proteins indicated that the WT protein attained equilibrium after 40 ns (Figure 4.28a) while Ost4V23D reached the equilibrium after 110 ns of MD simulation (Figure 4.28b). The RMSD of DPC molecules relative to the system was

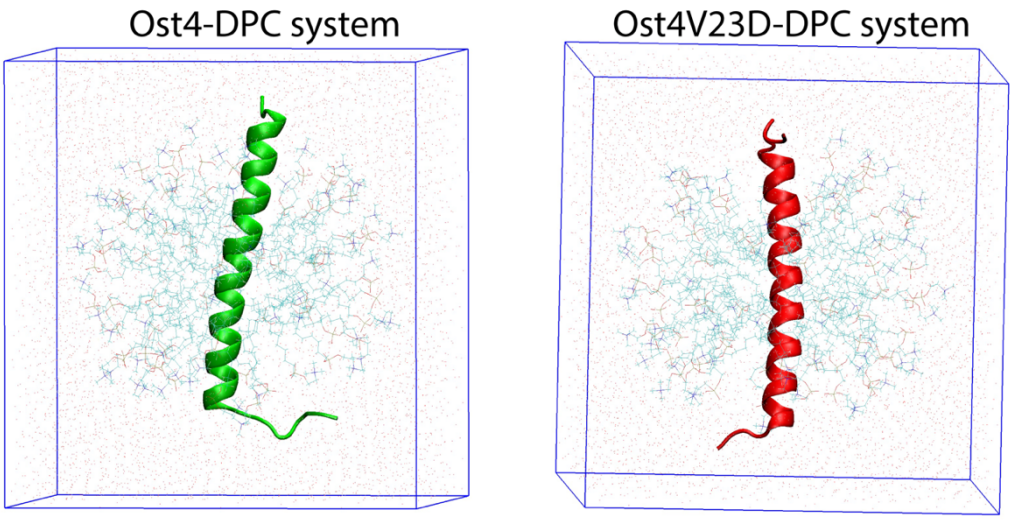


Figure 4.24: The initial system of Ost4 (left) and Ost4V23D (right) surrounded by 65 DPC molecules (cyan lines) and solvated with 7629 molecules of water (red dots).

calculated to confirm whether the DPC molecules reached the equilibrium state in the MD simulation. The RMSD plot (Figure 4.29 a and b) shows that not only protein but also the DPC molecules reached the equilibrium in MD simulation. For comparative study, the trajectories 110 ns to 150 ns were used for analyses. Although the NMR structures of Ost4 and Ost4V23D are nearly identical in DPC micelles, MD simulations of these proteins in

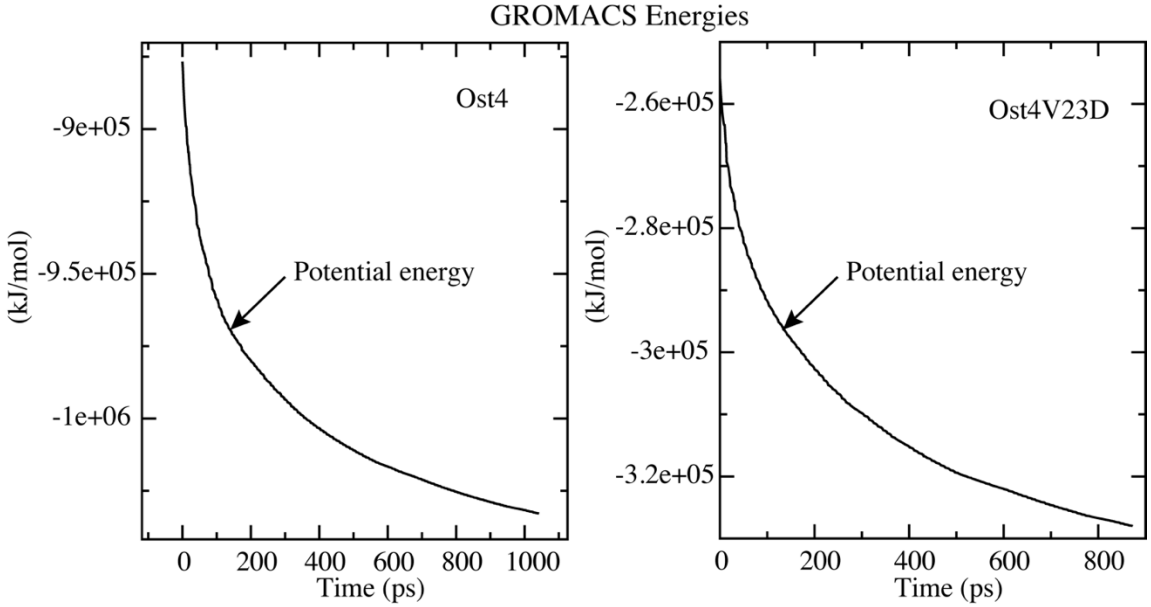


Figure 4.25: Energy diagram for energy minimization of Ost4-DPC (left) and Ost4V23D-DPC system (right). The negative potential energy on the order of 10^5-10^6 kJ/mole is an indication of a successful energy minimization step.

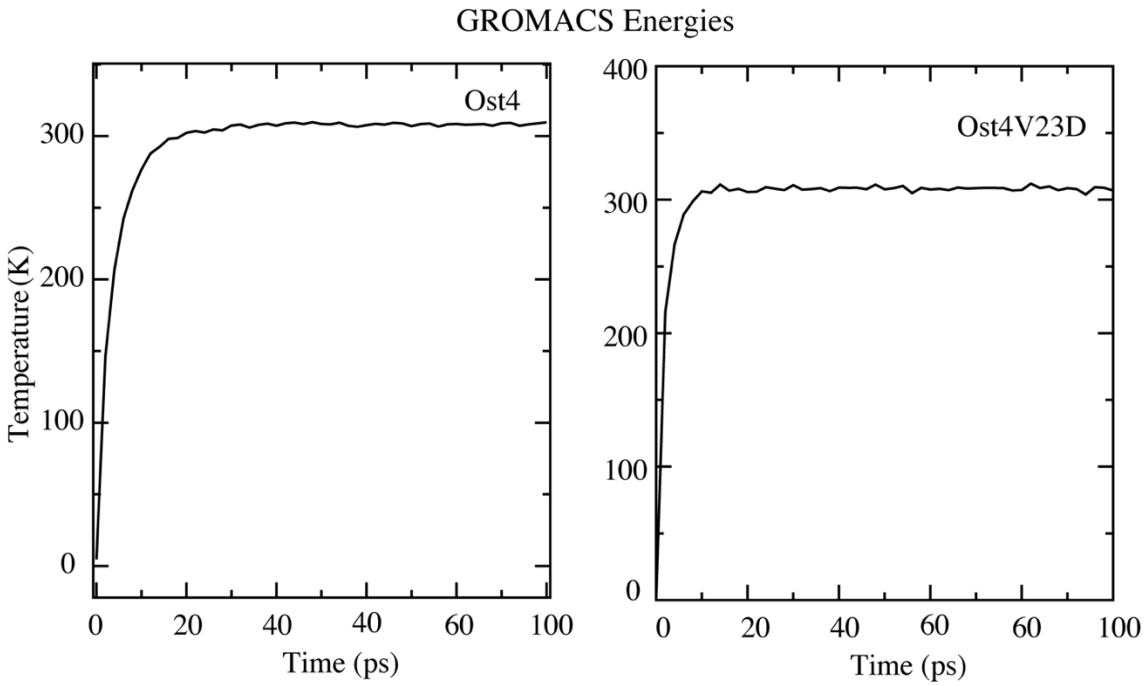


Figure 4.26: Energy diagram of equilibration under constant number, volume and temperature (NVT) condition. The NVT equilibration was carried out at 35 ºC.

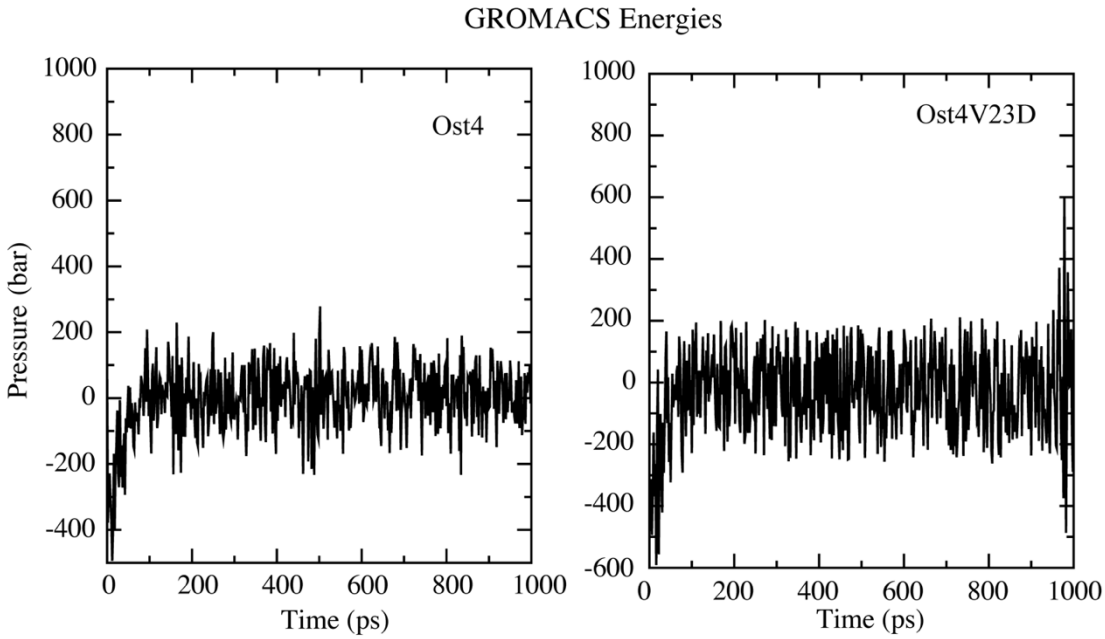


Figure 4.27: Energy diagram of equilibration of the Ost4-DPC (left) and Ost4V23D (right) systems under a constant number of molecules, pressure, and temperature conditions. This NPT equilibration was carried out at 1 bar pressure for 1 ns time.

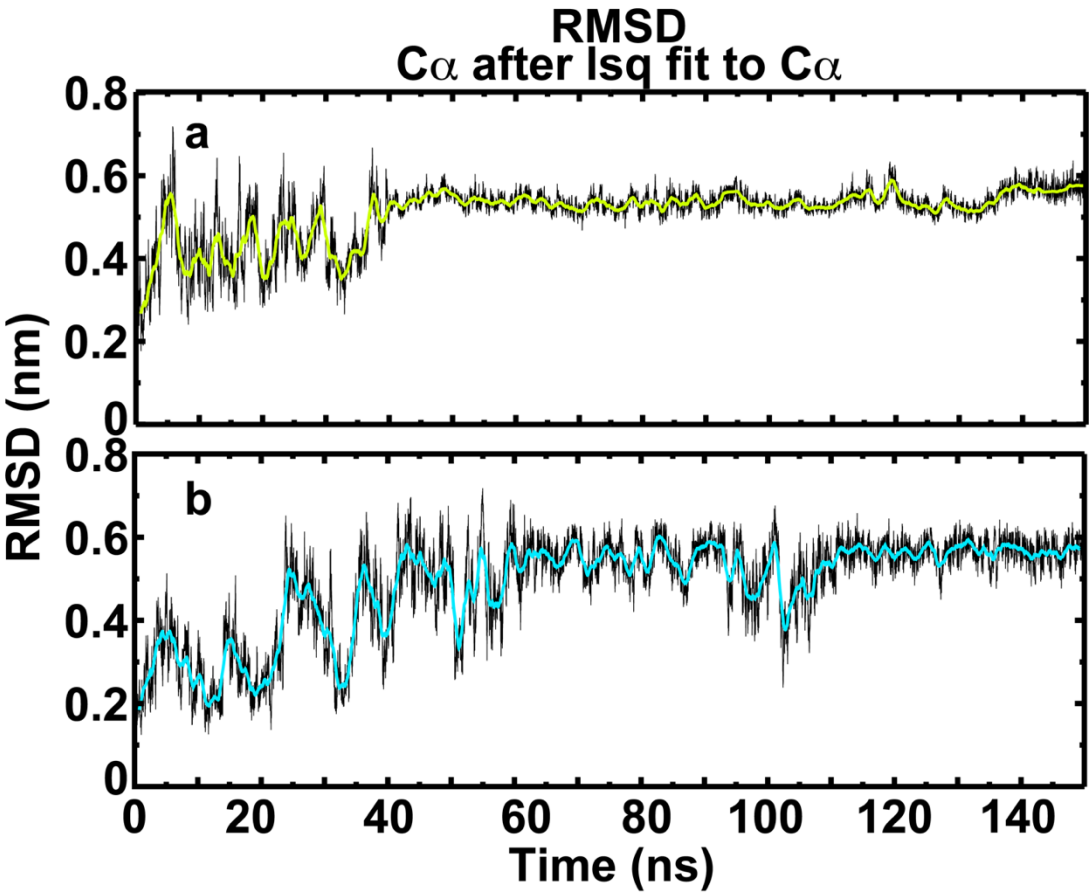


Figure 4.28: Backbone RMSD fluctuation of Cα of (a) Ost4 and (b) Ost4V23D showing the simulation attained equilibrium during a 150 ns MD run. For both the proteins, the MD was carried out by inserting proteins in 65 DPC micelles.

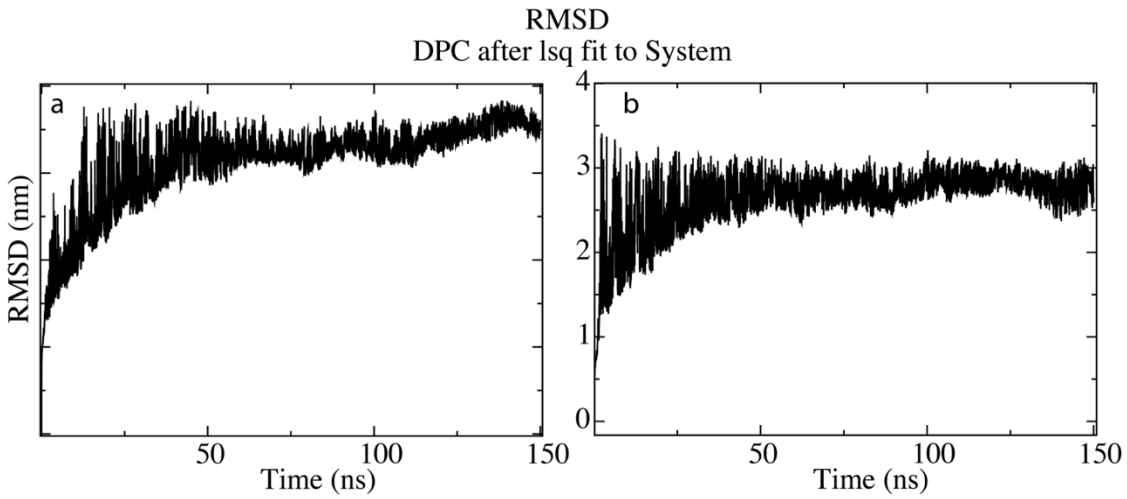


Figure 4.29: RMSD fluctuation of DPC micelles in (a) Ost4-DPC and (b) Ost4V23D-DPC systems showing the DPC molecules were at equilibrium during the 150 ns MD run.

DPC micelles resulted in an interesting outcome. At the end of the MD run, the WT Ost4 remained surrounded by DPC molecules where the protein sits at the core of micelles (Figure 4.30 a and b).

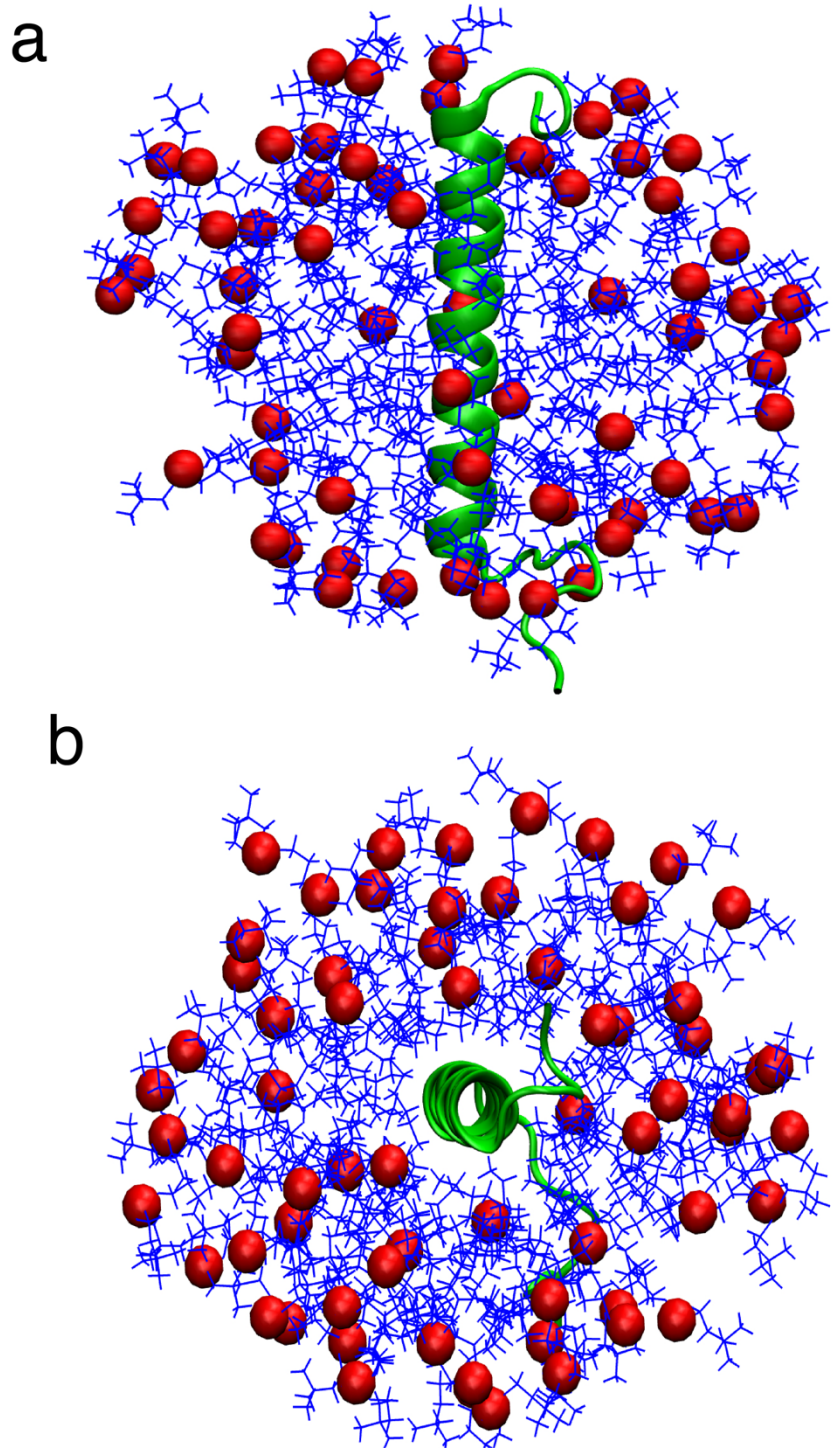


Figure 4.30: MD simulation results (a) side view and (b) top view of Ost4 in DPC micelles depicting protein inserted at the center of micelles.

Interestingly, mutant (Ost4V23D) protein showed an unexpected behavior. The mutant protein settled in such a way that the DPC micelles did not surround the protein from all sides. At the end of the MD simulation, the mutant protein was exposed to the solvent on one side while the other side was surrounded by DPC molecules (Figure 4.31 a and b).

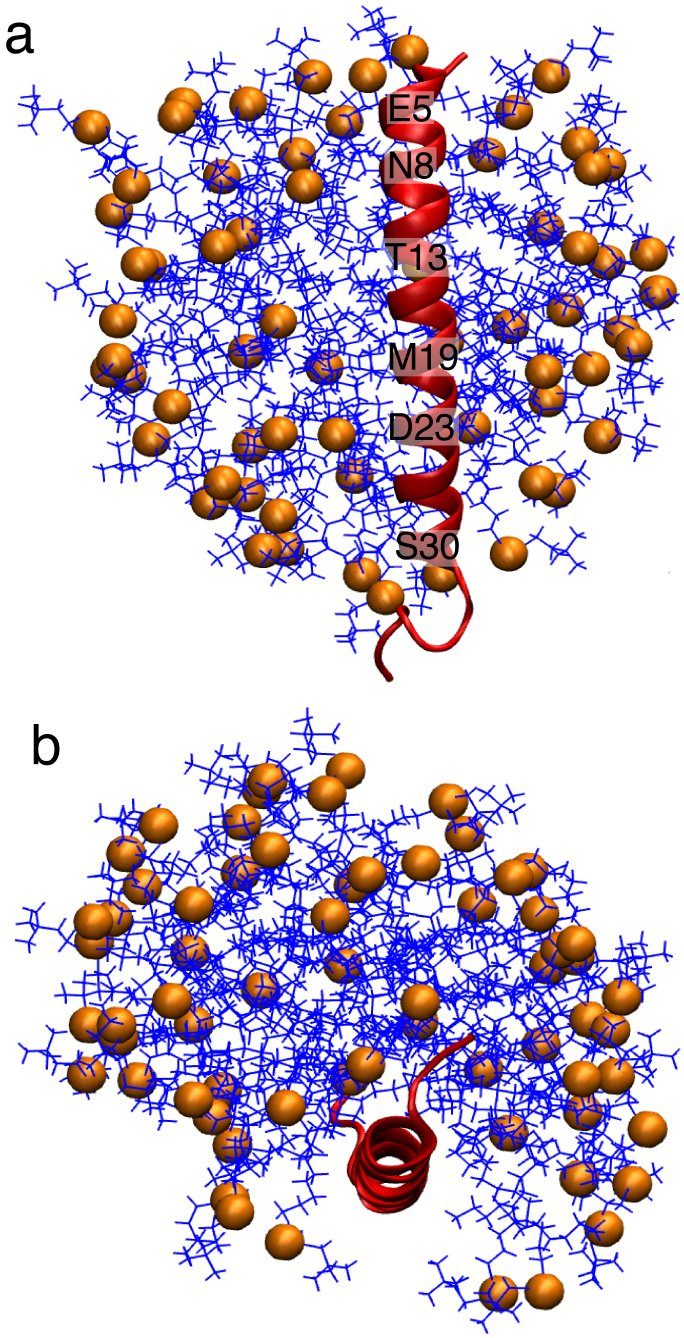


Figure 4.31: MD simulation results (a) side view and (b) top view of Ost4V23D in DPC micelles showing the protein to be moved towards the interface of micelles and solvent. Analysis of residues shows that the hydrophilic residues expose to the hydrophilic solvents.

The residues Leu^{10}-Val^{28} of the α-helix inserted in the DPC micelles exposing N-terminal Met^{1}-Ser^{9} residues and C-terminal Asp^{29}-Glu^{39} residues out to the solvent (Figure 4.32). A more detailed analysis of the simulation results revealed that the hydrophilic residues of the helix are exposed to solvent while hydrophobic residues are buried in the micellar environment (Figure 4.31 a).

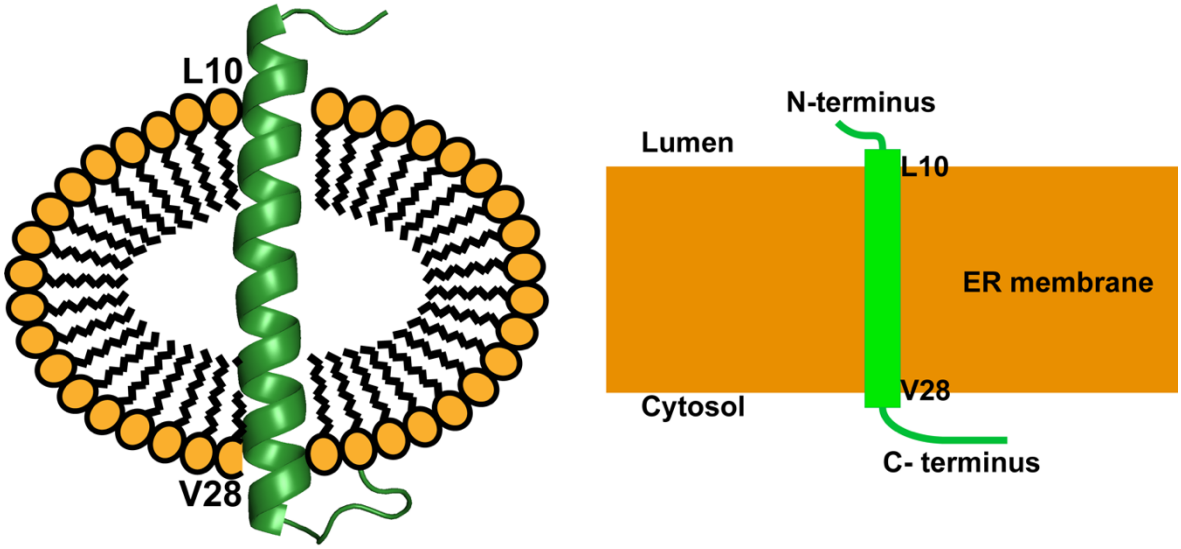


Figure 4.32: Schematic representation of the position of Ost4 in the DPC micelle (left) and membrane (right).

The Solvent Accessible Surface Area (SASA) of transmembrane residues (Leu^{10}-Val^{28}) of both proteins were quantified. SASA provides the surface area of the solvent-exposed residues and is defined as the ratio between the water exposed surface area and the total surface area of a given section of the peptide. The SASA fraction of hydrophobic and

hydrophilic residues of Ost4 were 0.0662 ± 0.0003 and 0.1718 ± 0.0009, respectively. Similarly, the SASA fraction of hydrophobic and hydrophilic residues of Ost4V23D were 0.2023 ± 0.0007, and 0.428 ± 0.001, respectively (Figure 4.33). The hydrophobic and hydrophilic SASA of Ost4 were smaller than that of Ost4V23D. Additionally, the number of interactions between DPC tails and residues Leu^{10}- Val^{28} were calculated for both proteins. About 64 ± 12 contacts were found between Ost4 residues Leu^{10}-Val^{28} and the DPC tails, whereas only 16 ± 6 contacts were observed between the same region of the Ost4V23D protein and the DPC tails (Figure 4.34).

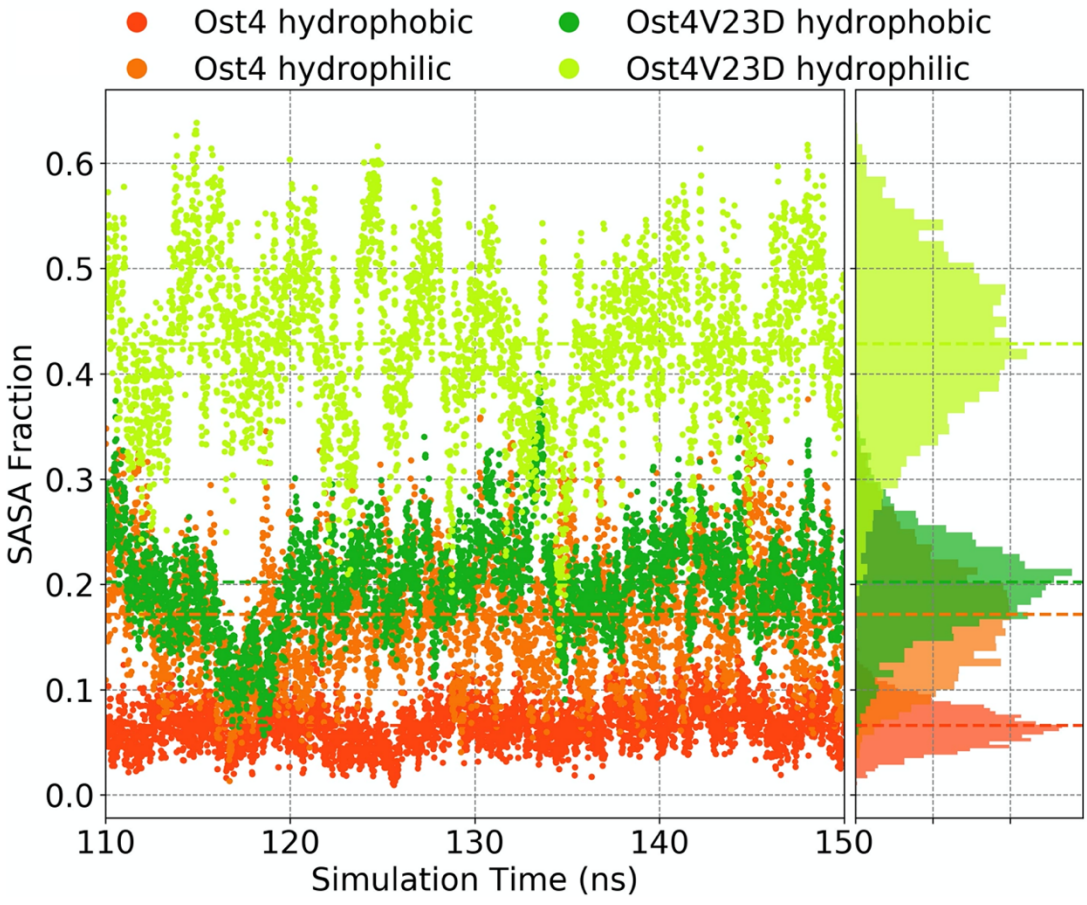


Figure 4.33: Solvent accessible surface area (SASA) of Ost4 and Ost4V23D. The hydrophilic and hydrophobic SASA of these proteins is labeled in the figure. The figure was adapted from reference (*23*).

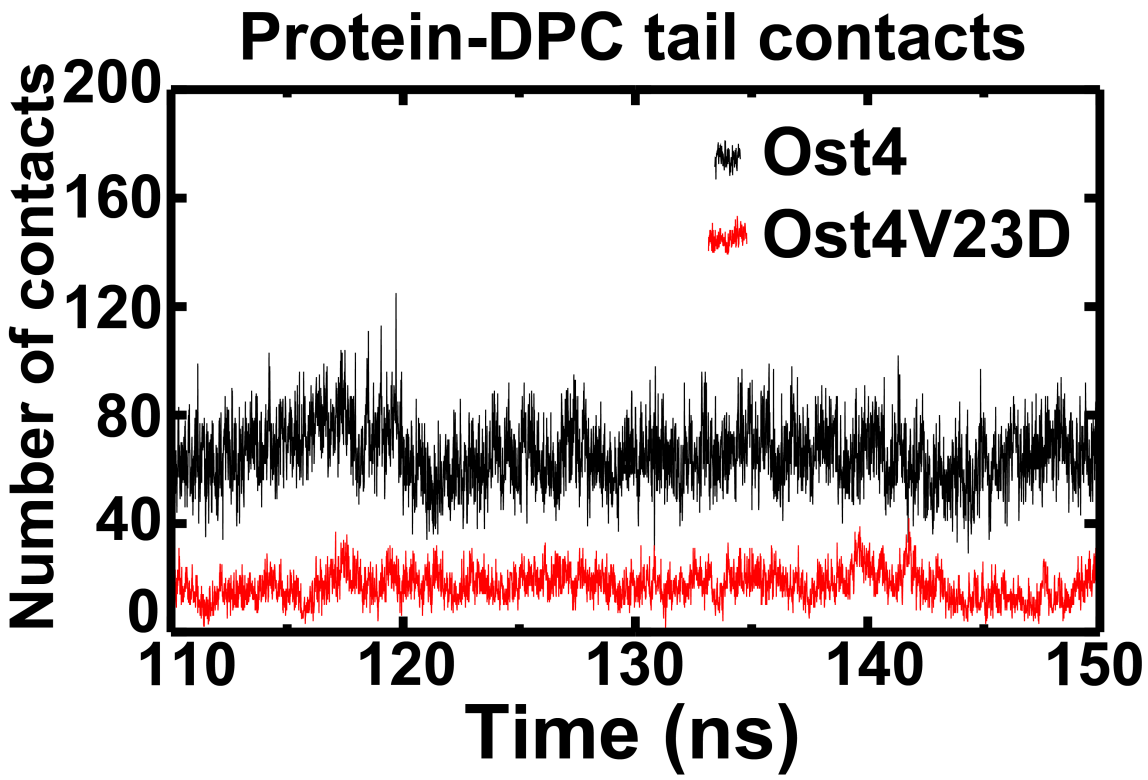


Figure 4.34: Number of DPC-tails that make contacts with the transmembrane domain of Ost4 and Ost4V23D.

4.3.6 Molecular dynamics simulation of membrane – bound OST complex

To investigate the structural perturbations of the Ost4V23D mutation, MD Simulations of membrane – bound OST complex containing TM helices Ost1, Ost3, Ost4/Ost4V23D,

Ost5, and the TMHs Stt3 other than TMH 9 were performed. The membrane-bound OST complex for the simulations was built by embedding the components from the CryoEM structure 6EZN (*34*) in a membrane of dipalmitoylphosphatidylcholine (DPPC) lipids. The simulation continued up to 600 ns. During the simulation, both the WT and Ost4V23D mutant remained embedded in the membrane. The analysis of MD trajectories of membrane-bound OST complex for WT simulation revealed several hydrophobic contacts between Ost4V23 (brown) with Stt3 TM12, residue F425 and TM13 (red), residues I456 and L459, and with lipid tails (cyan) (Figure 4.35a). A single contact between Ost4D23 and Stt3 I456 was maintained in the OstV23D. The formation of a new stable salt bridge with Stt3 K448 and the carboxylate functional group of the aspartate was observed in the mutant simulation (Figure 4.35b). The representative snapshots of the WT and mutant complex showed the local structural perturbations due to the Ost4V23D mutation (Figure 4.35a and b).

The increased hydrophilicity of the Ost4V23D mutant was quantified by calculating the average SASA of the V23 and D23 residue in the WT and the Ost4V23D, respectively (Figure 4.35c). The average SASA of D23 in the mutant was 0.15 nm^2 and that of V23 residue in the WT simulation was approximately zero indicating penetration of water in the mutant protein to a greater degree than in the WT. The increased solvent exposure in the mutant was due to the destabilization of the buried hydrophobic pocket as shown in Figure 4.35a. The hydrophobic contacts between the mutated D23 and the lipid tails were not observed.

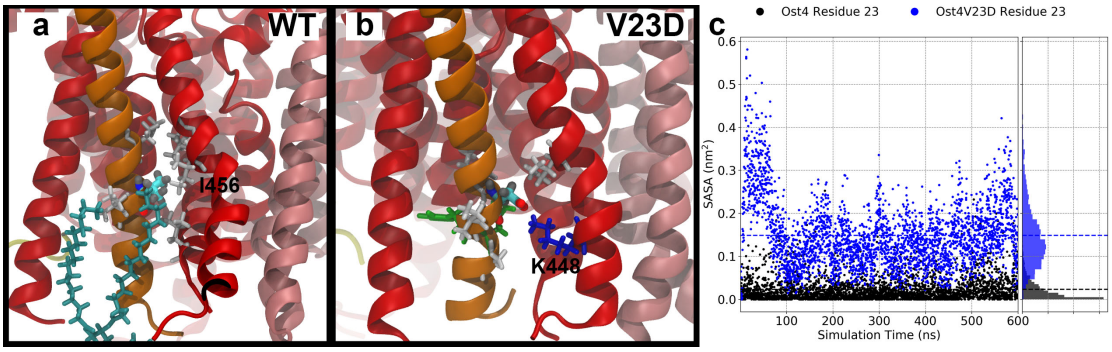


Figure 4.35: Molecular dynamics simulations of the OST complex in a dipalmitoylphosphatidylcholine (DPPC) membrane. (a) A representative snapshot of the WT simulation with highlighted Ost4V23 (brown) residue and associated hydrophobic contacts from Stt3 TM 12 and 13 (red), and lipid molecules (cyan). (b) Representative snapshot from Ost4V23D simulation with Ost4D23-Stt3 K448 salt-bridge highlighted. (c) SASA of residue 23 from 600 ns of simulation for both WT and mutant simulation. The Figure was adapted from reference (*23*).

4.4 Discussion

N-glycosylation is an essential and highly conserved post-translational protein modification in all domains of life. The high-resolution structures of both yeast and human OST complexes have been reported recently (*34, 39, 40*), however, the role of the requirement of the multiple subunits for this protein modification in eukaryotes is not yet clear. The major limitations have been the difficulty in the production of homogenous and pure recombinant proteins in milligram quantities along with their reconstitution in a suitable membrane mimicking environment. Once a suitable membrane mimetic system has been determined, it is possible to obtain atomic resolution structures of membrane proteins (*42*). In previous NMR structural studies of yeast and human Ost4, a mixture

containing 4:4:1 of chloroform, methanol, and water (*37, 38*) was used to mimic the membrane environment. Although such a solvent system is a marginal approximation of a non-polar membrane surrounded by an aqueous environment, phospholipid micelles are considerably better mimetics of biological membranes. They contain a completely nonpolar core region of a similar thickness to a membrane that is surrounded by a layer of charges at the water interface similar to that of the lipid bilayer. DPC is a zwitterionic detergent that has a structure similar to phospholipid bilayers and has been extensively used for NMR structure determination of membrane proteins (*43, 44*).

The isolation, purification to homogeneity, and initial NMR assignments of the smallest OST subunit, Ost4 and its critical point mutant, Ost4V23D have been discussed in Chapter 2 and 3. These chapters present the atomic resolution structures of recombinant wildtype Ost4 and its critical point mutant, Ost4V23D determined in DPC micelles with solution-state NMR . The effect of a point mutation that introduces a negative charge into the transmembrane helix of Ost4 is discussed. As shown in Figure 4.4, the mutation of Val^{23} to Asp in Ost4 caused chemical shift perturbation of almost all of the residues in the WT protein. The perturbation in the chemical shifts could be due to one or more reasons. The first possibility could be that the protein underwent a structural rearrangement. The second possibility could be due to a change in the chemical environment around the protein. The third possibility could be a structural rearrangement in a new chemical environment. Since the 3D structures of both Ost4 and Ost4V23D are nearly identical as calculated from completely separate NMR data sets, it became clear that the chemical environment around the mutant protein had changed. Each protein is composed of a single α–helix

encompassing residues Asp^{4}-Met^{32}. This structure is quite similar to the straight single helix that was reported for Ost4 as part of the OST complex solved in nanodisc using cryo-EM (*34, 39*). However, it was different than the structures of chemically synthesized Ost4 form both yeast and human that were solved in a mixture of chloroform/methanol/water (4:4:1) and was reported to contain a kinked helix (Figure 4.14) (*37, 38*). Ost4 structure in DPC micelles has an RMSD of 0.68 Å with the structure in nanodisc and digitonin. However, 2.92 Å when compared to the Ost4 structure determined in mixed organic solvents system . The observed kink in the Ost4 structure in mixed organic solvents can possibly a structural artifact most likely attributable to the thickness of the non-polar to polar interface within the mixed solvent system. These observations are consistent with a previous report on the Step 2 protein receptor fragment that had different behavior and helical tendencies in DPC micelles and organic-aqueous solvents (*45*). They likewise concluded that the DPC micelles provide an environment that appears closer to that of a membrane bilayer than organic-aqueous solvents for observing the conformational preferences of small membrane proteins (*45*). The hydrophobic thickness of a micelle composed of 65 molecules of DPC in an MD simulation was 30 ± 4 Å (*46*) which fits well with an observed hydrophobic length of ~25 Å for Ost4 and the average hydrophobic width is usually ~ 25-35 Å (*47*). It was previously reported that lipid rearrangement can occur to completely cover the hydrophobic region of OmpX, a small membrane protein in a similar micelle environment (*48*). The similarity of DPC micelles to the actual membrane thickness and the flexibility of DPC micelles to adjust with the width of the hydrophobic TM stands in contrast to a mixed solvent system that appears to have a narrower width and less flexibility in accommodating the transmembrane helix from Ost4. Therefore, the structure

of Ost4 in DPC micelles is likely closer to the structure in the native environment. This work also suggests that DPC provides a quite reasonable membrane-like environment.

Molecular dynamics simulations of Ost4 and Ost4V23D in DPC micelles was carried out to understand the effect of V23D mutation on the structure and function of Ost4. Molecular dynamics simulations have been frequently used to investigate the structural dynamics and orientation of membrane proteins in membrane environments (*49*-*52*). Although the structures of both the proteins were similar, the two proteins were in a different chemical environment as suggested by the CSP (Figure 4.4). The observed CSP due to mutation was justified by the molecular dynamics results. Both proteins were modeled in a starting configuration at the center of the DPC micelle. For the WT Ost4, the positioning was shown to be very stable in the center of the micelle (Figure 4.30 a and b). The simulation result showed that residues Leu^{10}-Val^{28} remained embedded in the micelles indicating these residues to be the part of the protein that transverses the membrane. Therefore, based on the established orientation of Ost4, the N-terminal residues Met^{1}-Ser^{9} of the Ost4 subunit remain in ER lumen, while the C-terminal residues Asp^{29}-Asn^{36} are exposed to the cytosol with residues Leu^{10}-Val^{28} forming the transmembrane domain. This result is in perfect agreement with the hydrophobicity profile and a previous report (*38*). The molecular dynamics simulation of Ost4V23D in DPC micelles also began with the protein in the center of the micelle, but over time, the simulation indicated that the position of the protein shifted towards the interface of DPC micelles and the solvent before stabilizing (Figure 31 a and b). This result suggests that the V23D mutation greatly destabilizes the protein with the negative charge from the aspartic acid seeking access to an aqueous environment. The

Ost4V23D mutation could lead to the disruption of important hydrophobic interactions resulting in the destabilization of the Stt3-Ost4-Ost3 sub-complex and impacting the function of the entire enzyme complex. The level of interaction of Ost4 and Ost4V23D proteins with the DPC micelles was determined by measuring the hydrophilic SASA and hydrophobic SASA of the transmembrane residues. The hydrophobic and hydrophilic SASA of Ost4 residues was smaller than that of Ost4V23D indicating that the Ost4 V23D transmembrane acts like an amphipathic helix.

MD simulations of both WT and Ost4V23D, with the neighboring membrane-spanning portion of the OST complex, were carried out to understand the impact of this amphipathic helix in the OST complex. The MD results indicated that the amphipathic helix did not dissociate from the Stt3 membrane-bound complex during the simulation time of the Ost4V23D mutant. However, disruption of the local hydrophobic contacts was observed in the WT complex for the mutant complex (Figure 4.35a and b). Additionally, the D23 residue of Ost4V23D protein formed a stable salt bridge with K448 of Stt3 TM13 as depicted in Figure 4.35b. This salt bridge formation is possible only when the D23 residue dislodges from the hydrophobic pocket. As a consequence of D23 dislodging, the observed contacts in the WT complex made to F425 of TM12 as well as L459 and I456 of TM13 from Stt3 and two specific lipid molecules from the membrane as shown in Figure 4.35a are broken resulting in destabilization of the complex. As evident from SASA of residue V23/D23, the hydrophilic aspartate residue becomes more solvent accessible.

The recent structures of the yeast and human OST complex demonstrate that the Ost4 subunit is stacked inside the transmembrane helices of the catalytic subunit, Stt3 (*39*). The

MD simulations of the membrane-bound complex of the WT and V23D mutated proteins demonstrated that the V23D mutation in Ost4 disrupts the hydrophobic interactions, and forms a new salt bridge increasing the solvent exposure of Ost4V23D (Figure 4.35) that results in the disruption of the *N*-linked glycosylation process.

4.5 Conclusion

The structure, and mechanism of function for individual subunits in the OST complex can be fully understood with the help of information gained from several approaches. In this study, NMR showed that the V23D mutation has minimal impact on the structure. However, MD simulation of both the proteins either isolated or in complex with other nearby subunits provides essential information on the role of the V23 residue in the stabilization of the catalytic subcomplex. This study could open a door for future research to understand the function of each subunit of the OST enzyme in the N-glycosylation pathway.

4.6 References

1. M. P. Williamson, T. F. Havel, K. Wuthrich, Solution conformation of proteinase inhibitor IIA from bull seminal plasma by 1H nuclear magnetic resonance and distance geometry. *J Mol Biol* **182**, 295-315 (1985).
2. D. Sakakibara *et al.*, Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature* **458**, 102-105 (2009).

3. K. Inomata *et al.*, High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. *Nature* **458**, 106-109 (2009).

4. F.-X. Theillet *et al.*, Structural disorder of monomeric α-synuclein persists in mammalian cells. *Nature* **530**, 45-50 (2016).

5. S. Ogino *et al.*, Observation of NMR Signals from Proteins Introduced into Living Mammalian Cells by Reversible Membrane Permeabilization Using a Pore-Forming Toxin, Streptolysin O. *Journal of the American Chemical Society* **131**, 10834-10835 (2009).

6. A. G. Palmer, 3rd, NMR characterization of the dynamics of biomacromolecules. *Chem Rev* **104**, 3623-3640 (2004).

7. B. Chaudhary, S. Mazumder, S. Mohanty, Production and biophysical characterization of a mini-membrane protein, Ost4V23D: A functionally important mutant of yeast oligosaccharyltransferase subunit Ost4p. *Protein Expr Purif* **139**, 43-48 (2017).

8. A. Kumar, P. Ward, U. V. Katre, S. Mohanty, A novel and simple method of production and biophysical characterization of a mini-membrane protein, Ost4p: a subunit of yeast oligosaccharyl transferase. *Biopolymers* **97**, 499-507 (2012).

9. T. J. Norwood, Boyd,J.,Heritage,J.E.,Soffe,N.,andCampbell, I. D., Comparison of techniques for 1H-detected heteronuclear 1H-15N Spectroscopy. *J. Magn. Reson.* **87**, 488-501 (1990).

10. O. Zhang, L. E. Kay, J. P. Olivier, J. D. Forman-Kay, Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and

unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *Journal of Biomolecular NMR* **4**, 845-858 (1994).

11. F. Delaglio *et al.*, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995).
12. W. Lee, M. Tonelli, J. L. Markley, NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325-1327 (2015).
13. B. A. Johnson, R. A. Blevins, NMR View: A computer program for the visualization and analysis of NMR data. *Journal of Biomolecular NMR* **4**, 603-614 (1994).
14. B. P. Chaudhary, D. Zoetewey, S. Mohanty, 1H, 13C, 15N resonance assignments and secondary structure of yeast oligosaccharyltransferase subunit Ost4 and its functionally important mutant Ost4V23D. *Biomolecular NMR Assignments*, (2020).
15. P. Guntert, L. Buchner, Combined automated NOE assignment and structure calculation with CYANA. *J Biomol NMR* **62**, 453-471 (2015).
16. Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* **44**, 213-223 (2009).
17. A. T. Brunger *et al.*, Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**, 905-921 (1998).

18. R. A. Laskowski, J. A. Rullmannn, M. W. MacArthur, R. Kaptein, J. M. Thornton, AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**, 477-486 (1996).

19. E. F. Pettersen *et al.*, UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612 (2004).

20. W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38 (1996).

21. L. Schrodinger, The Pymol Molecular Graphics System, Version~1.8. (2015).

22. R. Koradi, M. Billeter, K. Wüthrich, MOLMOL: A program for display and analysis of macromolecular structures. *Journal of Molecular Graphics* **14**, 51-55 (1996).

23. B. P. Chaudhary, D. L. Zoetewey, M. J. McCullagh, S. Mohanty, NMR and MD Simulations Reveal the Impact of the V23D Mutation on the Function of Yeast Oligosaccharyltransferase Subunit Ost4. *Glycobiology*, (2021).

24. P. Dosset, J.-C. Hus, M. Blackledge, D. Marion, Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *Journal of Biomolecular NMR* **16**, 23-28 (2000).

25. H. J. C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **91**, 43-56 (1995).

26. S. Jo, T. Kim, V. G. Iyer, W. Im, CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry* **29**, 1859-1865 (2008).

27. J. Lee *et al.*, CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of Chemical Theory and Computation* **12**, 405-413 (2016).

28. B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **18**, 1463-1472 (1997).

29. W. G. Hoover, Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A Gen Phys* **31**, 1695-1697 (1985).

30. S. Nosé, M. L. Klein, Constant pressure molecular dynamics for molecular systems. *Molecular Physics* **50**, 1055-1076 (1983).

31. M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **52**, 7182-7190 (1981).

32. U. Essmann *et al.*, A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **103**, 8577-8593 (1995).

33. J. Huang, A. D. MacKerell, Jr., CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem* **34**, 2135-2145 (2013).

34. R. Wild *et al.*, Structure of the yeast oligosaccharyltransferase complex gives insight into eukaryotic N-glycosylation. *Science* **359**, 545-550 (2018).

35. D. Case *et al.*, AMBER 2018; 2018. *University of California, San Francisco*.

36. B. P. Chaudhary, D. Zoetewey, S. Mohanty, (1)H, (13)C, (15)N resonance assignments and secondary structure of yeast oligosaccharyltransferase subunit

Ost4 and its functionally important mutant Ost4V23D. *Biomol NMR Assign*, (2020).

37. S. Gayen, C. Kang, Solution structure of a human minimembrane protein Ost4, a subunit of the oligosaccharyltransferase complex. *Biochem Biophys Res Commun* **409**, 572-576 (2011).

38. S. Zubkov, W. J. Lennarz, S. Mohanty, Structural basis for the function of a minimembrane protein subunit of yeast oligosaccharyltransferase. *Proc Natl Acad Sci U S A* **101**, 3821-3826 (2004).

39. L. Bai, T. Wang, G. Zhao, A. Kovach, H. Li, The atomic structure of a eukaryotic oligosaccharyltransferase complex. *Nature* **555**, 328-333 (2018).

40. A. S. Ramirez, J. Kowal, K. P. Locher, Cryo-electron microscopy structures of human oligosaccharyltransferase complexes OST-A and OST-B. *Science* **366**, 1372-1375 (2019).

41. J. Aramini. (http://www.nmr2.buffalo.edu/nesg.wiki/NMR_determined_Rotational_correlation_time, 2010).

42. B. Liang, L. K. Tamm, Solution NMR of SNAREs, complexin and alpha-synuclein in association with membrane-mimetics. *Prog Nucl Magn Reson Spectrosc* **105**, 41-53 (2018).

43. C. Kang, Q. Li, Solution NMR study of integral membrane proteins. *Current opinion in chemical biology* **15**, 560-569 (2011).

44. M. Renault, O. Saurel, P. Demange, V. Reat, A. Milon, in *Membrane Protein Structure Determination: Methods and Protocols,* J.-J. Lacapère, Ed. (Humana Press, Totowa, NJ, 2010), pp. 321-339.

45. A. Neumoin, B. Arshava, J. Becker, O. Zerbe, F. Naider, NMR studies in dodecylphosphocholine of a fragment containing the seventh transmembrane helix of a G-protein-coupled receptor from Saccharomyces cerevisiae. *Biophys J* **93**, 467-482 (2007).

46. D. P. Tieleman, D. van der Spoel, H. J. C. Berendsen, Molecular Dynamics Simulations of Dodecylphosphocholine Micelles at Three Different Aggregate Sizes: Micellar Structure and Chain Relaxation. *The Journal of Physical Chemistry B* **104**, 6380-6388 (2000).

47. A. Mashaghi *et al.*, Hydration strongly affects the molecular and electronic structure of membrane phospholipids. *The Journal of Chemical Physics* **136**, 114709 (2012).

48. C. Fernandez, C. Hilty, G. Wider, K. Wuthrich, Lipid-protein interactions in DHPC micelles containing the integral membrane protein OmpX investigated by NMR spectroscopy. *Proc Natl Acad Sci U S A* **99**, 13533-13537 (2002).

49. T. Hansson, C. Oostenbrink, W. van Gunsteren, Molecular dynamics simulations. *Current Opinion in Structural Biology* **12**, 190-196 (2002).

50. Y. Wang, E. Tajkhorshid, Electrostatic funneling of substrate in mitochondrial inner membrane carriers. *Proc Natl Acad Sci U S A* **105**, 9598-9603 (2008).

51. G. Hedger *et al.*, Lipid-Loving ANTs: Molecular Simulations of Cardiolipin Interactions and the Organization of the Adenine Nucleotide Translocase in Model Mitochondrial Membranes. *Biochemistry* **55**, 6238-6249 (2016).

52. V. Kurauskas *et al.*, How Detergent Impacts Membrane Proteins: Atomic-Level Views of Mitochondrial Carriers in Dodecylphosphocholine. *The Journal of Physical Chemistry Letters* **9**, 933-938 (2018).

CHAPTER 5

SOLID-STATE NMR (^{13}C, ^{15}N) RESONANCE ASSIGNMENT AND ATOMIC RESOLUTION STRUCTURE DETERMINATION OF Ost4 AND Ost4V23D IN LIPID BILAYER

5.1 Introduction

Membrane proteins are involved in various essential life processes such as signal and energy transduction, cell signaling, ion and nutrient transport, and biochemical reaction catalysis (*1*). As a consequence, nearly 25% of the genes encode for membrane proteins. Since membrane proteins are associated with a wide variety of diseases such as cancer, cystic fibrosis disorders, developmental delay, liver dysfunction, hypoglycemia, neurological disorders etc., 70 % of the membrane proteins are drug targets and 50% of them are new drug targets (*2, 3*). Membrane proteins perform their functions in the cell membrane environment and it's best to study their structure and function in their native environment. However, the complex nature of the membrane and its interactions with the other elements of the membrane create difficulty in interpretation of the results. Synthetic lipid bilayers provide the membrane-like environment for elucidating structural and functional aspects of the membrane proteins. Therefore, the reconstitution of membrane proteins into the native membrane-like system has become an essential step to study the structure and function of membrane proteins in a native membrane-like environment.

Study of the membrane protein in a lipid bilayer ensures that the protein sample represents a biologically relevant protein conformation. The 3D structure determinations of Ost4 and Ost4V23D in DPC micelles by using solution NMR have been discussed in Chapter 4. However, detergent micelles cannot provide the same near-native lipid bilayer environment as can be achieved by using phospholipid. The 3D structure of these membrane proteins can be determined in a near-native lipid bilayer environment by using solid-state NMR spectroscopy. The comparison of the structure of the same molecule determined in micelles and lipid bilayers has shown that the structure is quite different in the two different environments (*4*). Therefore, it is important to determine the structure of the membrane protein in a native-like lipid bilayer environment to know the actual structure and behavior of the protein in the real membrane.

The solution NMR requires the sample to be in the solution phase for data acquisition. The structure determination by solution NMR is based on isotropic interactions, such as isotropic chemical shift, J couplings, and relaxation information (T1, T2, and NOEs). Solution NMR provides much better resolution with line widths a hundred times narrower and much stronger signal intensity because of the rapid molecular tumbling (*5, 6*). Solution NMR has become the most frequently used technique for structure determination of biomolecules. However, its application is limited by the fact that the sample should be in a solution state. The solid-state NMR can be used to determine the structure of solid protein samples reconstituted in an appropriate membrane mimetic system.

To determine the structure of a membrane protein, its production and purification is the key step followed by reconstitution in an appropriate membrane mimetic system. Membrane protein production and purification to homogeneity have always been challenging. However, once the desired membrane protein is successfully expressed and purified to homogeneity, the next crucial step is the reconstitution of the protein into a membrane-like environment. The key goal of reconstitution is to incorporate the protein of interest into the lipid bilayers, a membrane-like environment. The synthetic lipid bilayers such as: POPC/POPE (1-palmitoyl-2-oleoyl-snglycero-3-phosphocholine/ 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine), DMPC/DMPG (1,2-dimyristoyl-sn-glycero-3-phosphocholine/1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol), POPC/POPG (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine/ 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)), and DOPC/DOPE (1,2-dioleoyl-sn-glycero-3-phosphocholine/ 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine) can be used for the reconstitution of membrane proteins (*7-10*). The mixture of either lipid provides a net negative charge on the lipid bilayer. The lipid bilayer contains both cylindrical head group provided by phosphocholine (PC) head groups and conical lipid head groups provided by phosphoethanolamine (PE) or phosphoglycerol (PG) head groups. A variety of molar ratio for the mixture of lipids can be used to obtain the required thickness for the lipid bilayer.

Another important parameter for a good reconstitution is the protein to lipid molar ratio. The appropriate protein to lipid molar ratio is dependent on molecular weight and the type of solid-state NMR experiments. Generally, the protein to lipid molar ratio should be low enough (1:80 to 1:100) to obtain a high degree of sample alignment, however, for smaller

protein and MAS solid-state NMR, the protein to lipid molar ratio of about 1:30 or even lower would be better (*7*). Once an appropriate protein to lipid ratio is determined for reconstitution, preparation of a large unilamellar vesicle (LUV) followed by the removal of detergent molecules used for the solubilization of the protein is an essential step. It is very important to remove the excess detergent while transferring the membrane protein into LUV, otherwise, the excess detergent can alter the structure of membrane protein (*11, 12*). In the detergent-based reconstitution process, the protein solubilized in detergent is transferred to the LUV dissolved in an appropriate buffer followed by the removal of excess detergent by using hydrophobic beads. The protein reconstituted in lipid vesicles is obtained either by ultra-centrifugation or by filtration using a molecular cutoff filter unit.

Although X-ray crystallography, Cryo-EM, and solution NMR techniques are frequently used for the determination of protein structures, solid-state NMR (ssNMR) is the most suited application used to study membrane proteins in a near native-like membrane environment, the lipid bilayers.

In this chapter, the reconstitution of Ost4 and Ost4V23D in a POPC/POPE lipid bilayer, 3D structure determination by using solid-state NMR, and molecular dynamics simulation results of Ost4 and Ost4V23D will be discussed. Additionally, the topology and behavior of the proteins are assessed by molecular dynamics simulation in membrane environments. Moreover, the comparison of structures of both the proteins in a variety of membrane mimicking systems is made to conclude the best-suited system to study the structural impacts of Ost4 and its critical mutant Ost4V23D protein.

5.2 Materials and methods

5.2.1 Protein expression and purification of Ost4 and Ost4V23D

Expression and purification of the proteins were carried as described previously in Chapter 2 and Chapter 3.

5.2.2 Reconstitution procedure

Both the proteins (Ost4 and Ost4V23D) were reconstituted in the synthetic lipid bilayer of a mixture of POPC and POPE in a 3:2 molar ratio. The initial four rounds of reconstitution out of seven reconstitutions were carried out in Dr. Donghua Zhou's lab in the physics department with the help of Dr. Hem Moktan. The first step of the reconstitution is the preparation of liposome or large unilamellar vesicles (LUV). LUV is prepared by the extrusion process as described below.

5.2.2.1 Preparation of LUV by the extrusion process

POPC and POPE powder (Avanti Polar Lipids, Inc.) was weight in the ratio of 3:2 in a 50 ml clean and dry round-bottomed flask and dissolved in 2 mL of chloroform. The bulk of chloroform was removed by evaporating the mixture under a flow of N_2 gas with gentle rotation. The film was incubated under a flow of nitrogen gas for 3 additional hours. The formed thin film of lipid mixture was then left overnight under vacuum to get rid of any traces of chloroform. The lipid film was then dissolved in 2 mL of NMR buffer (50 mM sodium phosphate, pH 6.5, containing 1mM EDTA and 0.01% NaN_3) by flushing with the

help of a pipette. The mixture of lipids in NMR buffer was then subjected to 10 freeze /thaw cycles by alternatively placing the sample e-tube in a -80 ^{0}C freezer and 40 ^{0}C warm water bath.

The freeze-thawed sample was loaded into one of the gas-tight syringes and was placed at one end of the Mini-Extruder (Avanti Polar Lipids, Inc.). Prior to the extrusion process, all the parts of the Mini-Extruder were soaked in water and the syringes were pre-wet by NMR buffer to reduce any dead volume. The empty gas-tight syringe was placed into the other end of the Mini-Extruder. A preassembled extruder apparatus containing a 0.1 μm membrane was inserted into the extruder stand. The plunger of the syringe containing lipid solution was pushed gently until all the solution was transferred to the syringe at the other end of the extruder. The alternate syringe was pushed back gently and the whole lipid solution was transferred to the initial syringe. This process was repeated 11 times. The prepared LUV was stored at 4 ^{0}C for later use.

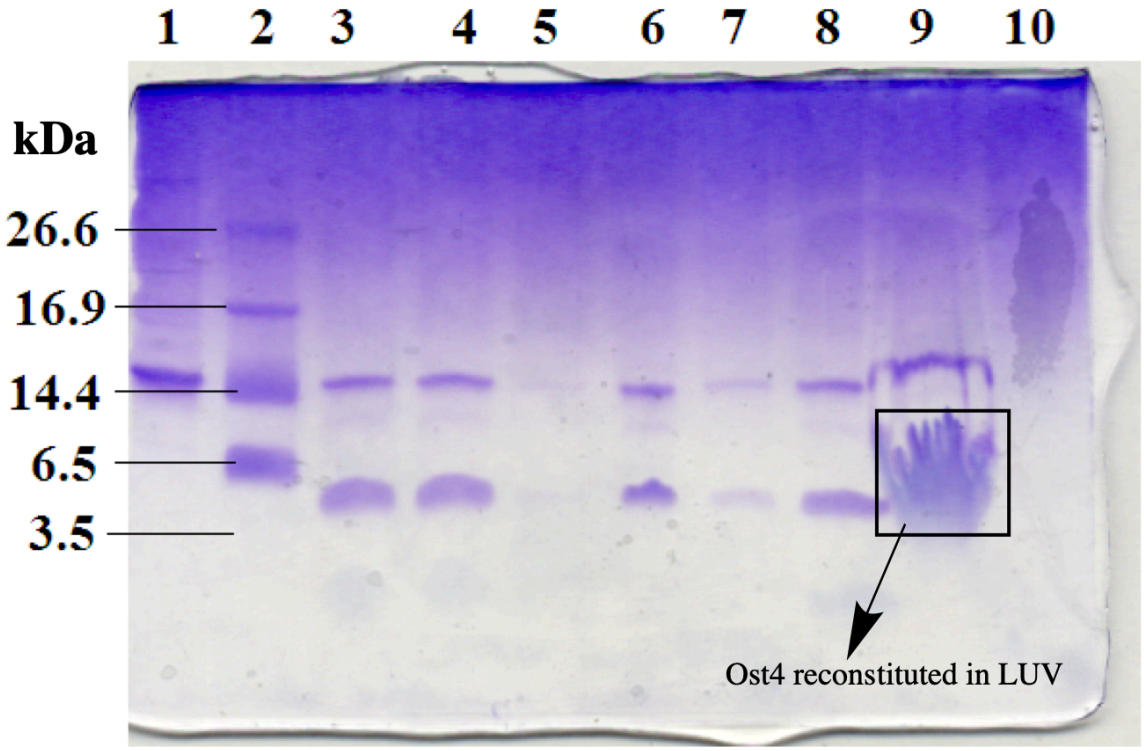
The purified dry proteins were solubilized in 15 mM CHAPS (3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate) (Avanti Polar Lipids) detergent maintaining CHAPS: Lipids mass ratio to 1:1.2. The CHAPS solubilized proteins were then reconstituted into an LUV of POPC-POPE by detergent removal using Bio-beads SM-2 (Bio-Rad).

5.2.3 Optimization of protein reconstitution

Initially, 1.8 mg of unlabeled Ost4 protein was reconstituted in 12.5 mg of POPC+POPE (3:2 w/w ratio) containing 7.5 mg of POPC and 5 mg of POPE. This yielded a protein to lipid molar ratio of 1:30. The lipid to CHAPS (w/w) ratio was maintained to 1.2:1 by adding

a total of 925 μL of 15 mM CHAPS (MW 614.88 g/mol) solution in NMR buffer. A 1 mL of LUV was prepared by using 12.5 mg of POPC+POPE by following the procedure described above. Ost4 pellet (1.8 mg)was dissolved 330 μL of CHAPS solution. 595 μL of 15 mM CHAPS was mixed with 1 mL LUV and incubated for 45 minutes at room temperature. The protein solution in CHAPS and LUV CHAPS solutions were mixed together followed by 2 h incubation at 4 ºC. 50 mg of prewet hydrophobic beads were added and incubated for 2 h at 4 ºC. In the next step, 50 mg of hydrophobic beads were added and incubated for 2 h at 4 ºC. Now, 125 mg of the hydrophobic beads were added and incubated for another 2 h at 4 ºC. The solution was drawn out into a separate e-tube using a syringe to separate the hydrophobic beads. The solution was centrifuged at 70000 rpm for 2 h at 4 ºC. The pellet containing Ost4 reconstituted in the LUV was separated from the supernatant liquid and dried by using a speedvac. During the reconstitution process, the supernatant liquid was observed to contain an additional pellet even after final centrifugation. The observed pellet was assumed to be LUV containing reconstituted Ost4 protein. Therefore, any leftover pellet in the supernatant liquid was recovered by filtering through a 10K molecular cutoff filter for 20 minutes at 16000 rcf. The success of reconstitution was confirmed by Tris-Tricine gel (Figure 5.1). Since the supernatant liquid after centrifugation (Figure 5.1, lane 8) showed a significant amount of residual Ost4 protein and was recovered after filtration (Figure 5.1, lane 9), the filtration procedure was performed for further reconstitution. The ultracentrifugation (centrifugation at 70K rpm) step was not performed to avoid the loss of reconstituted sample. Instead, the solution was filtered through a 3K molecular cut-off filter for the reconstitution of ^{13}C, ^{15}N labeled Ost4

and Ost4V23D in LUV. And, the protein to lipid molar ratio was 1:16 for the reconstitution of double labeled proteins.



1. Cell pellet
2. Protein Marker
3. Unlabeled Ost4 dissolved in CHAPS
4. Solution after adding 50 mg hydrophobic beads
5. Hydrphobic beads from second addition and incubation
6. Solution after adding 125 mg hydrophobic beads
7. Hydropbobic beads at the end of reconstitution
8. Supernatent liquid after centrifugation
9. Speed Vaccum dried pellet of Ost4 reconstituted in LUV
10. Filtrate after filtration of Supernatent liquid.

Figure 5.1: SDS-PAGE gel picture of reconstitution of unlabeled Ost4 in LUV. The information about lanes is labeled.

5.2.4 Reconstitution of ^{13}C, ^{15}N – labeled Ost4 and Ost4V23D in LUV

A total of 12 mg and 9.2 mg of Ost4 and Ost4V23D proteins were reconstituted in 42 mg (25.2 mg POPC + 16.8 mg POPE) and 33.2 mg (20.4 mg POPC + 12.8 mg POPE) of LUV,

respectively. For the reconstitution of Ost4 protein, LUV was prepared in two round-bottomed (RB) flasks. To each RB flask containing 13 mg POC and 8 mg POPE, 2 mL of chloroform was added. The RB flask was rotated under a gentle flow of N_2 gas until all the chloroform evaporated and a thin film of lipid was formed on the wall of the RB flasks. The thin film was incubated for another 3 h under the flow of N_2 gas. The RB flask containing formed thin film of lipid mixture was connected to the vacuum line and sealed airtight with a piece of parafilm. The film of the lipid mixture was then left overnight under vacuum to get rid of any traces of chloroform. The lipid film was then dissolved in 2 mL NMR buffer (50 mM sodium phosphate, pH 6.5, containing 1 mM EDTA and 0.01% NaN_3) by pipetting the solution up and down. The lipid mixture was incubated for 1 h at room temperature with shaking. The mixture of lipids in NMR buffer was then subjected for 10 freeze /thaw cycles by alternatively placing the sample e-tube in a -80 ºC freezer and 40 ºC warm water bath. LUV was prepared by an extrusion process as described earlier in this chapter. The extrusion procedure was performed four times which yielded a total of 3.5 mL of LUV. Figure 5.2 demonstrates the steps followed for the LUV preparation and reconstitution of the protein into LUV.

A total of 3.8 mL of 15 mM CHAPS was needed to maintain a lipid to CHAPS ratio of 1.2:1 for Ost4 reconstitution. 12 mg of dry pellet of ^{13}C, ^{15}N labeled Ost4 protein was dissolved in 2.3 mL of 15 mM CHAPS by vortexing followed by bath sonication. The reconstitution was carried out in 5 e-tubes. To each e-tube containing 700 μL LUV, 300 μL 15 mM CHAPS was added and incubated for 45 minutes at RT. And 460 μL of CHAPS containing dissolved Ost4 was added to the e-tube containing 700 μL of LUV and 300 μL

15 mM CHAPS followed by 2 h incubation with shaking at RT. 100 mg of wet hydrophobic beads were added and incubated for 2 h at RT with shaking. Another 100 mg of

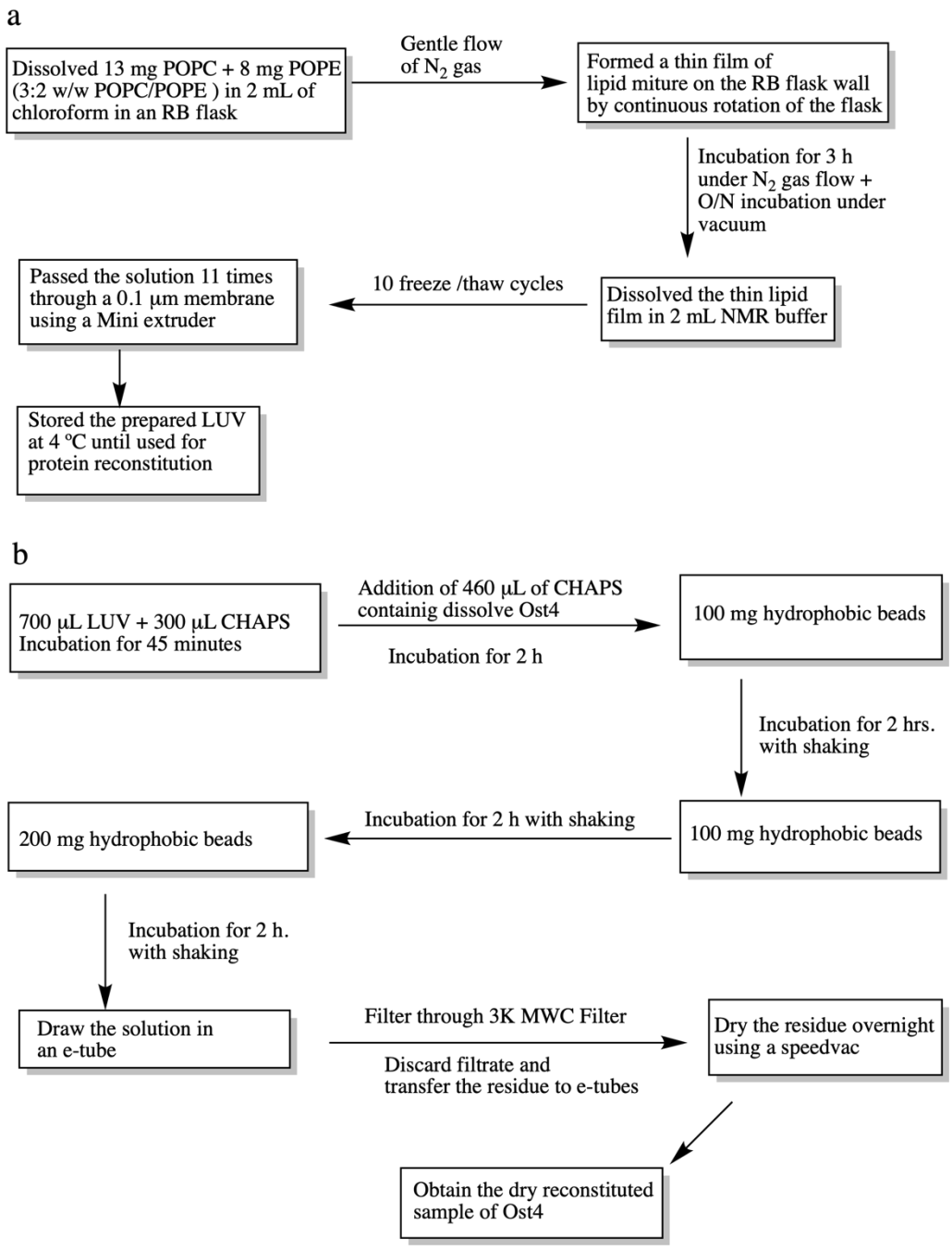


Figure 5.2: The flowchart demonstrating the steps of LUV preparation by extrusion process; (a) The LUV was prepared in two RB flasks. The steps show the amounts used

for one of the RB flasks. Reconstitution of Ost4 protein in the LUV; (b) The reconstitution was performed in five e-tubes. The figure shows the amounts used in one of the e-tubes.

hydrophobic beads were added and incubated for 2 h with shaking. Finally, 200 mg of hydrophobic beads was added to each e-tubes followed by 2 h incubation. The solution was separated from hydrophobic beads by using a syringe. The solution was filtered by using a 3K molecular weight cut off filter at 3000 rpm. The filtrate was discarded and the residue on the filter was transferred into a pre-weighed e-tube. The solid residue was dried overnight by using a SpeedVac. By following a similar procedure for Ost4 protein, 9.2 mg Ost4V23D protein was reconstituted in 33.2 mg of POPC-POPE lipid bilayer. A total of 27.8 mg and 22.7 mg of dry samples of ^{13}C, ^{15}N double labeled Ost4 and Ost4V23D, respectively, were obtained for NMR data collection.

5.2.5 Packing of samples of Ost4 and Ost4V23D in the rotor

A total of 13 mg and 16.9 mg dry samples of Ost4 and Ost4V23D, respectively, were added packed in two 3.2 mm sapphire rotors. 4 μL (~4 mg) and 4 μL (~5.1 mg) of water (30% w/w) was added to the rotor containing the Ost4 and Ost4V23D samples, respectively. The rotors were then spun at 1000 rpm for two minutes to complete the hydration of the samples in the rotors. The rotors were capped with a vessel cap. The rotor packing for ssNMR data collection was performed by Dr. Jochem Struppe, a senior application scientist at Bruker BioSpin Corporation.

5.2.6 Solid-state NMR data collection

All solid-state NMR (ssNMR) experiments were acquired using a 3.2 mm sapphire rotor and triple resonance (^{1}H, ^{13}C, ^{15}N) probe head, either at a static magnetic field of 14.1 T corresponding to 600 MHz proton resonance frequency or at a static magnetic field of 16.4 T corresponding to 700 MHz proton resonance frequency at Bruker Biospin Corporation, Billerica, MA, USA. All the ssNMR data were acquired by Dr Jochem Struppe, a senior application scientist at Bruker Biospin Corporation, Billerica, MA, USA. The backbone resonances were assigned using a combination of 2D and 3D correlation experiments. 2D [^{13}C-^{13}C] CC2D (*13*) experiments for both the proteins were performed with 50 ms DARR mixing time. The raw data size for Ost4 2D [^{13}C-^{13}C] experiment and Ost4V23D 2D [^{13}C-^{13}C] were 2048 x 1100 and 4096 x 1124, respectively. The 3D[^{15}N,^{13}CA-CX] (*14*), 3D [^{15}N,^{13}CO-CX] (*14*), and 3D [^{13}CA,^{15}N,^{13}CO-CX] (*15*) for both the proteins were recorded with ^{13}Ca-^{13}CX , ^{13}CO-^{13}CX, and ^{13}CO-^{13}CX DARR mixing time of 50 ms, respectively. The 3D[^{15}N,^{13}CA-CX] for Ost4 and Ost4V23D were recorded with a raw data size of 2048 x 160 x 50 and 2048 x 88 x 42, respectively, in the ^{13}C, ^{13}C, and ^{15}N dimensions. Similarly, the raw data size for the Ost4 and Ost4V23D 3D [^{15}N,^{13}CO-CX] experiment were 2048 x 58 x 56 and 2048 x 58 x 56, respectively, in ^{13}C, ^{15}N, and ^{13}C dimensions. The raw data sizes for the 3D [^{13}CA,^{15}N,^{13}CO-CX] experiments for Ost4 and Ost4V23D were 2048 x 50 x 58 and 2048 x 32 x 128, respectively, in the ^{13}C, ^{15}N, and ^{13}C dimensions. For ^{1}H-^{1}H distance measurement, 2D CHHC (*16*) for each protein were recorded with a ^{1}H-^{1}H mixing time of 300 ms and ^{13}C-^{1}H and ^{1}H-^{13}C cross polarization (CP) contact time of 150 ms. Similarly, 2D NHHC (*16*) experiments for each protein were collected with a ^{1}H-^{1}H mixing time of 300 ms. The raw data size for the 2D CHHC experiments for Ost4 and Ost4V23D

were 2048 x 750 and 2048 x 175, respectively. Similarly, the raw data size for the 2D NHHC experiments for Ost4 and Ost4V23D were 2048 x128 and 2048 x 58, respectively. Tables 5.1 and 5.2 demonstrate the list of experiments for Ost4 and Ost4V23D, respectively, with the important parameters used during the ssNMR data acquisition. All the experiments were performed at 12.5 kHz MAS at either 10 ^{0}C or -43 ^{0}C with ^{1}H decoupling power of -22.8 dB during acquisition. All the spectra were processed by using NmrPipe (*17*) and analyzed by using CcpNmr V2.1 (*18*).

Table 5.1: List of experiments for Ost4 with the important parameters used during the ssNMR data acquisition

Experiments	Dimension	Isotope	Carrier Frequency (ppm)	Spectral width (ppm)	Number of data points	Number of scans	Mixing time
CC2D	1	^{13}C	100.273	298.877	2048	48	^{13}C-^{13}C 50 ms
	2	^{13}C	100.235	213.483	1100		
NCACX	1	^{13}C	101.935	298.877	2048	128	^{13}Ca-^{13}CX 50 ms
	2	^{13}C	56.981	70.983	160		
	3	^{15}N	121.103	40.010	50		
NCOCX	1	^{13}C	102.015	298.877	2048	128	^{13}CO-^{13}CX 50 ms
	2	^{13}C	177.000	23.661	58		
	3	^{15}N	120.008	40.010	56		
CaNCOCX	1	^{13}C	100.000	298.877	2048	256	^{13}CO-^{13}CX 50 ms
	2	^{15}N	120.000	40.010	50		
	3	^{13}C	100.000	23.661	58		
CHHC	1	^{13}C	56.670	298.890	2048	96	^{1}H-^{1}H 300 μs
	2	^{13}C	56.717	141.973	750		

NHHC	1	^{13}C	54.919	298.890	2048	512	^{1}H-^{1}H 300 μs
	2	^{13}C	122.204	58.714	128		

Table 5.2: List of experiments for Ost4V23D with the important parameters used during the ssNMR data acquisition

Experiment	Dimension	Isotope	Carrier Frequency (ppm)	Spectral width (ppm)	Number of data points	Number of scans	Mixing time
CC2D	1	^{13}C	100.400	298.876	2048	16	^{13}C-^{13}C 50 ms
	2	^{13}C	100.400	212.983	1124		
NCACX	1	^{13}C	99.345	298.876	2048	128	^{13}Ca-^{13}CX 50 ms
	2	^{13}C	57.298	35.492	88		
	3	^{15}N	120.003	44.036	42		
NCOCX	1	^{13}C	100.015	298.876	2048	128	^{13}CO-^{13}CX 50 ms
	2	^{13}C	175.015	20.000	58		
	3	^{15}N	120.008	40.000	56		
CaNCOCX	1	^{13}C	100.000	298.877	2048	64	^{13}CO-^{13}CX 50 ms
	2	^{15}N	120.000	44.036	32		
	3	^{13}C	100.000	70.983	128		
CHHC	1	^{13}C	45.000	298.893	2048	128	^{1}H-^{1}H 300 μs
	2	^{13}C	45.000	70.987	175		
NHHC	1	^{13}C	55.195	298.890	2048	512	^{1}H-^{1}H 300 μs
	2	^{13}C	118.635	58.714	58		

5.2.7 ^{13}C, ^{15}N resonances assignment

The processed ssNMR spectra such as 2D [^{13}C-^{13}C] CC2D, 3D[^{15}N,^{13}CA-CX], 3D [^{15}N,^{13}CO-CX], 3D [^{13}CA,^{15}N,^{13}CO-CX] , 2D CHHC, and 2D NHHC of Ost4 and Ost4V23D were opened in the "Analysis" software package integrated with CcpNmr V2.1. The protein contains only one glycine residue present at position 15 of the protein sequence. The glycine residue contains only Cα with a chemical shift around 45 to 47 ppm. Therefore, the Cα Chemical shift of G15 was identified and the corresponding ^{15}N and ^{13}CO chemical shifts were identified in CC2D and 3D[^{15}N,^{13}CA-CX] spectra. The 3D[^{15}N,^{13}CA-CX] provides ^{15}N, and ^{13}C resonances belonging to a single (i) residue and 3D [^{15}N,^{13}CO-CX] links ^{15}N of its own (i) residue to ^{13}CO and other ^{13}C resonances of the previous (i-1) residue in the protein sequence. Therefore, once the ^{15}N ppm of G15 was identified, the ^{13}CO, Cα and Cβ chemical shifts of its previous residue F14 were identified in the [^{15}N,^{13}CO-CX] spectrum. Again, in the 3D[^{15}N,^{13}CA-CX] spectrum the Cα and Cβ chemical shifts of F14 and the corresponding ^{15}N chemical shift were identified. The same ^{13}C resonances were identified and confirmed in the CC2D spectrum as well. To find the ^{15}N, and ^{13}C peaks of the I16 (i+1) residue, the ^{13}CO, and the $^{13}C\alpha$ of G15 were searched in 3D[^{15}N,^{13}CO-CX] at difference ^{15}N ppm values. Once ^{15}N ppm was identified, the ^{13}C resonances of the I16 residue were identified in 3D[^{15}N,^{13}CA-CX] spectrum. In this way, the backbone ^{13}C, ^{15}N resonances of Ost4 and Ost4V23D protein sequence were assigned by using CC2D and connecting ^{15}N ppm of the (i) and (i-1) residues using 3D[^{15}N,^{13}CA-CX] and 3D[^{15}N,^{13}CO-CX] spectra. The lists of chemical shifts of Ost4 and Ost4V23D proteins obtained from the assignment project are provided in Tables 5.3 and 5.4, respectively.

Table 5.3: Chemical shifts of Ost4 protein obtained from ssNMR assignment

Residue number	Residue	N	CO	Cα	Cβ	Cγ	Cδ
1	Met	-	-	-	-	-	-
2	Ile	-	-	-	-	-	-
3	Ser	-	178.04	57.98	66.03	-	-
4	Asp	121.34	178.07	58.05	41	-	-
5	Glu	119.3	177.53	59.5	29.05	33	-
6	Gln	119.58	177.39	58.86	31.62	33	-
7	Leu	122.08	178.05	58.09	41.82	26	-
8	Asn	117.7	177.73	56.22	34.09	-	-
9	Ser	114.27	176.96	60.91	63.35	-	-
10	Leu	122.5	173.86	58.01	41.85	26	21.39
11	Ala	125.6	174.59	55.35	18.66	-	-
12	Ile	120.3	176.72	66.17	38.04	30.7	17.33
13	Thr	119.11	175.48	67.48	68.23	22	-
14	Phe	121.59	176.2	61.9	39.5	-	-
15	Gly	106.31	174.3	47.62	-	-	-
16	Ile	121.55	177.15	66.1	38.02	29.3	17.56
17	Val	122.5	177.9	66.99	31.26	21	-
18	Met	119.37	177.14	57.16	31.32	-	-
19	Met	118.12	176.63	56.27	31.31	-	-
20	Thr	117.19	175.93	67.31	68.77	21	-
21	Leu	119.06	174.4	55.56	37.03	-	22.93
22	Ile	121.87	176.59	65.38	38.22	17	-
23	Val	118.75	175.51	67.56	31.35	21	-
24	Ile	119.69	177.26	65.46	38.18	17	-
25	Tyr	119.69	176.89	56.61	39.69	-	-
26	His	115	175.23	55.91	30.08	-	-
27	Ala	123.12	179.07	55.4	18.49	-	-
28	Val	117.49	175.98	67.9	34.83	21	-
29	Asp	121.27	177.35	56.43	41.64	-	-
30	Ser	113.49	176.6	60.66	62.82	-	-
31	Thr	116.31	176.56	67.93	68.56	21	-
32	Met	119.75	176.57	56.77	32.59	29	-

33	Ser	118.18	177.2	59.32	60.28	-	-
34	Pro	120	177.19	67.11	31.37	22	51.24
35	Lys	120.31	178.14	57.53	35.72	26	28.65
36	Asn	117.81	177.56	56.18	34.1	-	-
37	Arg	127.19	174.17	59.89	34.97	27	42.35
38	Leu	126.88	177.27	58.26	41.53	-	-
39	Glu	125.63	174.27	53.85	32.28	33	-

Table 5.4: Chemical shifts of Ost4V23D protein obtained from ssNMR assignment

Residue number	Residue	N	CO	Cα	Cβ	Cγ	Cδ
1	Met	-	-	-	-	-	-
2	Ile	-	-	-	-	-	-
3	Ser	121.86	174.39	59.96	60.59	-	-
4	Asp	120.61	177.28	56.35	40.42	-	-
5	Glu	119.99	176.89	57.28	28.65	-	-
6	Gln	119.74	177.34	56.71	30.71	-	-
7	Leu	121.6	177.14	56.2	39.71	-	24.15
8	Asn	117.82	175.67	55.36	39.37	-	-
9	Ser	115.66	174.42	60.8	61.13	-	-
10	Leu	121.55	177.04	55.9	39.13	24.76	-
11	Ala	122.6	177.53	53.2	15.97	-	-
12	Ile	118.54	177.3	58.05	39.78	15.29,25.43	12.42
13	Thr	116.66	175.21	66.22	67.12	19.2	-
14	Phe	119.47	174.64	63.67	36.36	-	-
15	Gly	105.69	172.42	45.86	-	-	-
16	Ile	121.52	175.83	65.32	39.69	28.83,1	6.43
17	Val	-	175.98	65.02	28.93	21.90,21.42	
18	Met	118.9	175.64	54.61	28.54	-	-
19	Met	117.34	176.22	54.41	30.38	-	-
20	Thr	117.34	173.2	65.95	68.14	20.16	-
21	Leu	120.46	177.09	55.98	39.72	-	24.24
22	Ile	119.84	175.39	64	40.1	28.89	15.57
23	Asp	122.1	176.12	55.75	40.13	-	-
24	Ile	118.9	176.32	63.72	35.53	29.51,15.54	15.24

25	Tyr	120.46	175.24	59.81	39.05	-	-
26	His	115.28	-	59.12	29.08	-	-
27	Ala	121.97	177.16	52.9	17.45	-	-
28	Val	120.72	175.86	64.65	29.25	21.71	-
29	Asp	119.67	175.97	55.61	39.77	-	-
30	Ser	115.61	175.59	60.44	61.04	-	-
31	Thr	116.66	175.72	65.67	66.11	18.58	-
32	Met	116.33	175.61	54.02	31.65	-	-
33	Ser	115.04	175.89	60.28	61.16	-	-
35	Lys	126.04	172	57.88	33.37	19.68	-
36	Asn	126.66	172.17	57.86	40.74	-	-
37	Arg	126.35	171.7	58.53	32.8	-	-
38	Leu	120.41	176.25	55.94	39.21	-	24.88
39	Glu	127.91	171.7	58.69	33.01	-	-

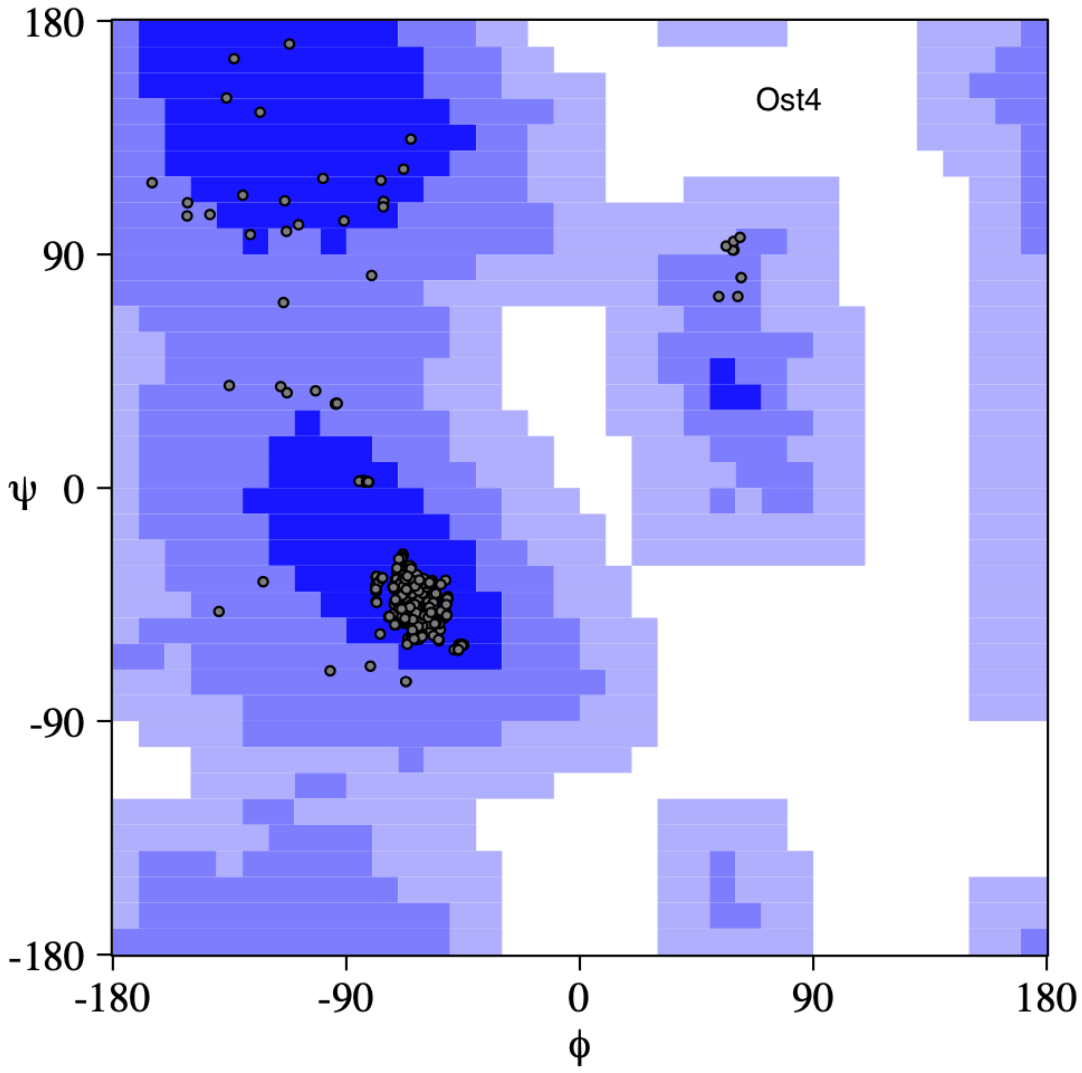
5.2.8 Resonance assignment of 2D CHHC and 2D NHHC spectra

Once the chemical shift values of ^{13}C, ^{15}N backbone and side-chain resonances were assigned, the ^{13}C – ^{13}C cross-peaks in the 2D CHHC and the ^{13}C – ^{15}N cross-peaks in 2D NHHC spectra were assigned either manually or by copying the assignment by using the "Copy Assignment" module integrated into the CcpNmr software. A total of 80 and 32 cross-peaks were assigned in the 2D CHHC and 2D NHHC spectra of Ost4. A similar strategy was used to assign ^{13}C – ^{13}C cross-peaks in the 2D CHHC and ^{13}C – ^{15}N cross-peaks in the 2D NHHC spectra of Ost4V23D protein. A total of 67 and 40 cross-peaks were assigned in the 2D CHHC and 2D NHHC of Ost4V23D.

5.2.9 Structure determination of Ost4 and Ost4V23D proteins in the lipid bilayer

The torsion angles – phi (Φ) and psi (Ψ) – of the residues of the protein define the geometry of attachment of a residue to its two adjacent residues. This torsion angle is also known as

the dihedral angle. The dihedral angle restraints are required as input in the structure calculation software for structure determination of the protein. The dihedral (Φ and Ψ) angle restraints were derived by using Dihedral Angles from Global Likelihood Estimates (DANGLE) (*19*) integrated into the CcpNMR program using ^{15}N, $^{13}C\alpha$, $^{13}C\beta$, and $^{13}C'$ resonances. Only the dihedral angles that existed in a good global likelihood estimates region were used as dihedral restraints for structure calculation. A Ramachandran plot is used to determine the permitted torsion angles to obtain insight into the protein structure. The good global likelihood estimates of dihedral angles of the residues of Ost4 and Ost4V23D proteins were assessed by using the Ramachandran plot (Figure 5.3 and 5.4). The Ramachandran statistics of the Ost4 and Ost4V23D proteins (Figure 5.3 and 5.4) show that all the residues have good global likelihood estimates of dihedral angles. The CcpNmr generated dihedral angle restraints were converted into CYANA readable format by using the Format Converter macro integrated into the CcpNmr software package.



95.0% in most favored regions
4.4% in additionally allowed regions
0.6% in generously allowed regions
0.0% in disallowed regions

Figure 5.3: Ramachandran plot of the Ost4 protein. This plot is used to analyze the good global likelihood estimates of dihedral angles.

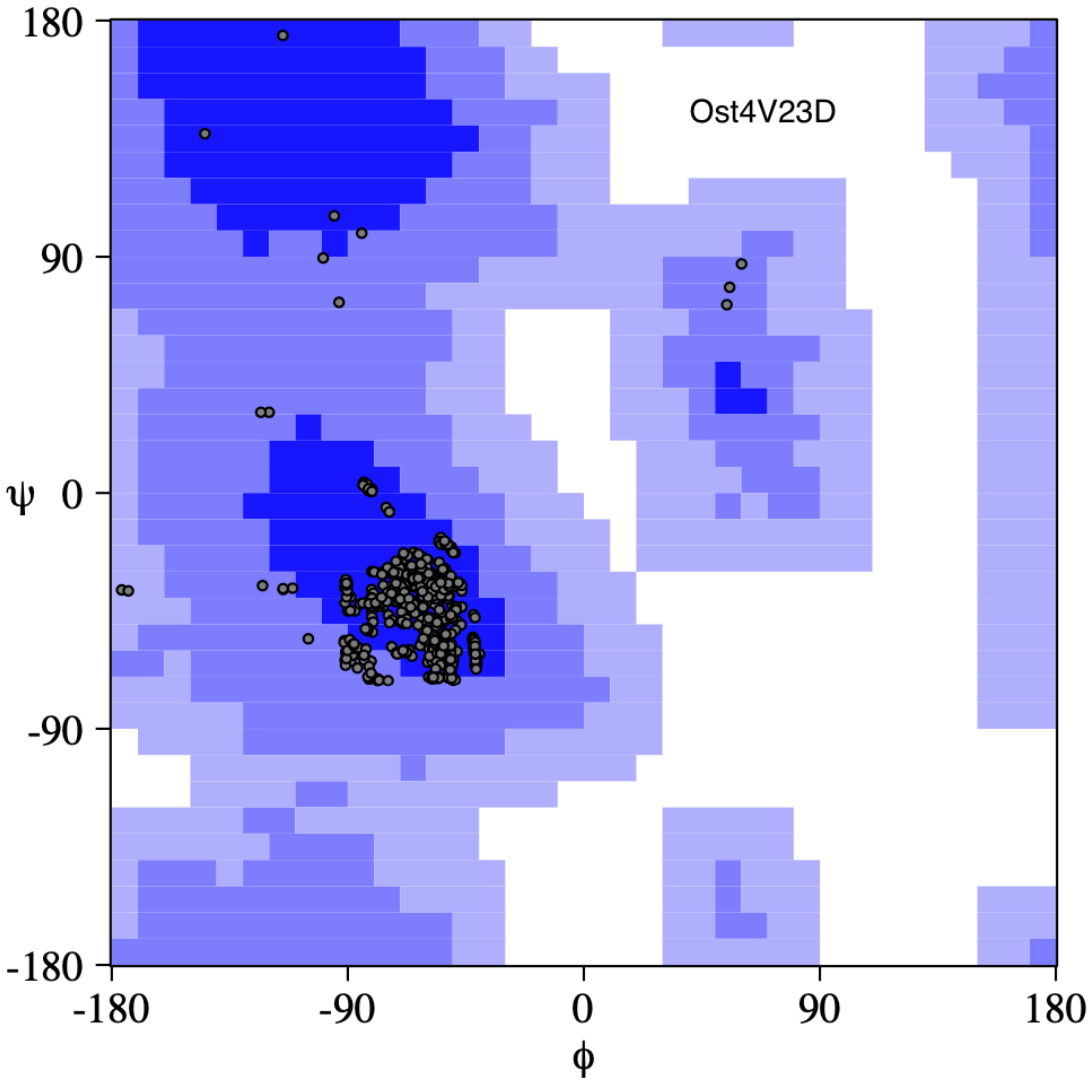


Figure 5.4: Ramachandran plot of the Ost4V23D protein. This plot is used to analyze the good global likelihood estimates of dihedral angles.

Although $^{13}C - ^{13}C$ and $^{15}N - ^{13}C$ cross-peaks in the 2D CHHC and 2D NHHC are assigned, they encode information about the $^{1}H - ^{1}H$ distances of the protons attached to the corresponding ^{13}C or ^{15}N nuclei. Therefore, ^{1}H-^{1}H distance restraints were derived from the assignments of the 2D CHHC and NHHC experiments. The ^{1}H-^{1}H distance restraints were generated from the assigned peaks of the CHHC and NHHC experiments by using the Distance bins function of the Restraint Distance Params section of CcpNMR. By using this function, the NOE peak intensities were related to the distance bounds of the generated distance restraints and the upper bound limit was defined as 5.0 Å. The interhelical hydrogen-bond restraints were generated for residues having a helical structure based on chemical shift and TALOS+. The distance between the H_{i+4} donor proton and the O_i acceptor oxygen was set to 1.8- 2.5 Å. A general scheme of structure calculation and refinement from solid state NMR data is outlined in Figure 5.5.

The structures of the Ost4 and Ost4V23D proteins were determined through simulated annealing procedure by using the CYANA program version 3.98.13 (*20*). During structure calculation, a total of 62 and 71 unambiguous distance restraints, 60 and 68 dihedral angle restraints, and 25 and 25 H-bonds were used for the Ost4 and Ost4V23D proteins, respectively. A total of 100 conformers were calculated using the standard simulated annealing schedule with 8000 torsion angle dynamics steps per conformer. Finally, the 20 conformers containing the lowest final target (energy) function values were analyzed using software packages Pymol (*21*), VMD(*22*), and Chimera (*23*). The quality of the Ost4 and Ost4V23D structures were evaluated using PROCHECK (*24*). The statistical parameters of the ssNMR structures of both the proteins are provided in Table 5.5.

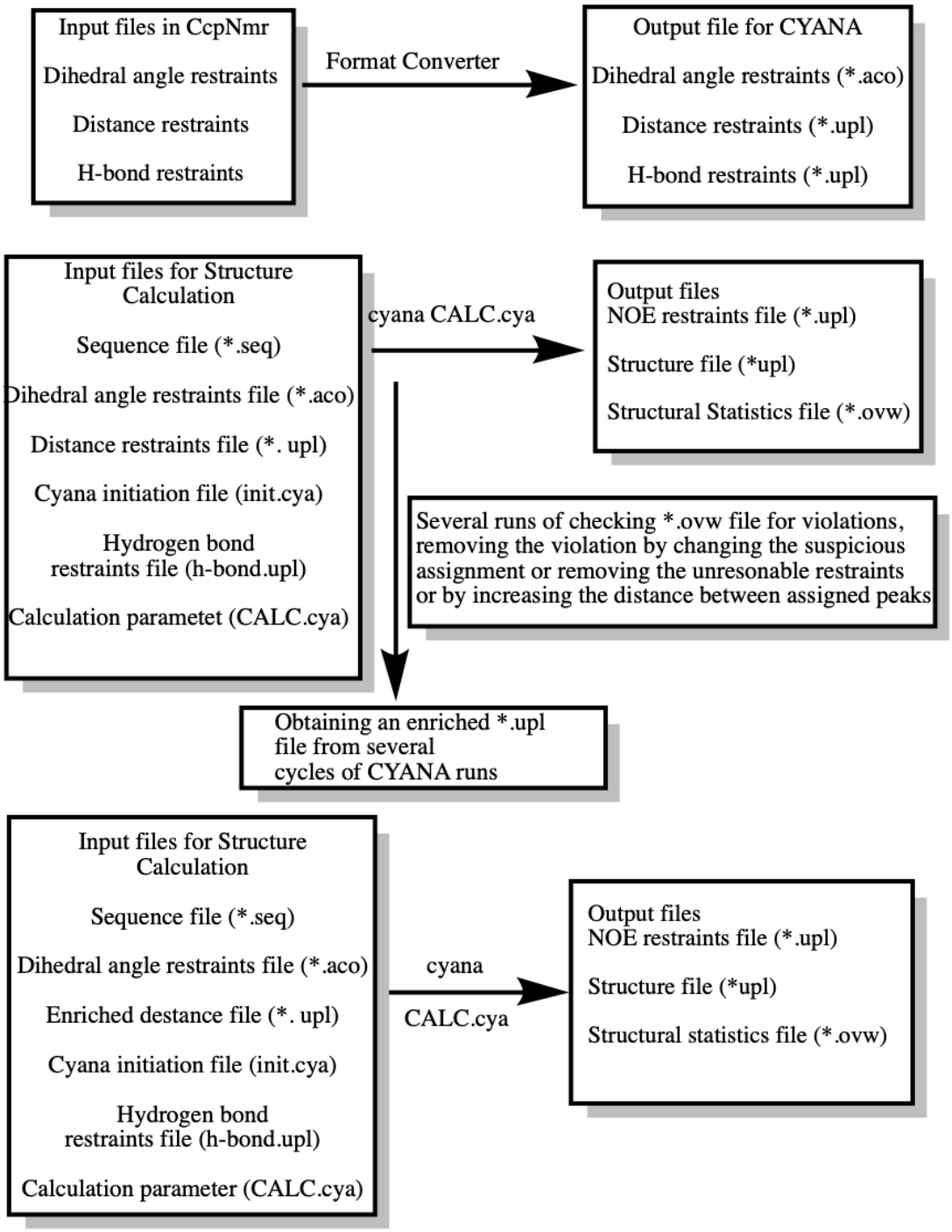


Figure 5.5: A schematic representation of the steps followed for the calculation of Ost4 and Ost4V23D structures by using ssNMR data.

Table 5.5: Structural statistics and experimental NMR restraints for 20 lowest energy conformers of Ost4 and Ost4V23D in the lipid bilayer

NOE	Ost4	Ost4V23D
Total	62	72
Intra residue, |i-j|=0	6	16
Sequential, |i-j|<=1	19	23
Medium range, 1<|i-j|<5	37	32
Long range, |i-j|>=5	0	1
Hydrogen bond restraints	25	25
Dihedral angle restraints	60	64
Inter residue distance restraints violations	0	0
RMSD to average structures, Å		
Backbone (residues 4-32)	0.17 ± 0.05	0.43 ± 0.15
Heavy atoms (residues 4-32)	0.69 ± 0.07	0.99 ± 0.09
Ramachandran plot outliers, %		
Residues in most favored regions	95 %	86.4 %
Residues in additionally allowed regions	4.4 %	13.3 %
Residues in generously allowed regions	0.6 %	0.3 %
Residues in disallowed regions	0 %	0 %

5.2.10 Computational details of Ost4 and Ost4V23D

5.2.10.1 System preparation

The protein-bilayer systems of the Ost4 and Ost4V23D proteins were built with the membrane builder module of the web-based graphical interface of CHARMM-GUI (*25, 26*) by using model 1 of the ssNMR ensemble and POPC-POPE lipids. During the preparation of the Ost4- and Ost4V23D-bilayer systems, the ratio of POPC-POPE was maintained to 3:2 to comply with the bilayer composition used for ssNMR experiments. This yielded the bilayer of a system containing 310 lipid molecules in the upper leaflet and 310 lipid molecules in the lower leaflet. The systems having the box dimensions of 10.08 Å x 10.08 Å x 9.96 Å were then solvated with 18627 TIP3P water molecules (*27*). The sodium ions were then randomly added by the solvent replacement method to neutralize the systems. The systems were then energy minimized by using 500 steps of the steepest descent method and CHARMM36 forcefield (*28*) using GROMACS (Version 2018) (*29*).

5.2.10.2 Molecular dynamics simulations

The molecular dynamics simulation for each system was performed under the isothermal-isobaric (NPT) periodic boundary conditions using the GROMACS simulation package version 2018 (*29*). A cutoff distance of 1.2 nm was used for nonbonded interactions and particle mesh Ewald summation method was applied to treat the long-range electrostatic interactions (*30*). All covalent bonds involving hydrogen atoms were constrained using the LINCS algorithm (*31*). The temperature was controlled at 303.15 K using Nose-Hoover coupling (*32*) with a 1ps coupling time constant and the pressure was controlled at 1 bar Parrinello-Rahman coupling (*33, 34*) with a pressure-relaxation time of 5 ps. A simulation

time step of 2 fs was used throughout the MD simulation to maintain the stability of the simulation with the protein bilayer system. After NVT equilibration, each system was NPT equilibrated by subjecting them to three steps of restrained MD simulations at 303.15 K without any restrain. Finally, each system was simulated for 250 ns under NPT equilibration condition. The molecular dynamics trajectories of each simulation were saved every 50 ps. The last 30 ns of each simulation were extracted from the whole trajectory for analysis. The root mean square deviation (RMSD) and the number of protein-lipid hydrogen bonds were used to verify the stability of the protein-lipid bilayer systems. Some of the computing for this project was performed at the OSU High Performance Computing Center at Oklahoma State University.

5.3 Results and Discussion

5.3.1 Reconstitution of Ost4 and Ost4V23D into the lipid bilayer

Lipid vesicles (liposomes) contain an enclosed lipid bilayer. Lipid bilayers have been frequently used to reconstitute membrane proteins. A variety of membrane proteins such as rhodopsin (*35*), P Protein kinase C (*36*), and G protein (*37*), have been successfully reconstituted into lipid bilayers. However, it has been difficult to control the concentration of inserted protein due to limited available area in the liposome. The factors: lipid/protein ratio, detergent concentration, and pH of the buffer for reconstitution were optimized. The pH of the buffer for solubilization of the protein into CHAPS detergent was set at 6.5. The CHAPS/lipid mass ratio was kept at 1:1.2. As shown in Figure 5.6, the reconstitution of both the proteins was successfully performed in four steps: preparation of large unilamellar vesicles, a saturation of preformed liposomes by CHAPS detergent, the

addition of CHAPS solubilized proteins into CHAPS saturated liposomes, and CHAPS detergent removal by using hydrophobic Bio beads.

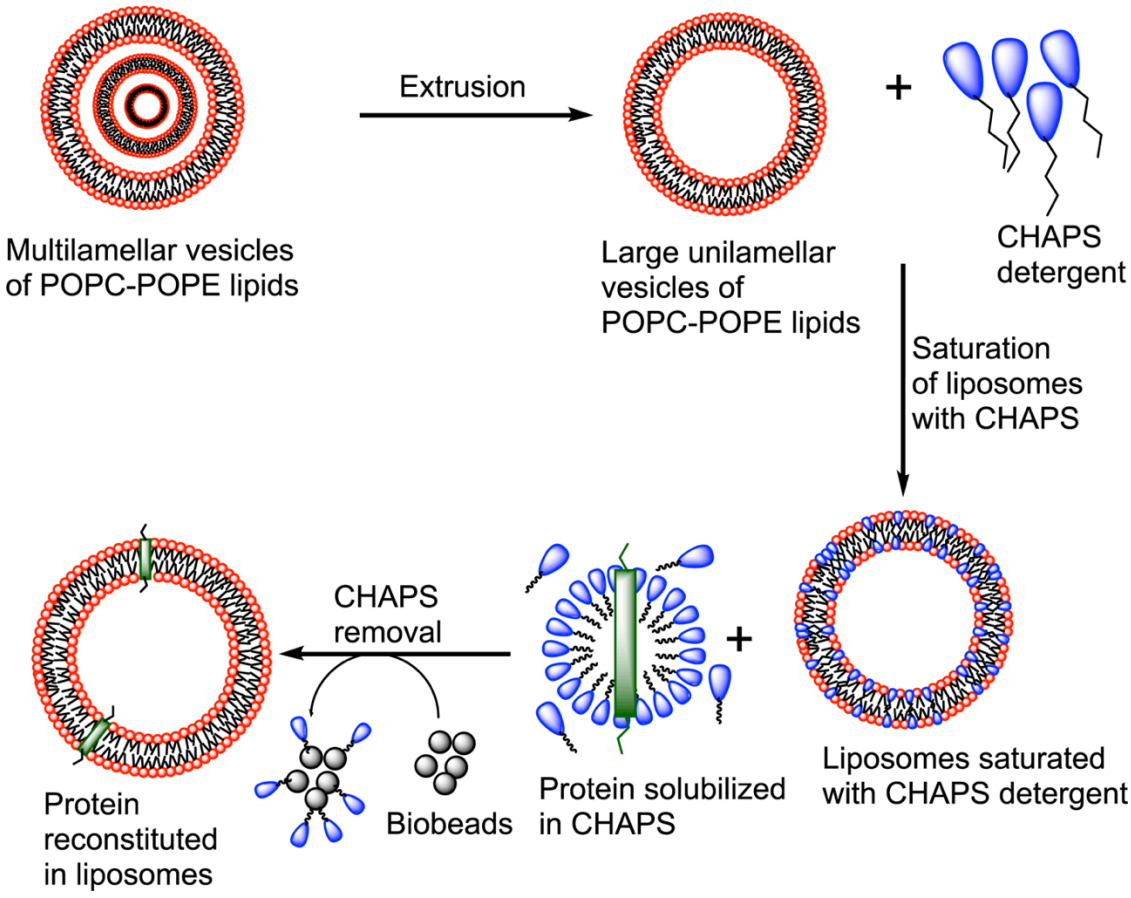


Figure 5.6: Schematic representation of steps of reconstitution of Ost4 and Ost4V23D proteins into liposome of POPC-POPE lipids by using Bio beads of CHAPS detergent removal. Large unilamellar vesicles of POPC-POPE lipids mixture in 3:2 ratio was prepared by an extrusion process using 0.1 μm membrane. The steps of reconstitution are labeled.

5.3.2 Backbone chemical shift assignment of Ost4 and Ost4V23D by ssNMR

For the characterization of Ost4 and Ost4V23D proteins, the backbone and sidechain (^{15}N and ^{13}C) chemical shifts were assigned by using 2D and 3D ssNMR experimental data sets. The assignments were completed by using ccpNMR Backbone and side-chain chemical shifts of Ost4 and Ost4V23D proteins have been provided in Table 5.3 and 5.4, respectively. Figure 5.7a and 5.7b show the results of 2D [^{13}C-^{13}C] DARR correlation experiments recorded on [^{13}C, ^{15}N] labeled Ost4 and Ost4V23D proteins. Intra-residue correlations in the 2D [^{13}C-^{13}C] experiments are labeled. As a representative, the strip plot displayed in Figures 5.8 and 5.9 provides a correlation between the resonances of the residues in the helical region of Ost4 and Ost4V23D, respectively. All the resonances of residues Met[1] and Ile[2] and the ^{15}N resonance of Ser[3] in Ost4 were not identified. The backbone ^{15}N chemical shifts of Val[17] and ^{13}CO chemical shifts of His[26], and all the chemical shifts of Pro[34] in Ost4V23D were not identified (Figure 5.10 a and b).

The secondary structures of Ost4 and Ost4V23D were accessed independently by using CcpNmr V2.1 (*18*) and the secondary chemical shifts $\Delta\delta C\alpha$ and $\Delta\delta C\alpha$-$\Delta\delta C\beta$. The secondary chemical shifts $\Delta\delta C\alpha$, and $\Delta\delta C\alpha$-$\Delta\delta C\beta$ for both the proteins were derived by using the SSP program (*38*). The chemical shifts of Ost4 deviated significantly upon mutation of Val[23] residue to Asp (Figure 5.11 a and b). The deviation of chemical shift upon V23D mutation in Ost4 indicates that the environment and 3D structures of these two proteins in the lipid bilayer might be different. The deviations in the chemical shifts between ssNMR and solution-NMR were observed throughout the sequences of both the proteins (Figure 5.12 a, b, c, and d). The analysis of secondary chemical shift resulted in a

similar backbone fold that was observed in solution-NMR (*39*) except that the helical portion in the ssNMR was a little longer in both the proteins. Comparison of the secondary structure of these proteins derived from the ΔδCα and ΔδCα-ΔδCβ secondary chemical shifts and predicted from CcpNmr results in a profile indicating a similar secondary structure with a minor variation in the helical pattern (Figure 5.13a and b).

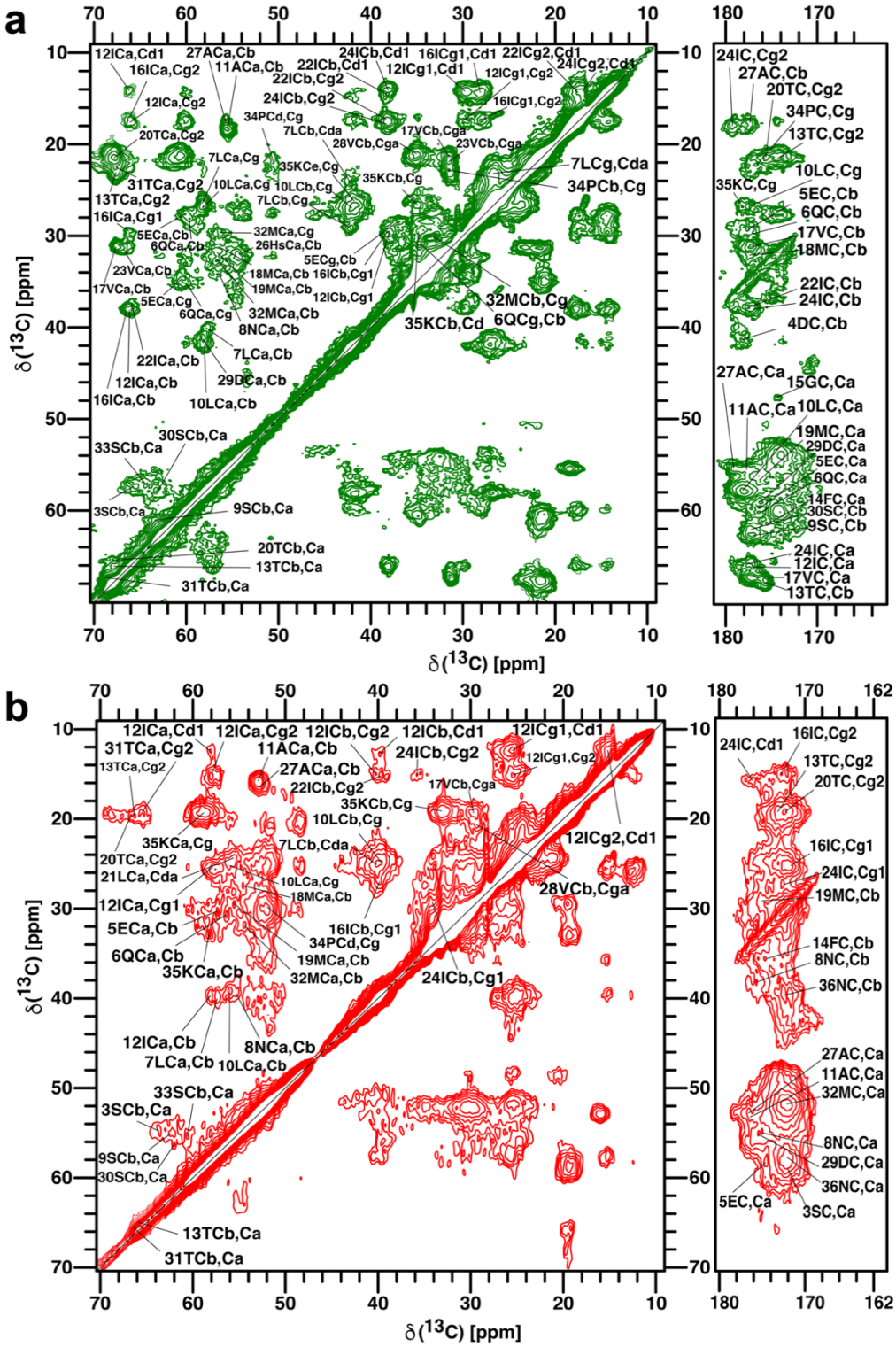


Figure 5.7: 2D [^{13}C-^{13}C] DARR correlation spectra of (a) Ost4 and (b) Ost4V23D with a DARR mixing time of 50 ms. The MAS ssNMR spectra were acquired on the samples prepared by reconstitution of these proteins into the POPC-POPE lipid bilayer system.

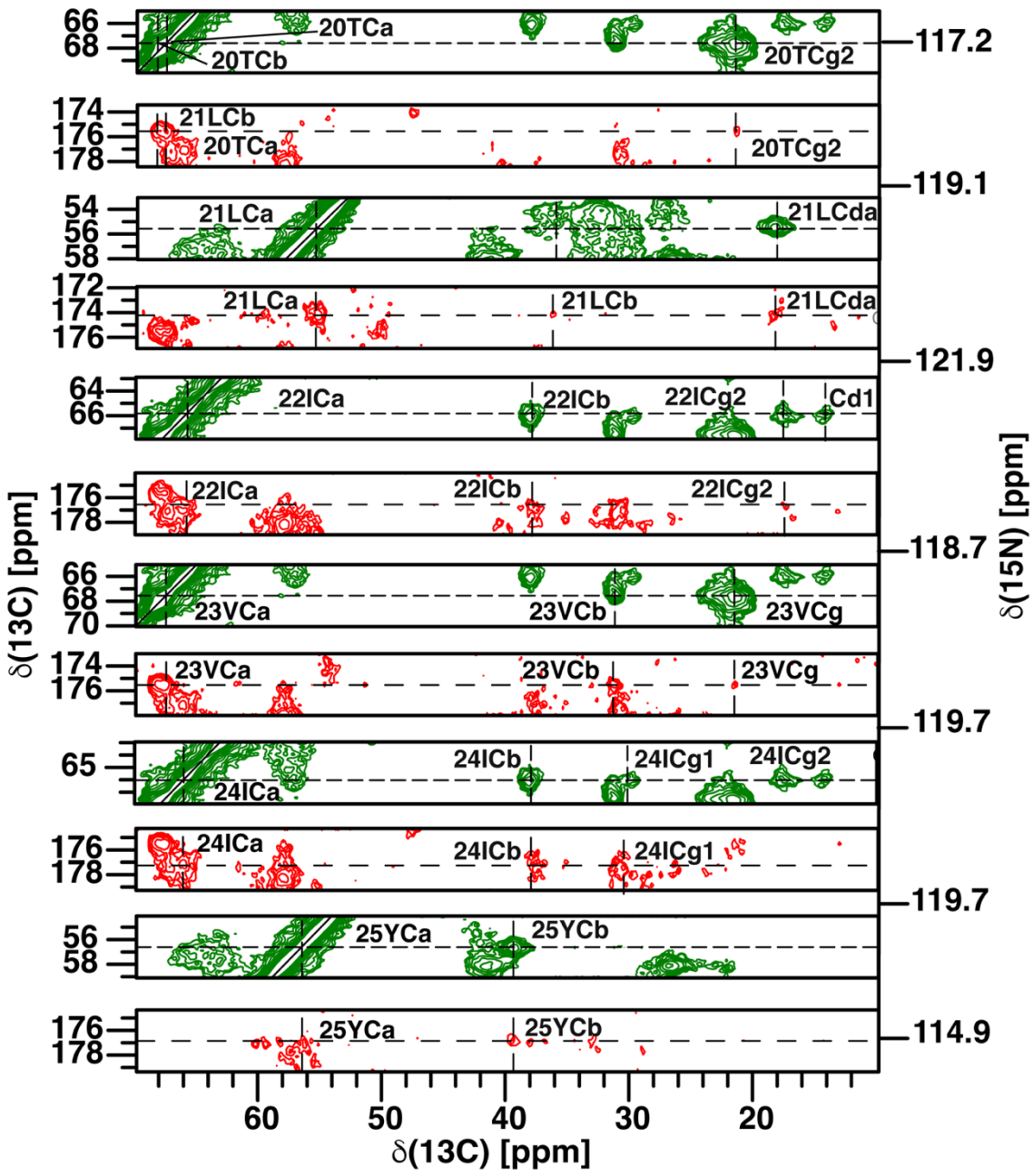
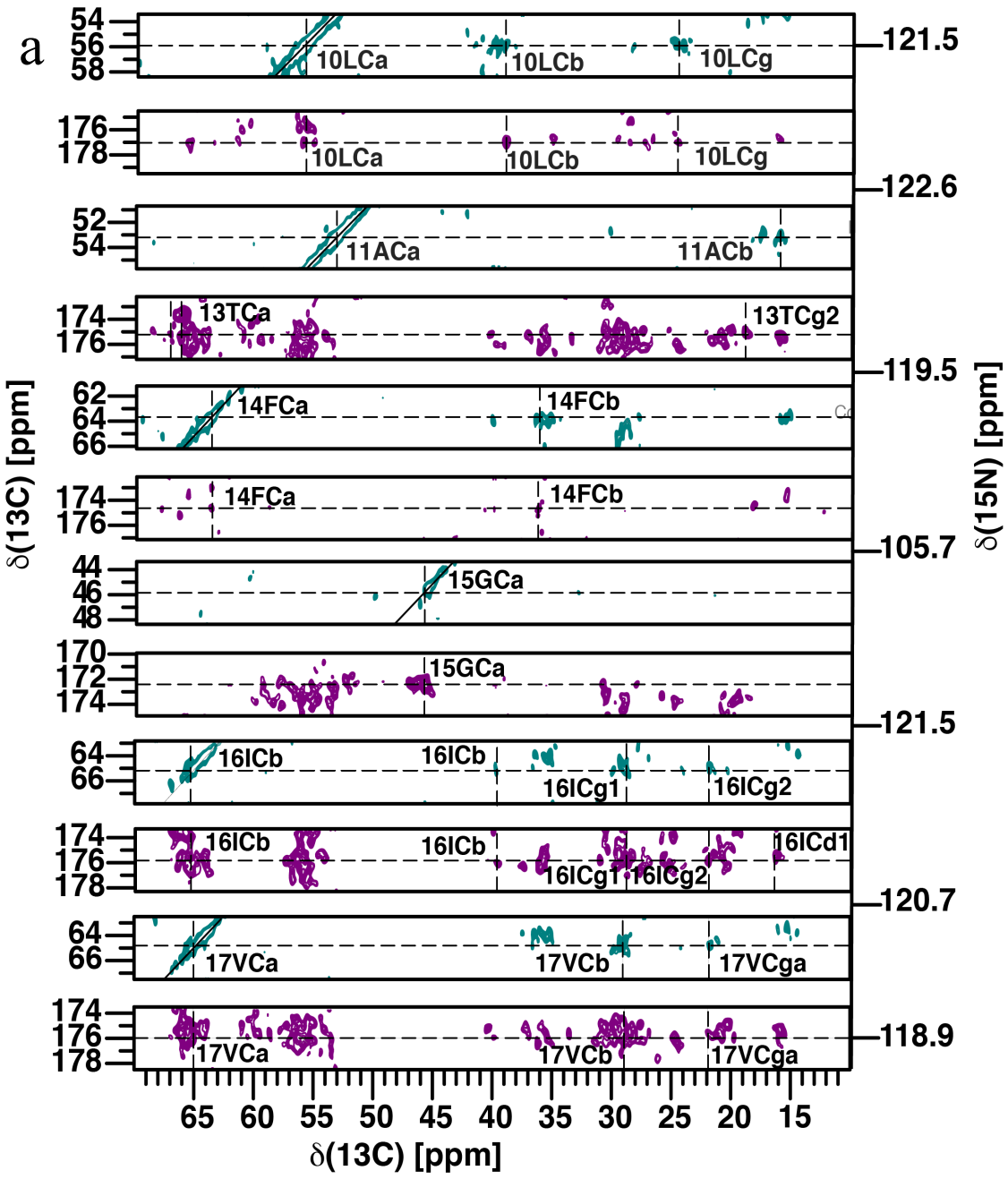


Figure 5.8: Representative strip plot of NCACX (green) and NCOCX (red) spectra of Ost4 protein for residues Thr^{20}-Tyr^{25} showing the sequential connectivity. The spectra were recorded at 700 MHz with 12.5 kHz MAS, and 50 ms DARR mixing time.


a
10LCa
10LCb
10LCg
11ACa
11ACb
13TCa
13TCg2
14FCa
14FCb
15GCa
16ICb
16ICg1
16ICg2
16ICd1
17VCa
17VCb
17VCga
δ(13C) [ppm]
δ(15N) [ppm]
121.5
122.6
119.5
105.7
121.5
120.7
118.9
65 60 55 50 45 40 35 30 25 20 15

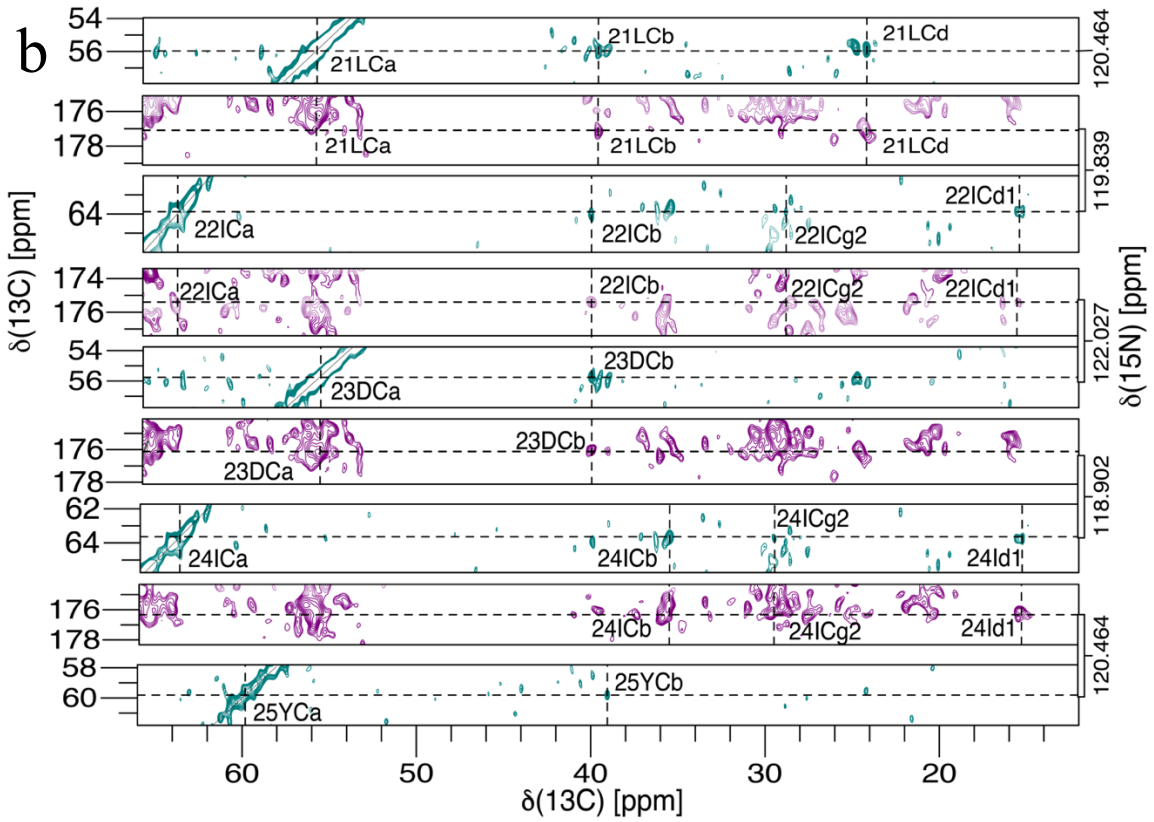


Figure 5.9: Representative strip plot of NCACX (teal) and NCOCX (magenta) spectra of Ost4V23D protein for residues Leu[10]-Val[17] (a) and residues Leu[21]-Tyr[25] (b) showing the sequential connectivity. The spectra were recorded at 700 MHz with 12.5 kHz MAS, and 50 ms DARR mixing time.

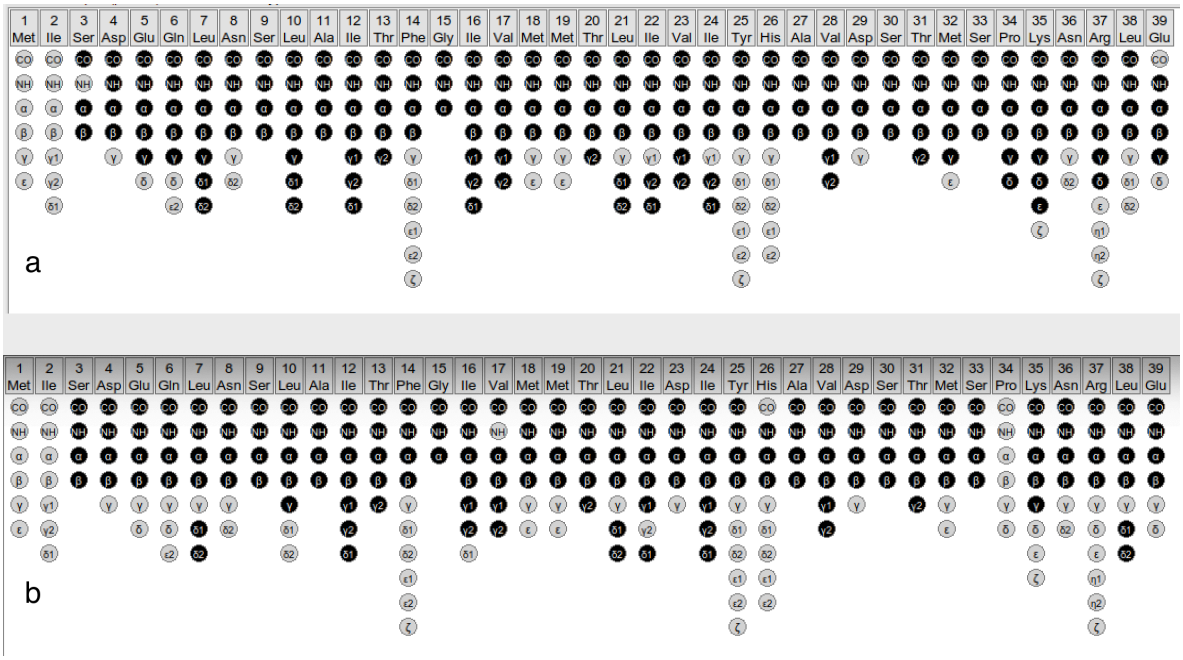


Figure 5.10: Assignment graph of Ost4 (a) and Ost4V23D (b) generated from CcpNMR software. The only resonances highlighted with black circles were identified by using 2D and 3D ssNMR data sets.

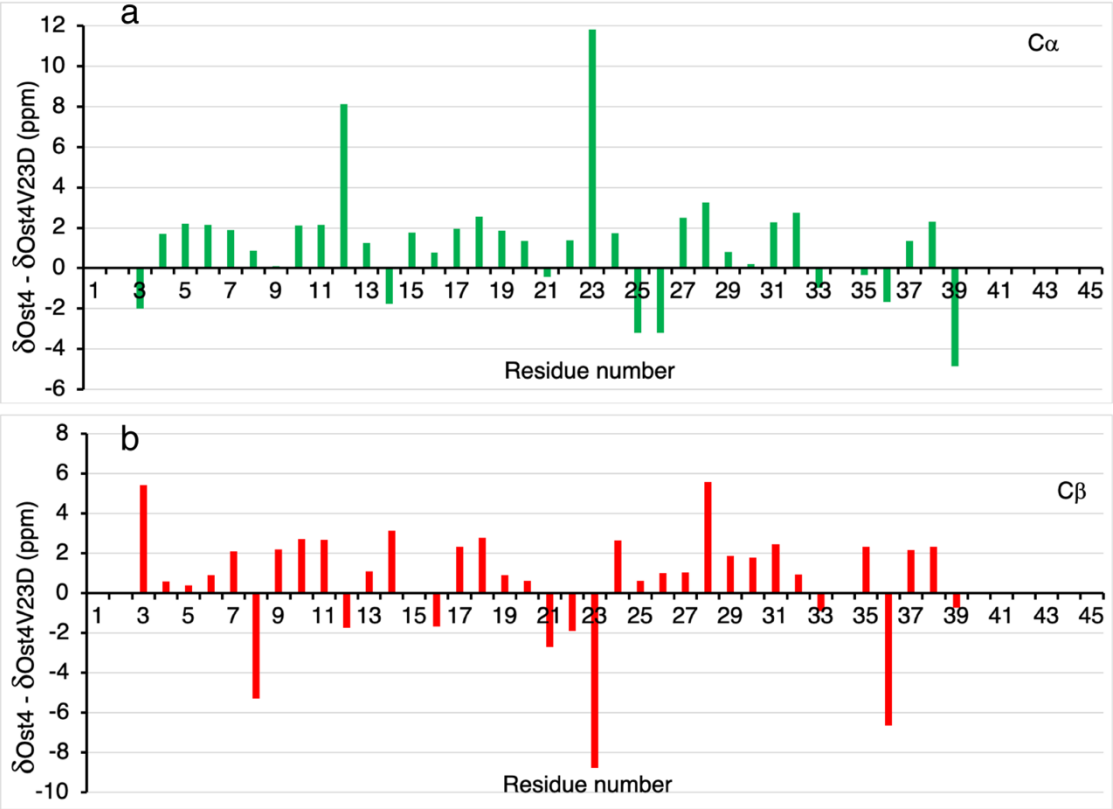


Figure 5.11: Comparison of ssNMR (a) Cα and (b) Cβ chemical shifts differences upon mutation of V23 residue to D in Ost4 protein. A significant deviation of Cα and Cβ chemical shifts upon mutation is an indication of the presence of either structure or environment variation due to mutation.

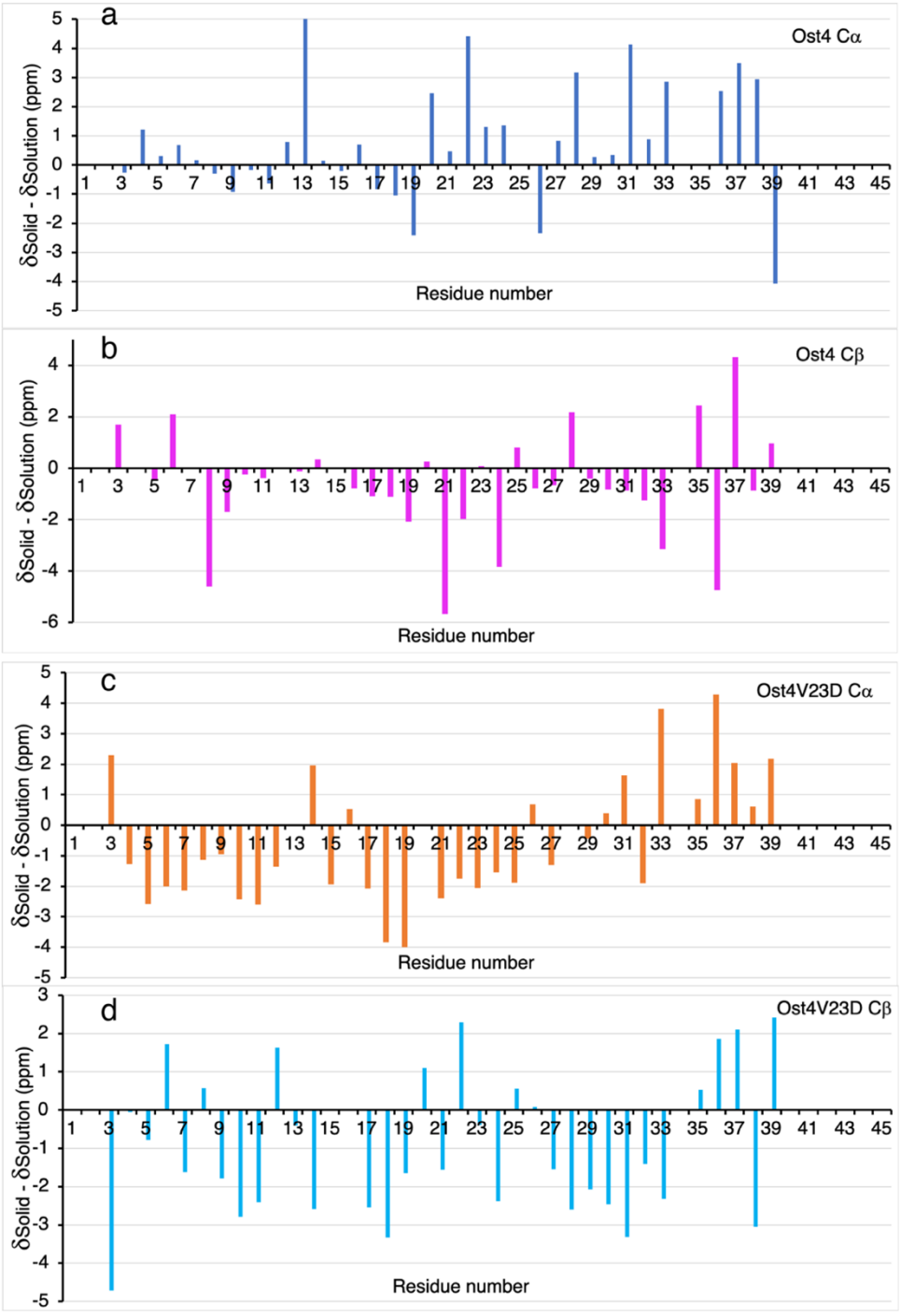


Figure 5.12: Comparison of chemical shift differences in (a) Cα, (b) Cβ of Ost4 (c) Cα, and (d) Cβ of Ost4V23D obtained by solution-state and solid-state NMR spectroscopy methods.

The difference in the helical propensities has been reported previously (*39, 40*). The analysis of helical propensity based on secondary chemical shifts indicates that the helical structure in these two proteins are different. However, the overall helical region in both the proteins spanned from Asp^{4}-Asn^{36} (Figure 5.13a and b). Excluding ^{1}H and the backbone resonances of the residues that do not belong to either Ost4 or Ost4V23D, a total of 93.5% and 89.8% of backbone assignments were obtained for Ost4 and Ost4V23d proteins, respectively. The chemical shift lists for the Ost4 and Ost4V23D proteins assigned by ssNMR 2D and 3D experiments are demonstrated in Tables 5.3 and 5.4, respectively.

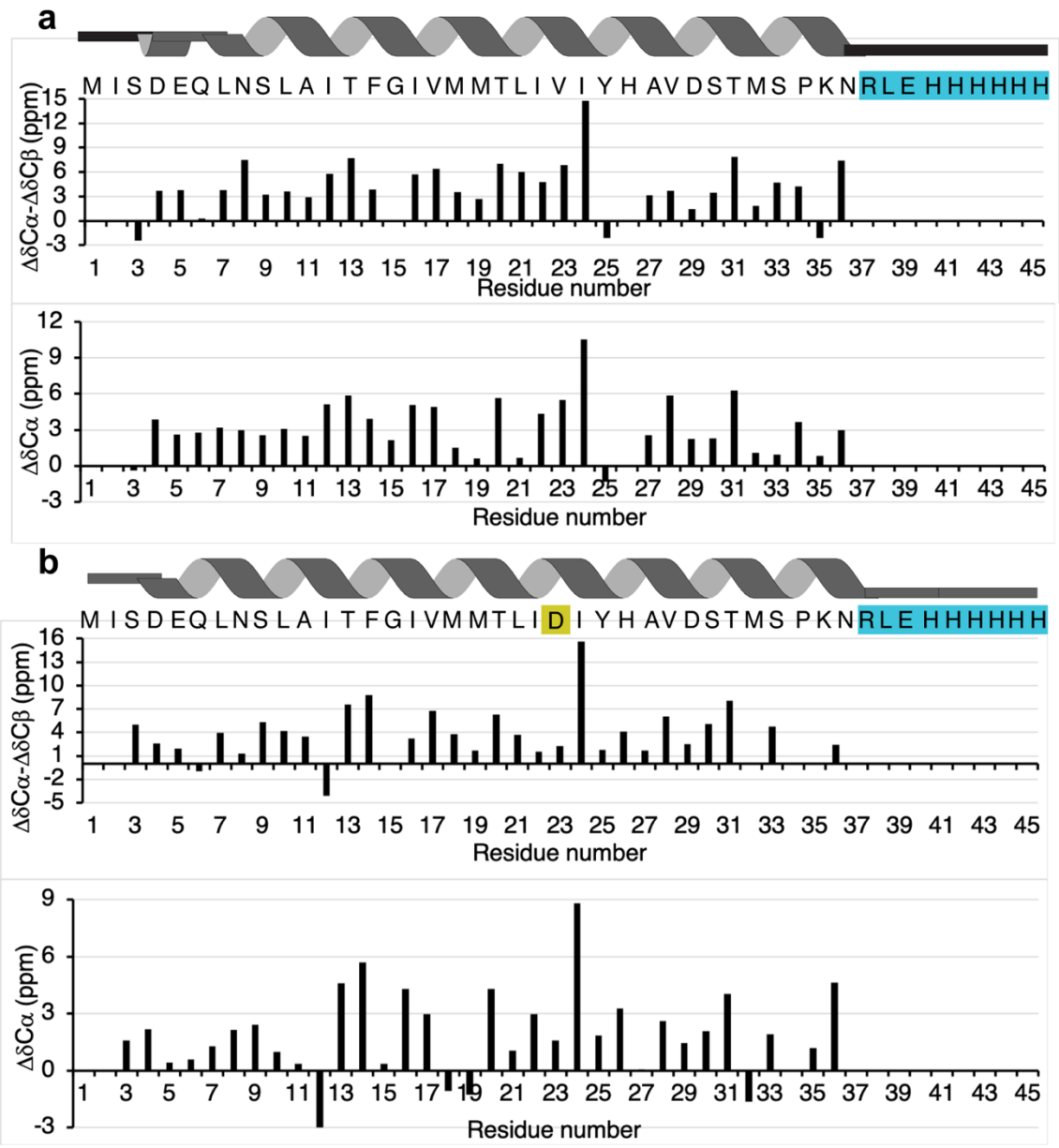


Figure 5.13: CcpNMR derived secondary structure, protein sequence, ΔδCα, and ΔδCα-ΔδCβ secondary chemical shifts: (a) Ost4 and (b) Ost4V23D. The secondary chemical shift values were obtained by using the SSP program (*38*). The residues that do not belong to Ost4 and Ost4V23D are highlighted in cyan. The mutated residue, D23, in the Ost4V23D protein is highlighted in yellow.

Solid-state NMR spectroscopy is well suited for the structure determination of membrane proteins in lipid bilayers. The membrane protein structures in the native-like system, the lipid bilayer, reflects the biologically relevant structure. Our results from ssNMR on the Ost4 and Ost4V23D proteins could provide a significant reference for correlating structural and functional implications of these proteins in a membrane environment.

5.3.3 Magic angle spinning ssNMR spectroscopy

The uniformly (^{13}C, ^{15}N)-labeled dry samples of Ost4 and OstV23D proteins reconstituted in a POPC-POPE lipid bilayer were packed into a 3.2 mm rotor for structure determination by MAS ssNMR. A combination of different multidimensional MAS ssNMR experiments was acquired, processed, and analyzed to solve the 3D structures of these proteins. The assignments of the ^{13}C, ^{15}N backbone and sidechain resonance of these proteins have been discussed earlier in this chapter. The distance restraints for structure determination were derived from a combination of 2D ^{13}C-detected DARR CHHC (carbon-proton-proton-carbon) experiments and 2D ^{13}C, ^{15}N-detected DARR NHHC (nitrogen-proton-proton-carbon) experiments (*16*). The NHHC and CHHC spectral resolutions of both proteins were sufficiently good for obtaining intra- and inter-residue distance restrains generation (Figure 5.14 and 5.15). The significant deviation of backbone and sidechain ^{13}C, ^{15}N chemical shifts were detected upon the V23 to D mutation in Ost4 indicating a significant change in wither protein structure or environment.

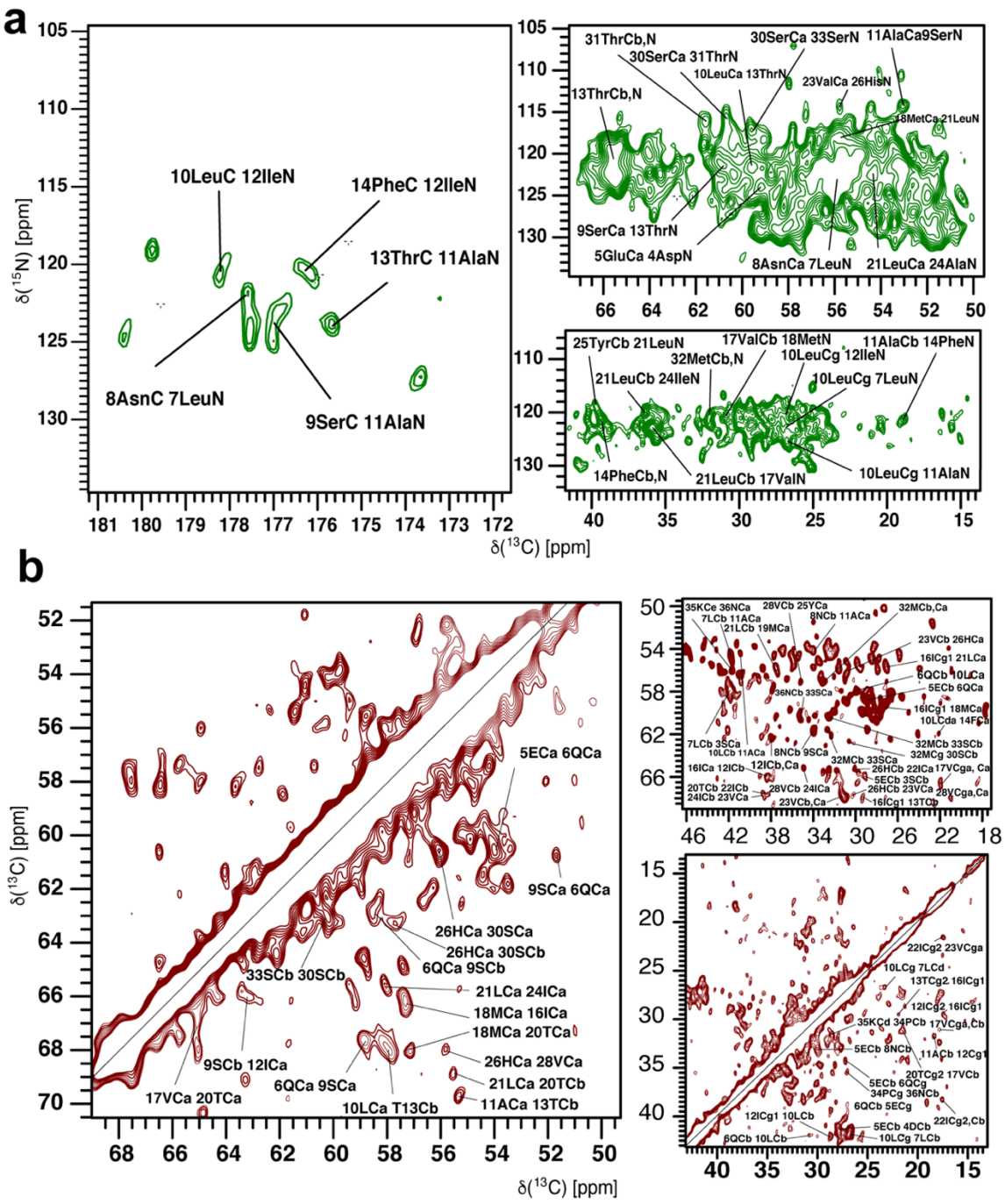


Figure 5.14: Parts of 2D [^{13}C, ^{13}N] DARR NHHC (a) and [^{13}C, ^{13}C] DARR CHHC (b) spectra of Ost4 acquired with a mixing time of 300 μs. The cross-peaks corresponding to inter- and intra-residue restraints of the protein are labeled. The peaks of the residues that do not belong to protein (R37, L38, E39, and H40-H45) are not labeled.

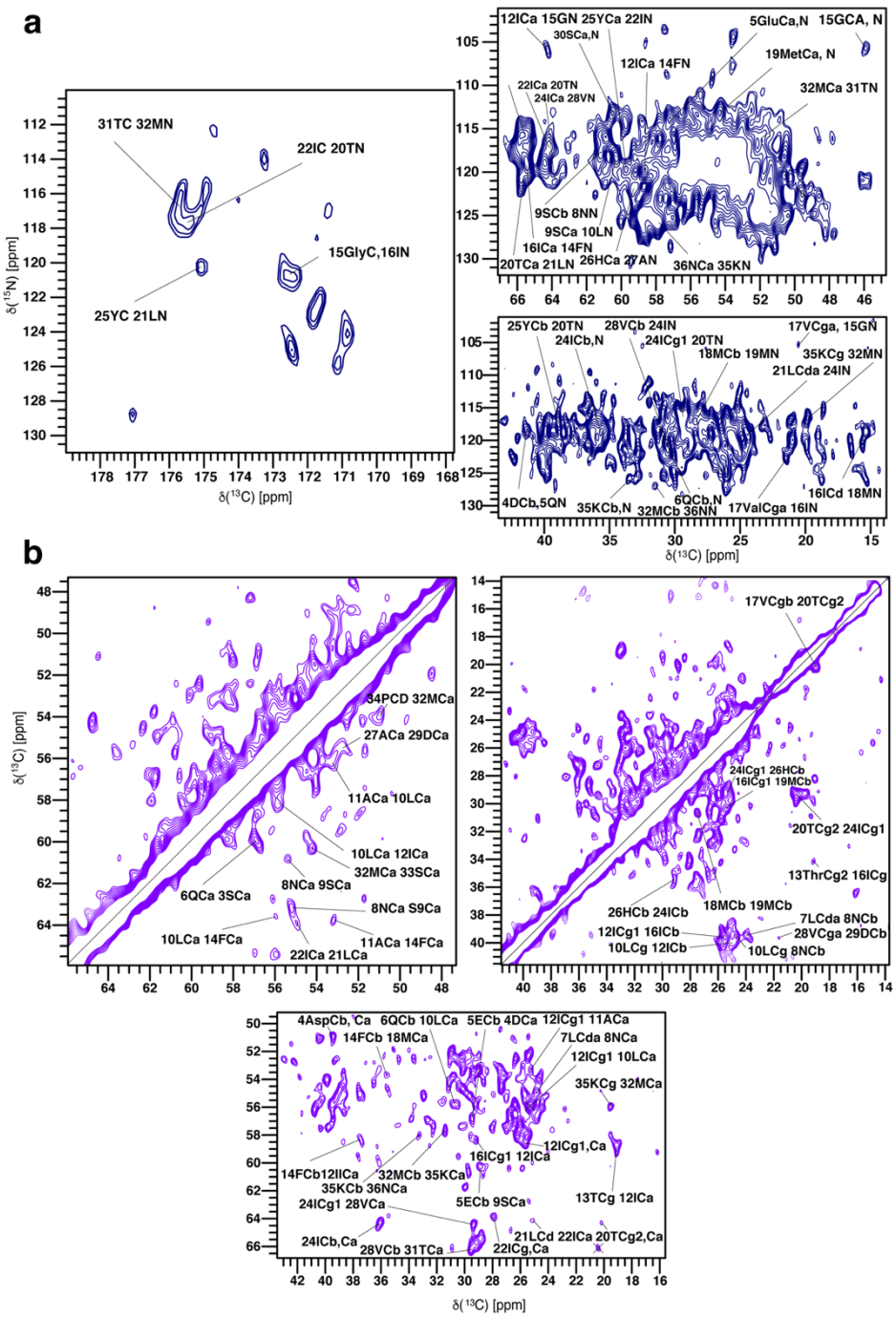


Figure 5.15: Parts of 2D [^{13}C, ^{15}N] DARR NHHC (a) and [^{13}C, ^{13}C] DARR CHHC (b) spectra of Ost4V23D acquired with a mixing time of 300 μs. The cross-peaks corresponding to inter- and intra-residue restraints of the protein are labeled. The peaks of the residues that do not belong to protein (R37, L38, E39, and H40-H45) are not labeled.

The secondary structure propensity (SSP) scores for individual residues in each protein were calculated using $^{13}C\alpha$ and $^{13}C\beta$ chemical shift values using the SSP program (*38*) (Figure 5.16). The SSP scores of these proteins demonstrated 48.9% and 57.6% α–structure propensity for Ost4 and Ost4V23D protein, respectively. This observation was consistent with the previously reported helical content for these proteins in micelles as seen by CD and solution-state NMR (*39, 40*). The consistency in the similar helical propensities for Ost4 and Ost4V23D proteins observed from different approaches indicates these proteins have similar behavior in a variety of membrane mimicking systems, however, Ost4 is either present in a different environment or has a different structure than Ost4V23D.

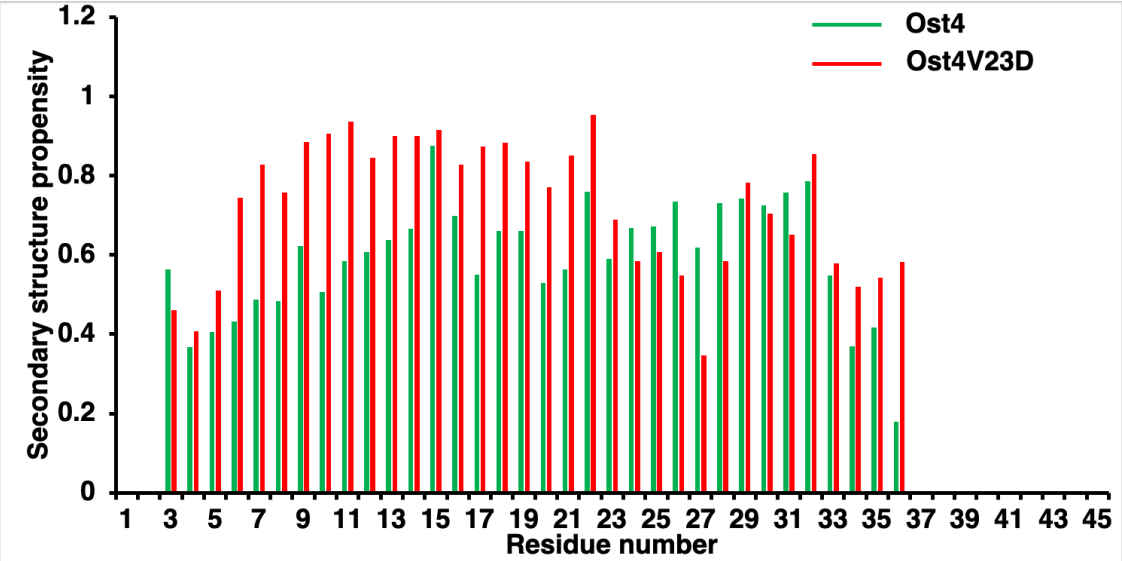


Figure 5.16: Residue-specific secondary structure propensities of Ost4 (green bars) and Ost4V23D (red bars). The SSP scores were calculated using $^{13}C\alpha$ and $^{13}C\beta$ chemical shift values with an SSP limit of 1.2. Overall helical content of Ost4 (48.9%) was increased to 57.6% upon mutation of V23 to D in the Ost4 protein. This observation was consistent with the previously reported results (*39, 40*).

5.3.4 Structure of Ost4 and Ost4V23D in the lipid bilayer

The atomic resolution structures of the Ost4 and Ost4V23D proteins were determined in the lipid bilayer by using ^{13}C-^{13}C and ^{13}C-^{15}N distance restraints which encodes information about ^{1}H-^{1}H distances attached to ^{13}C, and ^{15}N atoms from 2D [^{13}C, ^{13}C] CHHC and 2D [^{13}C, ^{15}N] NHHC MAS ssNMR experiments, respectively. The intra- and inter-residue distance restraints were generated from the ^{13}C-^{13}C and ^{13}C-^{15}N cross-peaks of these spectra by using the Make Distance Restraints module of the CcpNMR software package. All the distance restraints were combined and converted to CYANA format by using the "Format Converter" module of CcpNMR. The dihedral angle restraints were extracted using DANGLE. Additionally, hydrogen bonds between the backbone oxygen atom of the residue i and the backbone amide proton of the residue (i+4) were employed, as semirigid bodies, to improve the convergence of the structure. The structure of the Ost4 and Ost4V23D proteins were calculated by using these restraints in the structure determination program CYANA 3.98.13 (*20*).

The 20 best structures having the lowest energy functions were selected for analysis. Ost4 and Ost4V23D displayed backbone and heavy atoms (for residue Asp^{4}-Met^{32}) RMSDs of 0.17 ± 0.05 and 0.43 ± 0.15 Å and 0.69 ± 0.07 and 0.99 ± 0.09 Å, respectively (Table 5.5) indicating a tight ensemble of the 20 best representative structures. Figure 5.17 a, b, and c and d, e, and f display ensembles of the 20 conformers with the lowest energy function, ribbon representations of one the conformers, and helix orientations of Ost4 and Ost4V23D proteins, respectively. As shown in Figure 5.17, The structures calculated with all the restraints displayed a well-defined α-helical protein for both the proteins. Ost4 displayed

a straight α-helix spanning from residue 4-31 (Figure 5.18). This structure of Ost4 was consistent with the structure of Ost4 from the yeast OST complex determined in either nanodisc or digitonin by cryo-EM method (*41, 42*). This result indicates that the structure of Ost4 obtained from ssNMR data in the POPC-POPE bilayer can represent the structure present in the whole OST complex. In contrast to Ost4, which contained a well-defined straight α-helix (Figure 5.17b), the Ost4V23D ssNMR structure displayed a bent helix (Figure 5.17e) with a bent angle of 135.6º (Figure 5.17f). Although the helix in Ost4V23D started bending at the V23 mutation site (Figure 5.18), the helix angle was measured at H26. This suggests that V23D mutation is capable of creating a kink in the wildtype (WT) protein resulting in the disruption of important interactions of WT Ost4 to near transmembrane helix (TMH) of the Stt3 protein. The kinking due to the V23D mutation may be a key reason for the breakage of almost all the hydrophobic interaction that have been reported recently (*43*). In addition to a kink formation, the length of α-helix of Ost4V23D was calculated to be longer than that of Ost4. The helix in Ost4V23D mutated protein encompassed residues 3-35 (Figure 5.18). This result of an increase in helical content upon V23D mutation was consistent with the previously reported results (*39, 40*).

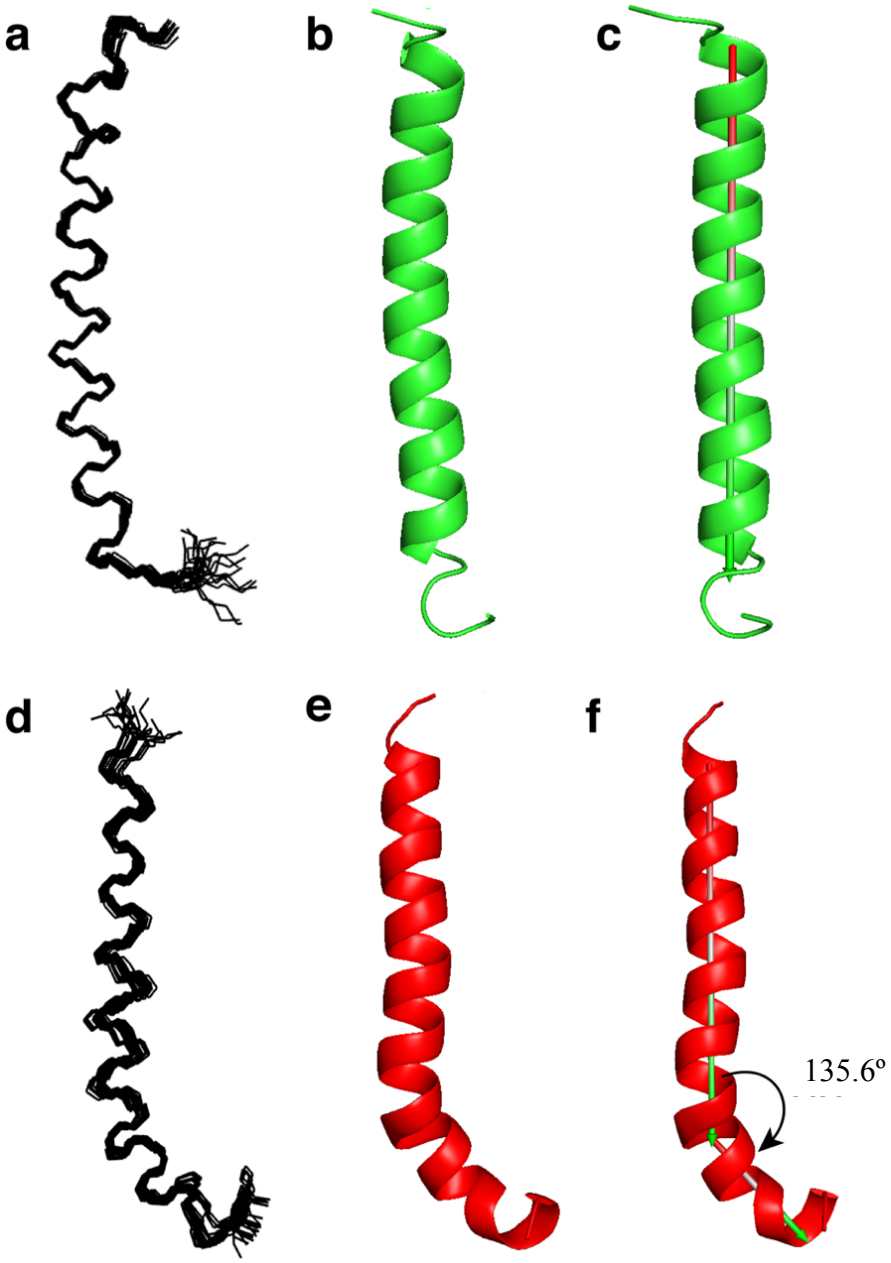


Figure 5.17: Atomic resolution NMR structural models of Ost4 and Ost4V23D determined by MAS solid-state NMR. Superposition of 20 conformations of Ost4 (a) and Ost4V23D (d) having the lowest target function. Ribbon representation of one of the structures of Ost4 (b) and Ost4V23D (c). Helix orientations of Ost4 (c) and Ost4V23D (f). The mutation of V23 to D in Ost4 results in bending the helix by 135.6º. The helix angle was measured by using the Pymol software package (*21*).

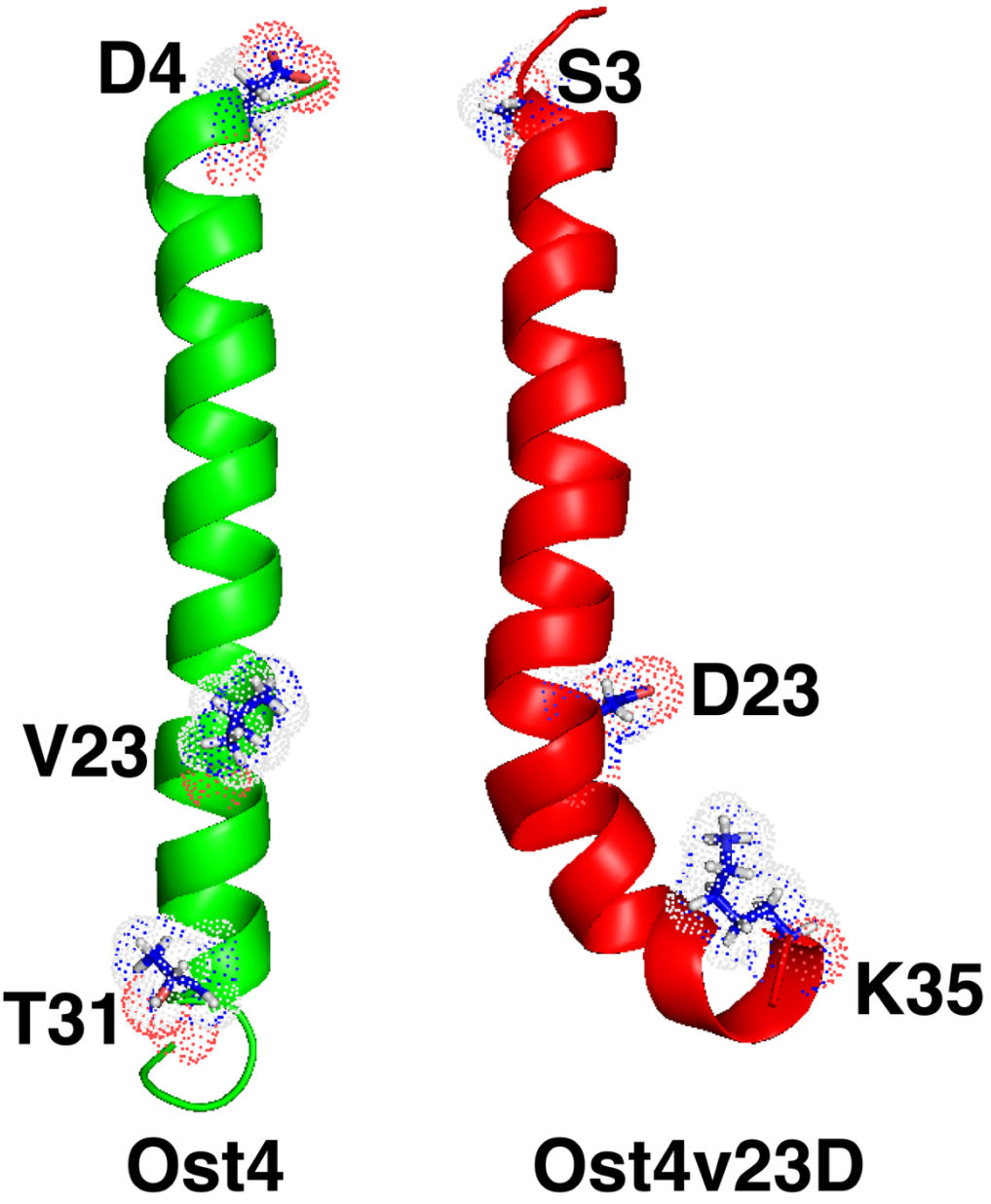


Figure 5.18: Ribbon representations of ssNMR structure of Ost4 (a) and Ost4V23D (b) proteins displaying the helix length and bending to helix upon mutation of V23 residue to D. Ost4 contains a straight α-helix encompassing from residue 4-31 and Ost4V23D contains a bent helix encompassing residues 3-35. The beginning, place of mutation, and ending of the Ost4 α–helix and the beginning, mutated residue, and ending residues of the Ost4V23D helix are highlighted with stick and dot representation.

5.3.5 Comparison of structures of Ost4 and Ost4V23D in lipid bilayer to that in micelles

Recent solution-state NMR structures of yeast Ost4 and Ost4V23D on recombinant proteins have shown both proteins to contain a straight α–helix encompassing residues Asp^{4}-Met^{32} (Figure 5.19a and b) (*43*). In contrast, structures of chemically synthesized yeast and human Ost4 in a mixed aqueous organic solvent system was reported to contain a kink in the transmembrane helix (Figure 5.19c and d) (*44, 45*). The recent solution NMR structure of yeast Ost4 in micelles contained a straight α–helix formed by residues Asp^{4}-Met^{32}. The present MAS solid-state NMR structure of Ost4 in the bilayer is also a straight α–helix formed by residues Asp^{4}-Thr^{31} (Figure 5.18). Thus, the helix in the ssNMR structure was shortened by one residue in the C-terminal end. However, this ssNMR structure of Ost4 was identical to the Ost4 structure extracted from the yeast OST complex that contained a straight α–helix with the same residues (Asp^{4}-Thr^{31}) (Figure 5.19e and f). The ssNMR structure of Ost4V23D determined in a lipid bilayer was quite different from the solution-state NMR structure determined in DPC micelles (*43*). While the solution NMR structure of Ost4V23D reconstituted in DPC micelles contained a straight α–helix formed of residues Asp^{4}-Met^{32}, solid-state NMR structure of the same protein determined in a lipid bilayer contained a bent helix formed of residues Ser^{3}-Lys^{35} (Figure 5.18 and 5.19b). The differences in the micelle and bilayer structures for Ost4V23D protein could be due to the difference in the hydrophobic environment of DPC micelles from the lipid bilayer. The lipid bilayer can provide a near native membrane-like environment.

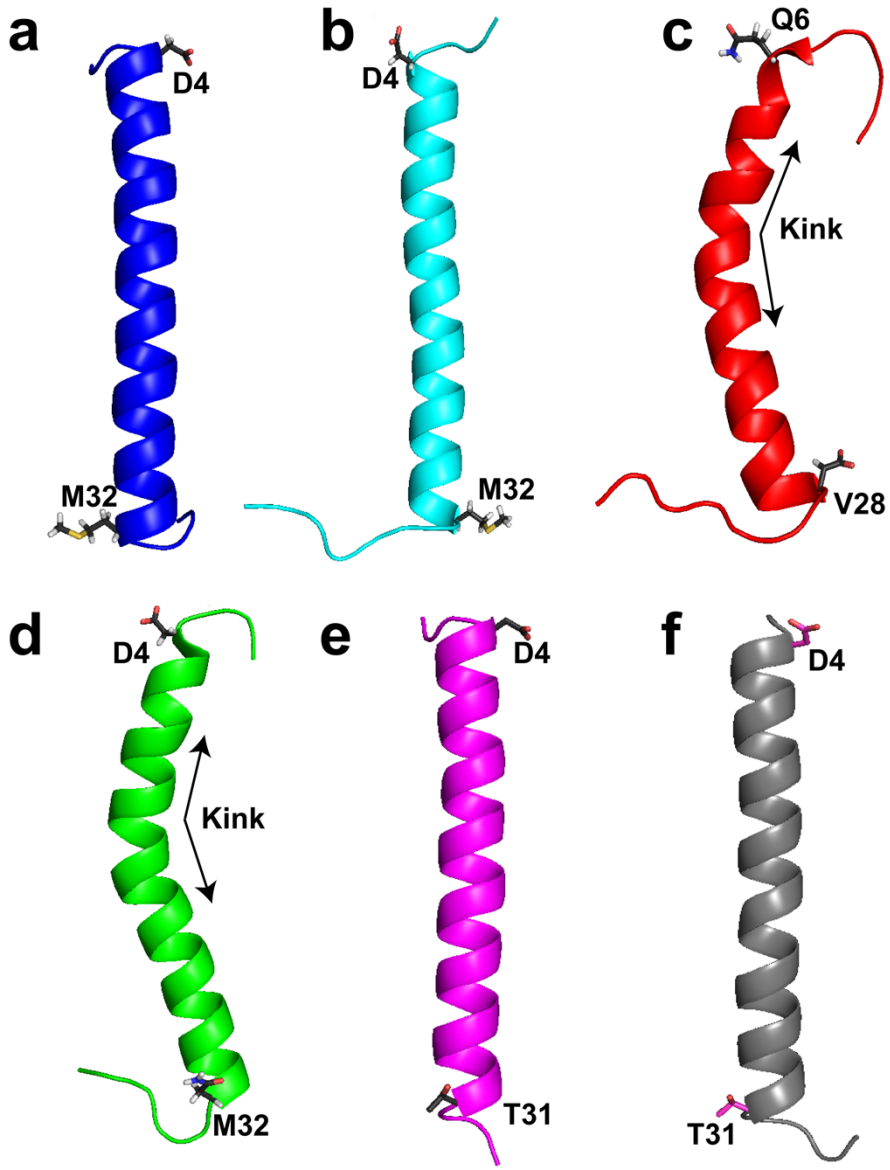


Figure 5.19: Ribbon representations of structures of Ost4 determined in different membrane mimetic systems using various techniques. Solution-state NMR structures of yeast Ost4 (PDB ID 6XCR) (a) and Ost4V23D (PDB ID 6XCU) (b) in DPC micelles contain a straight α–helix spanning residues 4-32. Solution-state NMR structures of yeast Ost4 (PDB ID 1RKL) (c) and human Ost4 (PDB ID 2LAT) (d) in mixed aqueous organic solvent system displayed a kinked helix. Structures of the Ost4 protein extracted from

cryo-EM structures of the yeast OST complex determined in nanodisc (PDB ID 6EZN) (e) in digitonin (PDB ID 6C26) (f) show a straight α-helix encompassing residues 4-31.

5.3.6 Molecular Dynamic studies of Ost4 and Ost4V23D in the lipid bilayer

Characterization of the topological properties of the Ost4 and Ost4V23D proteins in a native membrane-like environment is crucial for understanding the biological behavior of these proteins. In order to probe the structure function of Ost4 and Ost4V23D in the lipid bilayer, the MD simulations of these proteins were performed in the lipid bilayer system of the POPC-POPE lipid mixture for 250 ns. Since the protein-bilayer systems were equilibrated sufficiently, the average backbone RMSD of both the proteins exhibited a minimal fluctuation (~0.2 nm) until the end of the simulation time indicating a stable simulation system (Figure 5.20a). The number of H-bond formated between protein and lipid molecules was calculated by setting the hydrogen donor and acceptor distance of ≤ 3.5 Å. The number of stronger H-bonds between protein and lipid bilayer in Ost4 were higher than those of the Ost4V23D-lipid bilayer system (Figure 5.20b) suggesting that Ost4 favors being inside the hydrophobic environment of the lipid bilayer.

MD simulations of the Ost4-lipid bilayer system and Ost4V23D-lipid bilayer systems resulted in interesting outcomes. The Ost4 protein crossed the lipid bilayer in an α-helical conformation maximizing the hydrogen bonding between the peptide bonds following the unique orientation of the transmembrane proteins in the membrane. The Ost4 protein traversed through the lipid bilayer with a tilt angle of 23º with respect to the bilayer normal. The protein maintained the property of a regular α–helix (Figure 5.21a and b) during the

simulation time. Indeed, the helix length of Ost4 protein increased up to residue M32 (Figure 5.22a). The solution-state NMR structure of Ost4 in DPC micelles also contained a straight helix formed by residues Asp^{4} – Met^{32}. This result indicated that DPC micelles can serve as an alternative membrane mimicking system to characterize small membrane proteins. The Ost4V23D protein exhibited a unique behavior in the lipid bilayer. At the end of the simulation, the V23D mutated protein underwent a dramatic structural transformation. The protein in the lipid bilayer existed perpendicular to the membrane plane. However, the kink that was seen in the ssNMR structure of the Ost4V23D protein disappeared at the end of the MD simulation. The α–helical part of the protein formed by residues His^{26}-Lys^{35} transformed into a random coil structure (Figure 5.21c, d, and Figure 5.22b). The helix uncoiling of the Ost4V23D protein, resulting from MD simulation, could be another reason for the disruption of several hydrophobic interactions between Ost4 and TMH12 and TMH13 of Stt3 that are responsible for maintaining the stability of the OST complex.

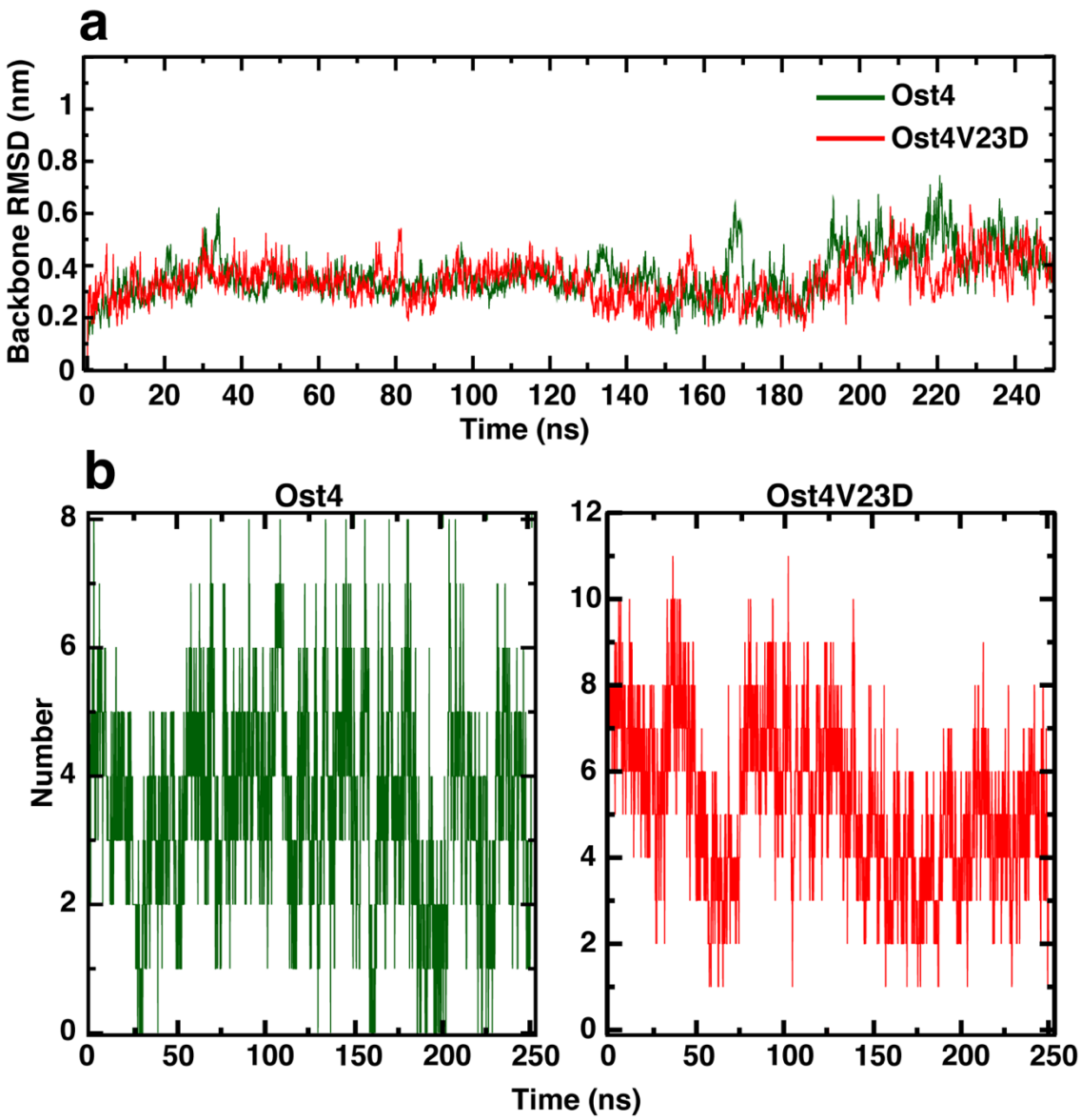


Figure 5.20: Backbone RMSD (a) and number of H-bonds (b) of Ost4 and Ost4V23D plotted along the 250 ns of MD simulation.

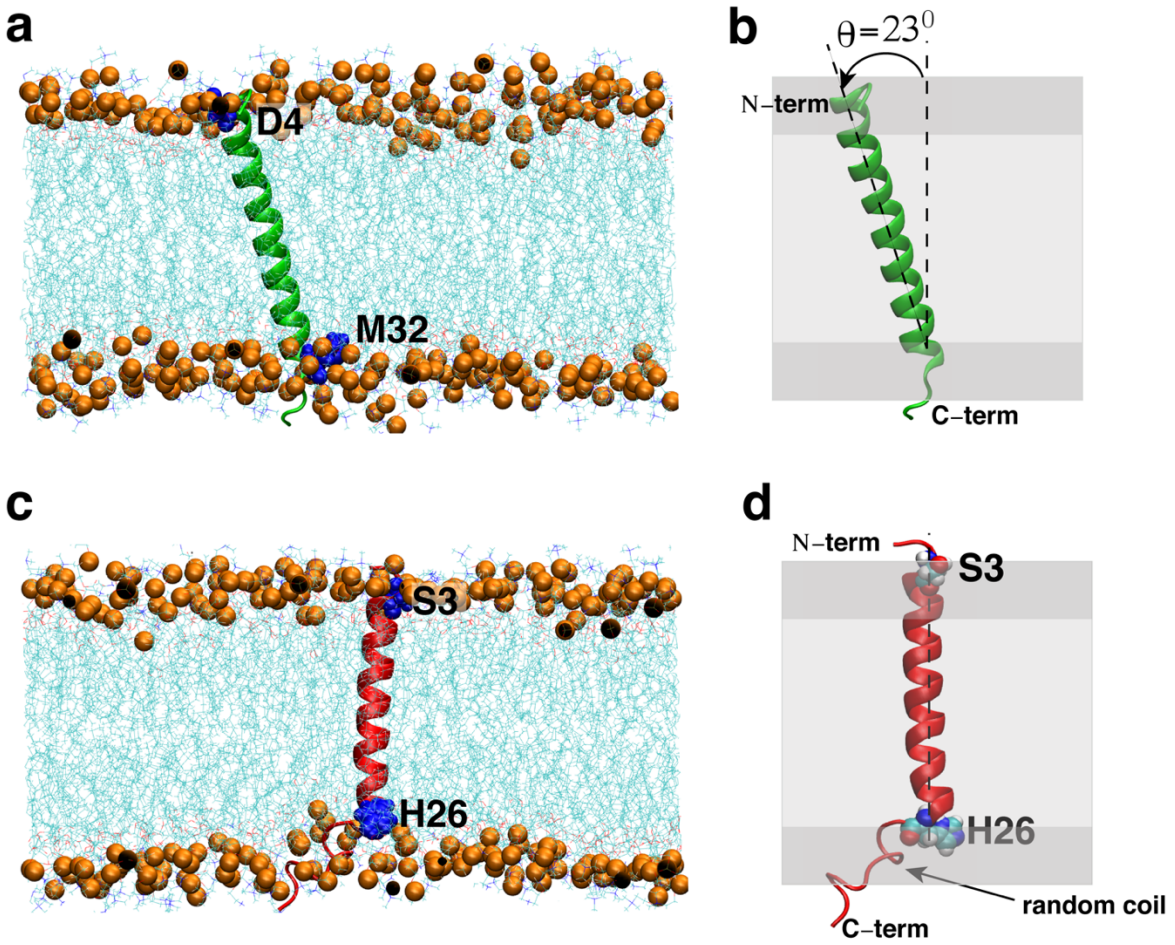


Figure 5.21: Topology of Ost4 and Ost4V23D in a POPC-POPE lipid bilayer. A snapshot of the last frame of the Ost4-lipid bilayer MD simulations (a), representation of Ost4 in lipid bilayer displaying tilt angle of the transmembrane domain (b), a snapshot of the last frame of the Ost4V23D-lipid bilayer MD simulations (c), and the orientation of Ost4V23D in the lipid bilayer. The α–helical regions of the proteins are labeled.

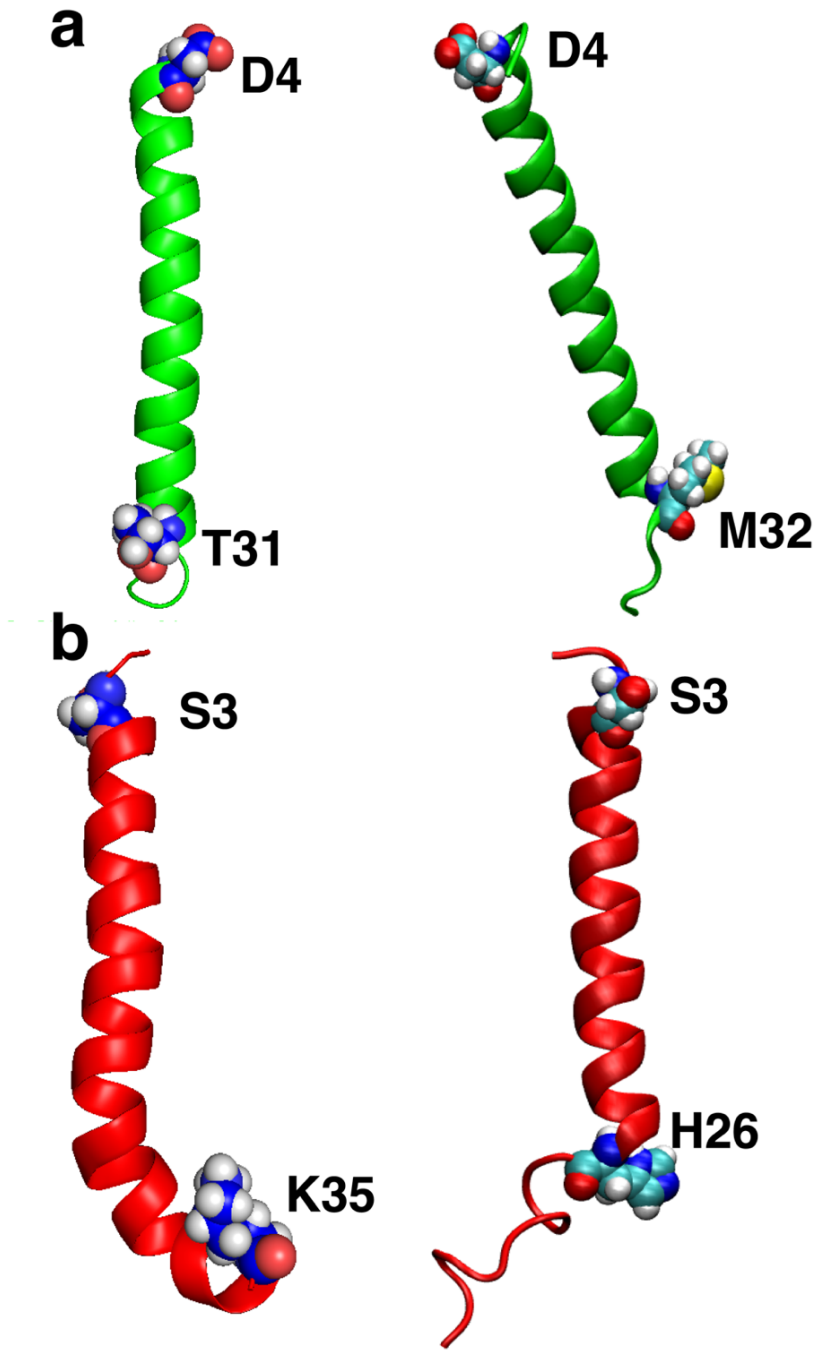


Figure 5.22: Ribbon representation of Ost4 (a) and Ost4V23D (b) indicating the structural transition upon performing MD simulations in the lipid bilayer of POPC-POPE. The starting and ending of helices are indicated by residue letter codes and sphere representations for the residues. The left panel and right panel contain the structure of the proteins prior to and after the MD simulations, respectively.

The bilayer thickness of each simulation system and helix lengths at the end of the simulations were measured to gain further insight into membrane behavior and protein topology in the membrane-like environment. The lipid bilayer thickness of the Ost4-bilayer and the Ost4V23D-bilayer systems were identical (4.0 nm) (Figure 5.23a) indicating that both the membranes were stable and behaved similar to each of the proteins during the simulation time. The helix length of Ost4 protein (3.7 nm) increased by one residue after simulation (Figure 5.23b). In contrast, the helix length of Ost4V23D after simulation was shortened and measured to be 2.8 nm (Figure 5.23c) due to helix unfolding. In order to further study the conformational behavior of these two proteins in the lipid bilayer, the MD trajectories were analyzed at different time points of the simulations. The initial states of the Ost4-bilayer and the Ost4V23D-bilayer systems show that the Ost4V23D protein contains higher helical content.

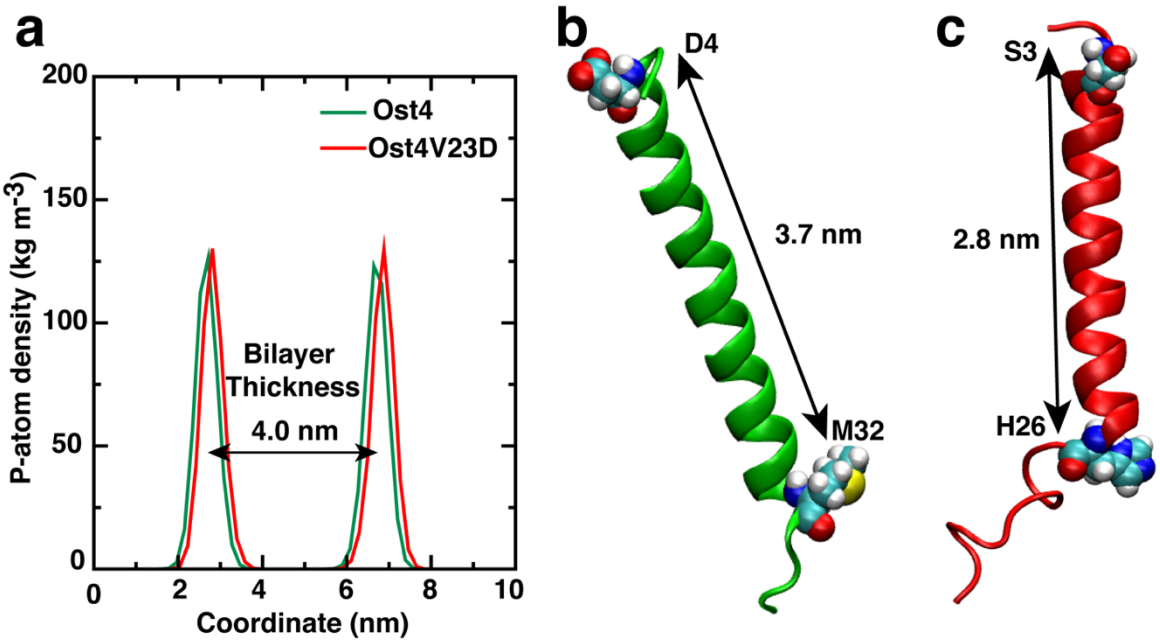


Figure 5.23: Analysis of bilayer thickness in MD simulations of Ost4-bilayer and Ost4V23D-bilayer systems of the Ost4 and the Ost4V23D proteins (a), Ost4 helix length (b), and Ost4V23D helix length (c) at the end of the simulation.

The initial α–helix of Ost4 was observed to be intact indicated by MD trajectories and secondary structure profile (Figure 5.24a and b). On the other hand, the Ost4V23D protein unfolded during the MD simulations and lost about 30% of its initial helicity (Figure 5.24c and d). This helix unfolding is an indication of an unfavorable α–helical structure of Ost4V23D in the membrane bilayer. This helix unfolding could result in the disruption of appropriate interactions of Ost4 to TMHs of Stt3 protein responsible for the stability of the Ost3 – Ost4 – Stt3 catalytic subcomplex and formation of inappropriate interactions thereby making the V23D mutant a temperature-sensitive phenotype in eukaryotes.

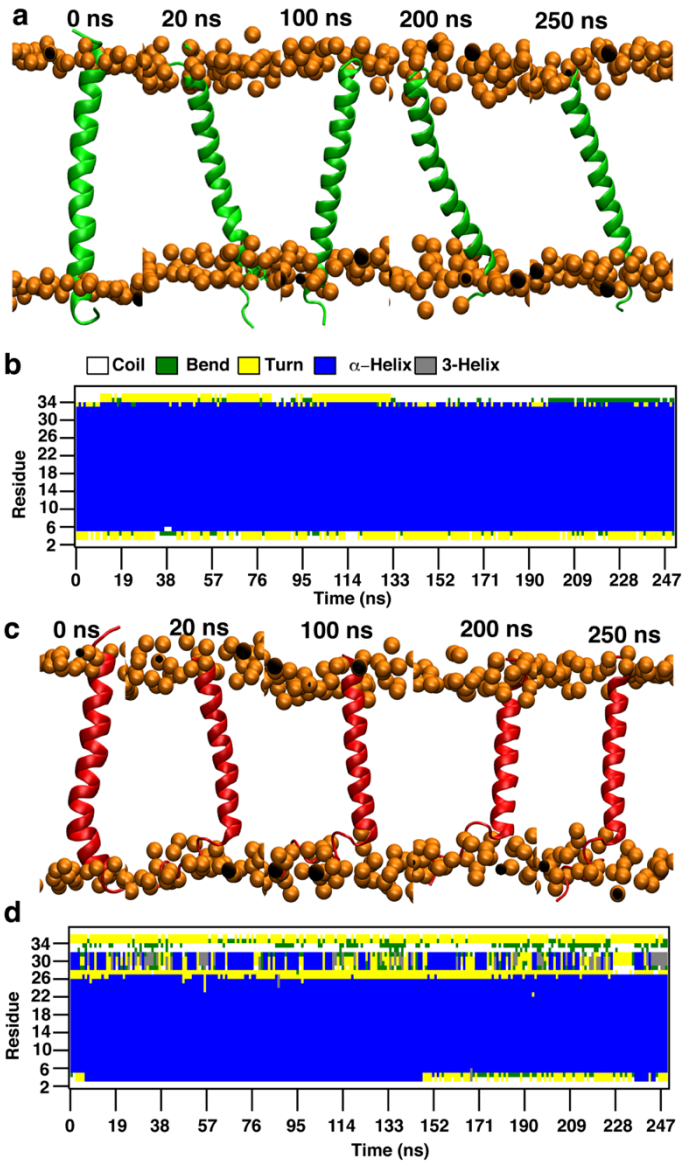


Figure 5.24: Analysis of the trajectories of Ost4 and Ost4V23D for the study of conformational changes of these proteins in the lipid bilayer. Snapshots of trajectories at five different time points in the MD simulation run of Ost4 (a) and Ost4V23D (c). The proteins are displayed by ribbon representations. The phosphorus atoms of POPC-POPE lipids are represented by spheres (orange color) to show the bilayer surface. The water molecules and other atoms of lipids are not shown for clarity purpose. The secondary structure profiles of Ost4 (b) and Ost4V23D (d) as a function of simulation time in the x-axis are also shown.

5.4 Conclusions

In this chapter, the reconstitution, ssNMR structures, MD simulation, and comparison of structures for Ost4 and its critical mutant Ost4V23D of yeast oligosaccharyltransferase have been discussed. The integral membrane protein Ost4 of catalytic subcomplex (Stt3-Ost4-Ost3/Ost6) and its critical mutant Ost4V23D were successfully reconstituted into the lipid bilayer of POPC-POPE mixtures. Since the functional activities of membrane proteins depend on the structure of the lipid molecules around the proteins, structural and functional aspects of the membrane proteins can be studied in a model membrane system that closely resembles the natural lipid bilayer. In addition to reconstitution in large unilamellar vesicles, the two proteins are characterized by ssNMR in a near native-like lipid bilayer setting. The ^{13}C and ^{15}N backbone side-chain resonances of Ost4 and Ost4V23D proteins were successfully assigned as a prerequisite for atomic resolution structure determination.

From the assigned ^{13}C and ^{15}N backbone and side-chain resonances, the distance restraints necessary for the structure calculation were generated. The ssNMR structures of Ost4 and Ost4V23D in the POPC-POPE lipid bilayer represent the structures in a membrane-like system. The ssNMR and MD simulation results of Ost4 and Ost4V23D in the lipid bilayer explained that V23D mutation creates structural deformation either by kink formation or by helix unfolding. The kinking or helix unfolding could be the key reason for the disruption of several hydrophobic interactions of Ost4 to TM12 and TM13 of the Stt3 protein. Our results open the door to investigate the structural and functional impact of Ost4V23D mutation in a natural membrane-like environment.

5.5 References

1. L. Wang, L. Tonggu, Membrane protein reconstitution for functional and structural studies. *Science China Life Sciences* **58**, 66-74 (2015).
2. H. H. Freeze, Disorders in protein glycosylation and potential therapy: Tip of an iceberg? *Journal of Pediatrics* **133**, 593-600 (1998).
3. V. Westphal, M. Xiao, P. Y. Kwok, H. H. Freeze, Identification of a frequent variant in ALG6, the cause of Congenital Disorder of Glycosylation-Ic. *Human mutation* **22**, 420-421 (2003).
4. H.-X. Zhou, T. A. Cross, Influences of membrane mimetic environments on membrane protein structures. *Annu Rev Biophys* **42**, 361-392 (2013).
5. D. D. Laws, H.-M. L. Bitter, A. Jerschow, Solid-State NMR Spectroscopic Methods in Chemistry. *Angewandte Chemie International Edition* **41**, 3096-3129 (2002).
6. D. G. Reid, NMR for physical and biological scientists, T. C. Pochapsky and S. S. Pochapsky. Taylor and Francis: New York, London. 2007. ISBN 9780815341031. *NMR in Biomedicine* **21**, 417-418 (2008).
7. N. Das, D. T. Murray, T. A. Cross, Lipid bilayer preparations of membrane proteins for oriented and magic-angle spinning solid-state NMR samples. *Nat Protoc* **8**, 2256-2270 (2013).
8. O. Domenech, S. Merino-Montero, M. T. Montero, J. Hernandez-Borrell, Surface planar bilayers of phospholipids used in protein membrane reconstitution: an atomic force microscopy study. *Colloids Surf B Biointerfaces* **47**, 102-106 (2006).

9. M. Lopes-Rodrigues, D. Zanuy, C. Aleman, C. Michaux, E. A. Perpete, 3D structure of a Brucella melitensis porin: molecular modelling in lipid membranes. *J Biomol Struct Dyn* **37**, 3923-3935 (2019).

10. M. Sharma *et al.*, Insight into the mechanism of the influenza A proton channel from a structure in a lipid bilayer. *Science* **330**, 509-512 (2010).

11. F. Separovic, J. A. Killian, M. Cotten, D. D. Busath, T. A. Cross, Modeling the membrane environment for membrane proteins. *Biophys J* **100**, 2073-2074; author reply 2075 (2011).

12. K. Oxenoid, J. J. Chou, The structure of phospholamban pentamer reveals a channel-like architecture in membranes. *Proc Natl Acad Sci U S A* **102**, 10870-10875 (2005).

13. K. Takegoshi, S. Nakamura, T. Terao, 13C–1H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chemical Physics Letters* **344**, 631-637 (2001).

14. J. Pauli, M. Baldus, B. van Rossum, H. de Groot, H. Oschkinat, Backbone and Side-Chain 13C and 15N Signal Assignments of the α-Spectrin SH3 Domain by Magic Angle Spinning Solid-State NMR at 17.6 Tesla. *ChemBioChem* **2**, 272-281 (2001).

15. W. T. Franks, K. D. Kloepper, B. J. Wylie, C. M. Rienstra, Four-dimensional heteronuclear correlation experiments for chemical shift assignment of solid proteins. *Journal of Biomolecular NMR* **39**, 107-131 (2007).

16. A. Lange, S. Luca, M. Baldus, Structural Constraints from Proton-Mediated Rare-Spin Correlation Spectroscopy in Rotating Solids. *Journal of the American Chemical Society* **124**, 9704-9705 (2002).

17. F. Delaglio *et al.*, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995).

18. T. J. Stevens *et al.*, A software framework for analysing solid-state MAS NMR data. *J Biomol NMR* **51**, 437-447 (2011).

19. M. S. Cheung, M. L. Maguire, T. J. Stevens, R. W. Broadhurst, DANGLE: A Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. *J Magn Reson* **202**, 223-233 (2010).

20. in *Protein NMR Spectroscopy: Practical Techniques and Applications*. pp. 159-192.

21. L. Schrodinger, The Pymol Molecular Graphics System, Version~1.8. (2015).

22. W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38 (1996).

23. E. F. Pettersen *et al.*, UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612 (2004).

24. R. A. Laskowski, J. A. Rullmannn, M. W. MacArthur, R. Kaptein, J. M. Thornton, AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**, 477-486 (1996).

25. S. Jo, T. Kim, V. G. Iyer, W. Im, CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry* **29**, 1859-1865 (2008).

26. J. Lee *et al.*, CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of Chemical Theory and Computation* **12**, 405-413 (2016).

27. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79**, 926-935 (1983).

28. J. Huang, A. D. MacKerell, Jr., CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem* **34**, 2135-2145 (2013).

29. H. J. C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **91**, 43-56 (1995).

30. U. Essmann *et al.*, A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **103**, 8577-8593 (1995).

31. B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **18**, 1463-1472 (1997).

32. W. G. Hoover, Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A Gen Phys* **31**, 1695-1697 (1985).

33. S. Nosé, M. L. Klein, Constant pressure molecular dynamics for molecular systems. *Molecular Physics* **50**, 1055-1076 (1983).

34. M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **52**, 7182-7190 (1981).

35. M. F. Brown, Modulation of rhodopsin function by properties of the membrane bilayer. *Chem Phys Lipids* **73**, 159-180 (1994).

36. R. M. Epand, D. S. Lester, The role of membrane biophysical properties in the regulation of protein kinase C activity. *Trends Pharmacol Sci* **11**, 317-320 (1990).

37. P. V. Escribá *et al.*, Role of lipid polymorphism in G protein-membrane interactions: nonlamellar-prone phospholipids and peripheral protein binding to membranes. *Proc Natl Acad Sci U S A* **94**, 11375-11380 (1997).

38. J. A. Marsh, V. K. Singh, Z. Jia, J. D. Forman-Kay, Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Protein Sci* **15**, 2795-2804 (2006).

39. B. P. Chaudhary, D. Zoetewey, S. Mohanty, (1)H, (13)C, (15)N resonance assignments and secondary structure of yeast oligosaccharyltransferase subunit Ost4 and its functionally important mutant Ost4V23D. *Biomol NMR Assign*, (2020).

40. B. Chaudhary, S. Mazumder, S. Mohanty, Production and biophysical characterization of a mini-membrane protein, Ost4V23D: A functionally important mutant of yeast oligosaccharyltransferase subunit Ost4p. *Protein Expr Purif* **139**, 43-48 (2017).

41. L. Bai, T. Wang, G. Zhao, A. Kovach, H. Li, The atomic structure of a eukaryotic oligosaccharyltransferase complex. *Nature* **555**, 328-333 (2018).

42. R. Wild *et al.*, Structure of the yeast oligosaccharyltransferase complex gives insight into eukaryotic N-glycosylation. *Science* **359**, 545-550 (2018).

43. B. P. Chaudhary, D. L. Zoetewey, M. J. McCullagh, S. Mohanty, NMR and MD Simulations Reveal the Impact of the V23D Mutation on the Function of Yeast Oligosaccharyltransferase Subunit Ost4. *Glycobiology*, (2021).

44. S. Gayen, C. Kang, Solution structure of a human minimembrane protein Ost4, a subunit of the oligosaccharyltransferase complex. *Biochem Biophys Res Commun* **409**, 572-576 (2011).

45. S. Zubkov, W. J. Lennarz, S. Mohanty, Structural basis for the function of a minimembrane protein subunit of yeast oligosaccharyltransferase. *Proc Natl Acad Sci U S A* **101**, 3821-3826 (2004).

Appendix Tables

Appendix Table A-1: Chemical shift values of nuclei obtained from assignment by using solution NMR experiments of Ost4 protein

SN	Chemical Shift (ppm)	Chemical shift Deviation from average	Resonance	Residue Number	Residue
1	122.254	0.176	N	1	Met
2	8.917	0.008	H	1	Met
3	57.246	0.093	CA	1	Met
4	4.422	0.018	HA	1	Met
5	34.095	0.393	CB	1	Met
6	2.087	0.019	QB	1	Met
7	2.622	0.032	QG	1	Met
8	116.424	0.125	N	2	Ile
9	7.792	0.019	H	2	Ile
10	60.23	0.544	CA	2	Ile
11	4.373	0.018	HA	2	Ile
12	39.082	0.343	CB	2	Ile
13	1.869	0.016	HB	2	Ile
14	1.155	0.029	QG2	2	Ile
15	17.684	0	CG2	2	Ile
16	27.355	0	CG1	2	Ile
17	1.488	0.019	QG1	2	Ile
18	0.91	0.016	QD1	2	Ile
19	13.055	0	CD1	2	Ile
20	120.699	0.145	N	3	Ser
21	8.585	0.01	H	3	Ser
22	58.249	0.129	CA	3	Ser
23	4.517	0.021	HA	3	Ser

24	64.34	0.303	CB	3	Ser
25	4.206	0.008	HB2	3	Ser
26	3.911	0.03	HB3	3	Ser
27	122.16	0.114	N	4	Asp
28	8.607	0.008	H	4	Asp
29	56.838	0.158	CA	4	Asp
30	4.417	0.031	HA	4	Asp
31	41	0.224	CB	4	Asp
32	2.672	0.019	QB	4	Asp
33	120.055	0.138	N	5	Glu
34	8.648	0.014	H	5	Glu
35	59.2	0.088	CA	5	Glu
36	4.155	0.017	HA	5	Glu
37	29.543	0.037	CB	5	Glu
38	2.028	0.027	QB	5	Glu
39	2.341	0.022	QG	5	Glu
40	119.589	0.166	N	6	Gln
41	7.948	0.021	H	6	Gln
42	58.177	0.067	CA	6	Gln
43	4.177	0.021	HA	6	Gln
44	29.514	0.345	CB	6	Gln
45	2.044	0.027	QB	6	Gln
46	35.909	0.71	CG	6	Gln
47	2.28	0.027	QG	6	Gln
48	111.055	0.01	NE2	6	Gln
49	7.595	0.002	HE21	6	Gln
50	6.764	0	HE22	6	Gln
51	120.736	0.159	N	7	Leu
52	8.242	0.014	H	7	Leu
53	57.925	0.066	CA	7	Leu
54	4.082	0.016	HA	7	Leu
55	41.856	0.23	CB	7	Leu
56	1.749	0.023	QB	7	Leu
57	0.988	0.016	QQD	7	Leu
58	117.457	0.11	N	8	Asn
59	8.218	0.016	H	8	Asn
60	56.511	0.042	CA	8	Asn
61	4.435	0.02	HA	8	Asn
62	38.683	0.127	CB	8	Asn

63	2.941	0.012	HB2	8	Asn
64	2.758	0.032	HB3	8	Asn
65	114.425	0.124	N	9	Ser
66	7.979	0.016	H	9	Ser
67	61.829	0.118	CA	9	Ser
68	4.409	0.018	HA	9	Ser
69	65.043	1.252	CB	9	Ser
70	4.209	0.014	HB2	9	Ser
71	4.005	0.019	HB3	9	Ser
72	124.404	0.191	N	10	Leu
73	8.365	0.024	H	10	Leu
74	58.181	0.074	CA	10	Leu
75	4.088	0.026	HA	10	Leu
76	42.102	0.135	CB	10	Leu
77	1.874	0.025	QB	10	Leu
78	29.232	0	CG	10	Leu
79	1.665	0.015	HG	10	Leu
80	24.983	0	CD1	10	Leu
81	24.457	0	CD2	10	Leu
82	121.434	0.117	N	11	Ala
83	8.458	0.025	H	11	Ala
84	55.993	0.07	CA	11	Ala
85	4.029	0.026	HA	11	Ala
86	1.539	0.02	QB	11	Ala
87	19.042	0.461	CB	11	Ala
88	116.233	0.204	N	12	Ile
89	8.036	0.028	H	12	Ile
90	65.386	0.23	CA	12	Ile
91	3.711	0.019	HA	12	Ile
92	2.059	0.014	HB	12	Ile
93	17.596	0	CG2	12	Ile
94	1.95	0.029	QG1	12	Ile
95	0.988	0	QD1	12	Ile
96	61.741	0	CA	13	Thr
97	4.29	0.009	HA	13	Thr
98	68.351	0.036	CB	13	Thr
99	3.77	0.035	HB	13	Thr
100	1.131	0.01	QG2	13	Thr
101	21.573	0.034	CG2	13	Thr

102	120.212	0.131	N	14	Phe
103	8.788	0.012	H	14	Phe
104	61.747	0.162	CA	14	Phe
105	4.058	0.019	HA	14	Phe
106	39.165	0.085	CB	14	Phe
107	3.157	0.02	QB	14	Phe
108	106.494	0.15	N	15	Gly
109	8.622	0.019	H	15	Gly
110	47.834	0.068	CA	15	Gly
111	3.632	0.007	QA	15	Gly
112	121.267	0.15	N	16	Ile
113	8.407	0.025	H	16	Ile
114	65.404	0.09	CA	16	Ile
115	4.24	0.004	HA	16	Ile
116	38.815	0.56	CB	16	Ile
117	1.886	0.013	HB	16	Ile
118	1.211	0	QG2	16	Ile
119	1.522	0.005	QG1	16	Ile
120	0.84	0.025	QD1	16	Ile
121	122.705	0.163	N	17	Val
122	8.404	0.014	H	17	Val
123	67.832	0.128	CA	17	Val
124	3.423	0.011	HA	17	Val
125	32.349	0.75	CB	17	Val
126	2.111	0.025	HB	17	Val
127	21.265	0	CG1	17	Val
128	20.423	0	CG2	17	Val
129	0.98	0.019	QQG	17	Val
130	117.194	0.155	N	18	Met
131	8.464	0.015	H	18	Met
132	58.201	0.045	CA	18	Met
133	4.082	0.012	HA	18	Met
134	32.437	0.784	CB	18	Met
135	2.13	0.021	QB	18	Met
136	30.839	0	CG	18	Met
137	2.538	0.029	QG	18	Met
138	116.501	0.152	N	19	Met
139	8.352	0.019	H	19	Met
140	58.694	0	CA	19	Met

141	4.066	0.01	HA	19	Met
142	33.399	0.262	CB	19	Met
143	2.134	0.012	QB	19	Met
144	30.802	0	CG	19	Met
145	2.656	0	QG	19	Met
146	60.8	0	CA	20	Thr
147	4.308	0.021	HA	20	Thr
148	68.502	0.036	CB	20	Thr
149	3.657	0.026	HB	20	Thr
150	1.096	0.022	QG2	20	Thr
151	21.545	0	CG2	20	Thr
152	119.574	0.155	N	21	Leu
153	8.509	0.014	H	21	Leu
154	55.102	0.234	CA	21	Leu
155	3.963	0.009	HA	21	Leu
156	42.709	0.333	CB	21	Leu
157	1.647	0.007	QB	21	Leu
158	26.595	0	CG	21	Leu
159	24.271	0	CD1	21	Leu
160	0.948	0.023	QQD	21	Leu
161	115.874	0.161	N	22	Ile
162	8.105	0.02	H	22	Ile
163	60.959	0	CA	22	Ile
164	4.061	0.036	HA	22	Ile
165	40.207	0.998	CB	22	Ile
166	2.011	0.026	HB	22	Ile
167	1.054	0.032	QG2	22	Ile
168	17.036	0.021	CG2	22	Ile
169	26.611	0.122	CG1	22	Ile
170	1.552	0.022	HG12	22	Ile
171	1.23	0.016	HG13	22	Ile
172	0.807	0.02	QD1	22	Ile
173	13.662	0.202	CD1	22	Ile
174	118.53	0.174	N	23	Val
175	8.098	0.015	H	23	Val
176	67.112	0.712	CA	23	Val
177	3.695	0.018	HA	23	Val
178	31.269	0.17	CB	23	Val
179	2.387	0.011	HB	23	Val

180	0.999	0.018	QG1	23	Val
181	1.128	0	QG2	23	Val
182	21.132	0	CG1	23	Val
183	119.155	0.176	N	24	Ile
184	8.676	0.015	H	24	Ile
185	57.984	0.174	CA	24	Ile
186	3.991	0.029	HA	24	Ile
187	42.016	0.135	CB	24	Ile
188	1.94	0.036	HB	24	Ile
189	1.706	0.024	QG2	24	Ile
190	1.781	0.014	QG1	24	Ile
191	0.847	0.025	QD1	24	Ile
192	118.647	0.144	N	25	Tyr
193	8.244	0.022	H	25	Tyr
194	56.619	0.135	CA	25	Tyr
195	4.706	0.024	HA	25	Tyr
196	38.877	0.078	CB	25	Tyr
197	2.846	0.03	QB	25	Tyr
198	57.38	0.522	CA	26	His
199	4.126	0.023	HA	26	His
200	30.873	0.368	CB	26	His
201	3.184	0.017	HB2	26	His
202	3.125	0.029	HB3	26	His
203	122.039	0.18	N	27	Ala
204	8.547	0.018	H	27	Ala
205	54.573	0.441	CA	27	Ala
206	4.121	0.019	HA	27	Ala
207	1.511	0.021	QB	27	Ala
208	19.159	0.351	CB	27	Ala
209	116.636	0.088	N	28	Val
210	8.127	0.023	H	28	Val
211	61.61	0	CA	28	Val
212	3.809	0.015	HA	28	Val
213	32.654	0.554	CB	28	Val
214	2.174	0.023	HB	28	Val
215	0.943	0.02	QG1	28	Val
216	1.046	0.033	QG2	28	Val
217	20.592	0	CG1	28	Val
218	19.708	0	CG2	28	Val

219	121.21	0.163	N	29	Asp
220	8.255	0.014	H	29	Asp
221	55.928	0.057	CA	29	Asp
222	4.381	0.01	HA	29	Asp
223	42.051	0.112	CB	29	Asp
224	2.518	0.032	QB	29	Asp
225	113.228	0.173	N	30	Ser
226	8.074	0.018	H	30	Ser
227	60.327	0.061	CA	30	Ser
228	4.274	0.032	HA	30	Ser
229	63.642	0	CB	30	Ser
230	3.877	0.034	QB	30	Ser
231	114.056	0.136	N	31	Thr
232	7.936	0.014	H	31	Thr
233	64.123	0.506	CA	31	Thr
234	4.227	0.027	HA	31	Thr
235	69.43	0.036	CB	31	Thr
236	4.083	0	HB	31	Thr
237	1.268	0.012	QG2	31	Thr
238	21.958	0	CG2	31	Thr
239	119.704	0.23	N	32	Met
240	8.031	0.029	H	32	Met
241	55.889	0.028	CA	32	Met
242	4.439	0.018	HA	32	Met
243	33.838	0.663	CB	32	Met
244	2.103	0.019	QB	32	Met
245	2.575	0.028	QG	32	Met
246	116.057	0.254	N	33	Ser
247	7.928	0.014	H	33	Ser
248	56.468	0	CA	33	Ser
249	4.276	0.02	HA	33	Ser
250	63.423	0	CB	33	Ser
251	3.881	0.006	QB	33	Ser
252	63.878	0.217	CA	34	Pro
253	4.261	0.011	HA	34	Pro
254	31.886	0.002	CB	34	Pro
255	2.228	0.037	QB	34	Pro
256	120.067	0.325	N	35	Lys
257	8.282	0.02	H	35	Lys

258	57.557	0.188	CA	35	Lys
259	4.207	0.031	HA	35	Lys
260	33.288	0.543	CB	35	Lys
261	1.85	0.02	QB	35	Lys
262	24.798	0.06	CG	35	Lys
263	1.44	0.018	QG	35	Lys
264	29.157	0	CD	35	Lys
265	1.725	0.018	QD	35	Lys
266	42.28	0	CE	35	Lys
267	3.014	0.007	QE	35	Lys
268	118.062	0.247	N	36	Asn
269	8.179	0.018	H	36	Asn
270	54.344	0.535	CA	36	Asn
271	4.481	0.023	HA	36	Asn
272	38.838	0.07	CB	36	Asn
273	2.855	0.009	HB2	36	Asn
274	2.783	0.002	HB3	36	Asn
275	120.568	0.199	N	37	Arg
276	8.124	0.023	H	37	Arg
277	56.391	0.17	CA	37	Arg
278	4.338	0.014	HA	37	Arg
279	30.661	0.139	CB	37	Arg
280	1.86	0.016	HB2	37	Arg
281	1.939	0.005	HB3	37	Arg
282	1.767	0.015	HG2	37	Arg
283	1.645	0.017	HG3	37	Arg
284	123.261	0.204	N	38	Leu
285	8.248	0.013	H	38	Leu
286	55.323	0.141	CA	38	Leu
287	4.367	0.016	HA	38	Leu
288	42.407	0.227	CB	38	Leu
289	1.654	0.006	QB	38	Leu
290	27.094	0	CG	38	Leu
291	1.061	0	HG	38	Leu
292	24.861	0	CD1	38	Leu
293	23.699	0	CD2	38	Leu
294	0.918	0	QQD	38	Leu
295	125.863	0.128	N	39	Glu
296	7.729	0.012	H	39	Glu

297	57.929	0.123	CA	39	Glu
298	4.139	0.009	HA	39	Glu
299	31.309	0	CB	39	Glu
300	1.935	0.03	QB	39	Glu
301	2.179	0	QG	39	Glu
302	55.9	0	CA	44	His
303	30.09	0	CB	44	His
304	125.3	0.062	N	45	His
305	8.056	0.009	H	45	His
306	57.14	0	CA	45	His
307	4.395	0	HA	45	His
308	30.33	0	CB	45	His
309	3.061	0	QB	45	His

Appendix Table A-2: Chemical shift values of nuclei obtained from assignment by using solution NMR experiments of Ost4V23D protein

SN	Chemical Shift (ppm)	Chemical shift Deviation from average	Resonance	Residue Number	Residue
1	122.383	0.065	N	1	Met
2	8.972	0.013	H	1	Met
3	57.365	0.12	CA	1	Met
4	4.352	0.008	HA	1	Met
5	33.709	0.055	CB	1	Met
6	2.579	0.008	QB	1	Met
7	2.652	0.003	QG	1	Met
8	115.77	0.077	N	2	Ile
9	7.695	0.008	H	2	Ile
10	59.753	0.084	CA	2	Ile
11	4.447	0.007	HA	2	Ile
12	39.988	0.05	CB	2	Ile
13	1.829	0.006	HB	2	Ile
14	1.192	0.011	QG2	2	Ile
15	1.544	0.007	QG1	2	Ile
16	0.918	0.012	QD1	2	Ile

17	121.57	0.042	N	3	Ser
18	8.713	0.011	H	3	Ser
19	57.656	0.005	CA	3	Ser
20	4.335	0.013	HA	3	Ser
21	65.31	0.013	CB	3	Ser
22	3.979	0.009	QB	3	Ser
23	122.05	0.031	N	4	Asp
24	8.802	0.012	H	4	Asp
25	57.626	0.009	CA	4	Asp
26	4.523	0.019	HA	4	Asp
27	40.466	0.117	CB	4	Asp
28	2.647	0.021	QB	4	Asp
29	119.632	0.066	N	5	Glu
30	8.686	0.011	H	5	Glu
31	59.855	0.065	CA	5	Glu
32	4.095	0.007	HA	5	Glu
33	29.438	0.007	CB	5	Glu
34	2.052	0.022	QB	5	Glu
35	2.356	0.02	QG	5	Glu
36	120.224	0.035	N	6	Gln
37	7.878	0.009	H	6	Gln
38	58.719	0.027	CA	6	Gln
39	4.125	0.006	HA	6	Gln
40	28.98	0.041	CB	6	Gln
41	2.412	0	QB	6	Gln
42	3.063	0	QG	6	Gln
43	120.563	0.017	N	7	Leu
44	8.343	0.02	H	7	Leu
45	58.342	0.131	CA	7	Leu
46	4.032	0.015	HA	7	Leu
47	41.33	0.029	CB	7	Leu
48	1.743	0.014	QB	7	Leu
49	1.446	0	HG	7	Leu
50	0.899	0.007	QQD	7	Leu
51	117.175	0.064	N	8	Asn
52	8.227	0.012	H	8	Asn
53	56.489	0.012	CA	8	Asn
54	4.431	0.008	HA	8	Asn
55	38.799	0.05	CB	8	Asn

56	2.858	0.017	QB	8	Asn
57	114.43	0.035	N	9	Ser
58	7.915	0.01	H	9	Ser
59	61.742	0.033	CA	9	Ser
60	4.264	0.004	HA	9	Ser
61	62.921	0.003	CB	9	Ser
62	4.03	0.004	QB	9	Ser
63	123.137	0.026	N	10	Leu
64	8.293	0.01	H	10	Leu
65	58.332	0.166	CA	10	Leu
66	4.05	0.026	HA	10	Leu
67	41.918	0.058	CB	10	Leu
68	1.932	0.016	QB	10	Leu
69	1.587	0.004	HG	10	Leu
70	0.856	0	QQD	10	Leu
71	121.402	0.044	N	11	Ala
72	8.354	0.007	H	11	Ala
73	55.803	0.091	CA	11	Ala
74	4.027	0.017	HA	11	Ala
75	1.58	0.01	QB	11	Ala
76	18.382	0.02	CB	11	Ala
77	118.362	0.041	N	12	Ile
78	8.003	0.021	H	12	Ile
79	59.411	0	CA	12	Ile
80	4.011	0.016	HA	12	Ile
81	38.152	0.047	CB	12	Ile
82	2.063	0.016	HB	12	Ile
83	1.27	0.011	QG2	12	Ile
84	1.799	0.014	QG1	12	Ile
85	0.939	0.004	QD1	12	Ile
86	4.251	0.014	HA	13	Thr
87	67.518	0	CB	13	Thr
88	1.765	0.006	QG2	13	Thr
89	120.496	0.032	N	14	Phe
90	8.963	0.009	H	14	Phe
91	61.712	0.035	CA	14	Phe
92	3.169	0.011	HA	14	Phe
93	38.942	0.04	CB	14	Phe
94	2.696	0	QB	14	Phe

95	106.719	0.045	N	15	Gly
96	8.3	0.012	H	15	Gly
97	47.794	0.041	CA	15	Gly
98	4.005	0.011	HA2	15	Gly
99	3.712	0.006	HA3	15	Gly
100	122.078	0.06	N	16	Ile
101	7.932	0.008	H	16	Ile
102	64.792	0.111	CA	16	Ile
103	3.862	0.021	HA	16	Ile
104	2.072	0.017	HB	16	Ile
105	1.179	0.03	QG2	16	Ile
106	0.929	0.014	QD1	16	Ile
107	122.677	0.028	N	17	Val
108	8.281	0.009	H	17	Val
109	67.096	0.033	CA	17	Val
110	3.514	0.009	HA	17	Val
111	31.481	0.123	CB	17	Val
112	2.229	0.009	HB	17	Val
113	0.874	0.011	QG1	17	Val
114	1.038	0.017	QG2	17	Val
115	21.285	0.028	CG1	17	Val
116	117.751	0.079	N	18	Met
117	8.443	0.006	H	18	Met
118	58.451	0.028	CA	18	Met
119	4.091	0.027	HA	18	Met
120	31.873	0.032	CB	18	Met
121	1.962	0.02	QB	18	Met
122	32.111	0.029	CG	18	Met
123	2.286	0.013	QG	18	Met
124	116.21	0.065	N	19	Met
125	8.17	0.013	H	19	Met
126	58.413	0.037	CA	19	Met
127	4.181	0.011	HA	19	Met
128	32.031	0.079	CB	19	Met
129	2.199	0.011	QB	19	Met
130	2.699	0.013	QG	19	Met
131	116.639	0.047	N	20	Thr
132	7.943	0.008	H	20	Thr
133	4.282	0	HA	20	Thr

134	67.041	0	CB	20	Thr
135	4.382	0.009	HB	20	Thr
136	1.786	0.006	QG2	20	Thr
137	121.081	0.037	N	21	Leu
138	8.23	0.011	H	21	Leu
139	58.379	0.112	CA	21	Leu
140	4.023	0.015	HA	21	Leu
141	41.272	0.051	CB	21	Leu
142	1.427	0.01	QB	21	Leu
143	1.249	0.009	HG	21	Leu
144	0.838	0.016	QQD	21	Leu
145	119.199	0.052	N	22	Ile
146	8.153	0.011	H	22	Ile
147	65.745	0.04	CA	22	Ile
148	3.597	0.015	HA	22	Ile
149	37.809	0.077	CB	22	Ile
150	1.828	0.012	HB	22	Ile
151	0.905	0.016	QD1	22	Ile
152	120.548	0.043	N	23	Asp
153	7.853	0.016	H	23	Asp
154	57.799	0.002	CA	23	Asp
155	4.269	0.008	HA	23	Asp
156	40.525	0.124	CB	23	Asp
157	1.547	0	QB	23	Asp
158	121.19	0.095	N	24	Ile
159	8.294	0.012	H	24	Ile
160	65.267	0.002	CA	24	Ile
161	3.761	0.013	HA	24	Ile
162	37.918	0.17	CB	24	Ile
163	2.036	0.017	HB	24	Ile
164	1.268	0	QG2	24	Ile
165	1.569	0.014	QG1	24	Ile
166	0.883	0.011	QD1	24	Ile
167	120.678	0.04	N	25	Tyr
168	8.498	0.01	H	25	Tyr
169	61.694	0.084	CA	25	Tyr
170	4.131	0.018	HA	25	Tyr
171	38.487	0.299	CB	25	Tyr
172	3.216	0.023	QB	25	Tyr

173	115.11	0.077	N	26	His
174	8.357	0.015	H	26	His
175	58.438	0.351	CA	26	His
176	4.042	0.022	HA	26	His
177	28.994	0.025	CB	26	His
178	3.004	0.014	QB	26	His
179	122.845	0.093	N	27	Ala
180	8.033	0.017	H	27	Ala
181	54.208	0.026	CA	27	Ala
182	4.302	0.011	HA	27	Ala
183	1.551	0.012	QB	27	Ala
184	18.992	0.106	CB	27	Ala
185	117.825	0.048	N	28	Val
186	8.047	0.014	H	28	Val
187	64.676	0.029	CA	28	Val
188	4.417	0.014	HA	28	Val
189	31.849	0.169	CB	28	Val
190	2.136	0.008	HB	28	Val
191	1.007	0.02	QG1	28	Val
192	0.935	0.007	QG2	28	Val
193	22.25	0	CG1	28	Val
194	22.25	0	CG2	28	Val
195	121.444	0.079	N	29	Asp
196	8.284	0.016	H	29	Asp
197	56.05	0.127	CA	29	Asp
198	4.358	0.032	HA	29	Asp
199	41.846	0.006	CB	29	Asp
200	2.564	0.009	HB2	29	Asp
201	2.417	0.005	HB3	29	Asp
202	113.709	0.076	N	30	Ser
203	8.112	0.017	H	30	Ser
204	60.037	0.001	CA	30	Ser
205	3.946	0.021	HA	30	Ser
206	63.506	0.02	CB	30	Ser
207	4.312	0	QB	30	Ser
208	115.09	0.069	N	31	Thr
209	7.937	0.017	H	31	Thr
210	64.028	0.025	CA	31	Thr
211	4.248	0.024	HA	31	Thr

212	69.435	0.04	CB	31	Thr
213	2.659	0	QG2	31	Thr
214	119.973	0.097	N	32	Met
215	8.008	0.018	H	32	Met
216	55.925	0.025	CA	32	Met
217	4.444	0.005	HA	32	Met
218	33.062	0.024	CB	32	Met
219	2.583	0	QB	32	Met
220	116.167	0.104	N	33	Ser
221	7.948	0.01	H	33	Ser
222	56.46	0	CA	33	Ser
223	4.688	0.012	HA	33	Ser
224	63.477	0	CB	33	Ser
225	3.861	0.016	QB	33	Ser
226	63.587	0.166	CA	34	Pro
227	4.427	0.013	HA	34	Pro
228	31.943	0	CB	34	Pro
229	2.042	0.017	QB	34	Pro
230	2.29	0	QG	34	Pro
231	119.959	0.076	N	35	Lys
232	8.274	0.014	H	35	Lys
233	57.033	0.01	CA	35	Lys
234	4.194	0.008	HA	35	Lys
235	32.843	0.088	CB	35	Lys
236	1.762	0.002	QB	35	Lys
237	118.318	0.058	N	36	Asn
238	8.176	0.01	H	36	Asn
239	53.566	0.012	CA	36	Asn
240	38.881	0.005	CB	36	Asn
241	2.784	0	QB	36	Asn
242	120.72	0.066	N	37	Arg
243	8.137	0.02	H	37	Arg
244	56.489	0.01	CA	37	Arg
245	4.251	0	HA	37	Arg
246	30.7	0.001	CB	37	Arg
247	1.868	0	QB	37	Arg
248	1.577	0	QG	37	Arg
249	3.154	0	QD	37	Arg
250	121.981	0.058	N	38	Leu

251	8.12	0.009	H	38	Leu
252	55.324	0.004	CA	38	Leu
253	4.27	0.006	HA	38	Leu
254	42.252	0.047	CB	38	Leu
255	1.57	0	QB	38	Leu
256	1.599	0.002	HG	38	Leu
257	0.865	0.01	QQD	38	Leu
258	120.743	0.094	N	39	Glu
259	8.155	0.022	H	39	Glu
260	56.513	0	CA	39	Glu
261	4.046	0.028	HA	39	Glu
262	30.598	0	CB	39	Glu
263	1.898	0.017	QB	39	Glu
264	55.91	0	CA	44	His
265	30.13	0	CB	44	His
266	3.091	0	QB	44	His
267	125.4	0.067	N	45	His
268	8.012	0.013	H	45	His
269	57.26	0	CA	45	His
270	4.395	0	HA	45	His
271	30.48	0	CB	45	His

VITA

Bharat Prasad Chaudhary

Candidate for the Degree of

Doctor of Philosophy

Thesis: EXPRESSION, PURIFICATION, NMR STRUCTURE, AND MOLECULAR DYNAMICS STUDIES OF Ost4 AND Ost4V23D: A CRITICAL SUBUNIT OF YEAST OLIGOSACCHARYLTRANSFERASE

Major Field: Chemistry

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in chemistry at Oklahoma State University, Stillwater, Oklahoma in May, 2021.

Completed the requirements for the Master of Science in chemistry at Tribhuwan University, Kirtipur, Nepal in 2007.

Completed the requirements for the Bachelor of Science in chemistry at Butwal Multiple Campus, Butwal, Nepal in 2005.

Experience:

- QC officer at Siddhartha Pharmaceuticals Pvt. Ltd, 2008-2009
- Chemistry Lecturer at Khwapa College, Bhaktapur, Nepal 2009-2010
- Chemistry Lecturer at SS College, Nayathimi, Nepal 2010-2014
- Graduate Teaching and Research Assistant, Oklahoma State University 2015-2020

Awards: Outstanding graduated student award (2020)