

**Residue level characterization of molecular
interactions of intermembrane space domains
governing the preprotein import into the
mitochondrial matrix**

Dissertation

for the award of the degree

‘Doctor rerum naturalium’ (Dr. rer. nat.)

Division of Mathematics and Natural Sciences

of Georg-August-Universität, Göttingen

Submitted by

Rakhi Bajaj

From New Delhi, India

Göttingen, 2013

Members of the Thesis Committee:

Prof. Dr. Markus Zweckstetter, (1st reviewer)
Department of NMR based Structural Biology,
Max Planck Institute for Biophysical Chemistry
Göttingen

Prof. Dr. Peter Rehling, (2nd reviewer)
Göttingen University Medical School
Dept. Biochemistry II,
Georg-August-University, Göttingen

Prof. Dr. Kai Tittmann,
Department of Bio-Analytics,
Georg-August-University, Göttingen

Date of the oral examination: 1st March, 2013

Affidavit

I hereby declare that this dissertation has been written independently and with no other sources and aids than quoted.

.....

Rakhi Bajaj

Göttingen

ABSTRACT

Mitochondrial matrix targeting proteins are translated as preproteins (carrying an N-terminal 20-50 residue presequence) in the cytoplasm. Post-translationally, they are imported into the mitochondrial matrix through multi-subunit protein machineries called translocases. The intermembrane space domains (ims) of both the outer and inner mitochondrial membrane translocases perform multiple functions including presequence-receptor binding, translocation contact site constituent and regulation of channel activity across the inner mitochondrial membrane.

TIM23 is the translocase of the inner mitochondrial membrane comprising of Tim17, Tim21, Tim23, Tim50 and motor subunits in *S.cerevisiae*. Here, we have characterized the intermembrane space domains (ims) of Tim23 and Tim21 in solution. We show that Tim21ims is a folded protein and exhibits monomer-dimer equilibrium in solution. The monomeric Tim21ims undergoes dynamics in β -sheet that might play important role in its dimerization.

Tim23ims is intrinsically disordered and exists as monomer in solution and is engrossed in a multitude of interactions in the intermembrane space of mitochondria to facilitate import of matrix targeted preproteins. Using solution NMR spectroscopy, the atomic details of binding sites of disordered Tim23 with functionally important ligands (Tim21, Tom22, Tom40, Tim50, mitochondrial membrane mimics and presequence) are elucidated. We have identified five hydrophobic linear motifs in Tim23 involved in binding the aforementioned ligands. We also demonstrate that in disordered Tim23, residues 58-78 acts as hub region and its interactions with aforesaid ligands is regulated by well-placed multivalent hydrophobic motifs, which clearly provide the molecular basis for working of exuberant Tim23 in mitochondria.

The role and structural characterization of Tim23ims in two different complexes, one involving membrane mimics (Tim23-DHPC micelles) and another with Tim21, corroborate the lack of regular secondary structure in its bound form. We propose a model for the weak interaction of Tim21-Tim23 intermembrane space domain that describes the binding of short linear hydrophobic motifs of Tim23ims at single site in Tim21ims. Our interaction studies underline the importance of interplay of hydrophobic linear motifs in providing specificity and explicit affinity in interactions involving disordered proteins. This study establishes the need of multiple hydrophobic binding motifs of disordered domains to interact in synergistic manner in transiently weak complex of Tim21-Tim23 and highlights the precise selectivity of hydrophobic motifs as in Tim23-DHPC complex.

Acknowledgements

It is a great pleasure to express my sincere thanks to all those who made this thesis possible.

I am grateful to Prof. Dr. Markus Zweckstetter for providing me a wonderful opportunity to work on this challenging project. He has enlightened me in the field of Biomolecular NMR spectroscopy. I express my sincere gratitude for his constant support, suggestions, encouragement and guidance throughout my thesis. I am grateful to him for giving me enough time for learning the available experimental techniques that laid the foundation for my future career in research.

I deeply thank Dr. Stefan Becker for his guidance and discussions in the molecular biology part of the thesis. Without his support and lab facilities, the large part of the thesis could not be accomplished. I also thank Karin Giller for her constant support in the wet lab.

I extend my sincere thanks to Prof. Dr. Christian Griesinger for providing the excellent research infrastructure and NMR spectrometer facilities in our department.

I am thankful to my thesis committee advisors Prof. Dr. Peter Rehling and Prof. Dr. Kai Tittmann for useful discussions and suggestions during thesis committee meetings.

I am grateful to Laura de la Cruz for her initial studies on Tim23ims. I would like to thank Dr. Lukasz Jaremko and Dr. Mariusz Jaremko for the useful discussions and cooperation in Tim21ims project.

I am thankful to Dr. Dana Hutu and Dr. Oleksandr Lytovchenko from Prof. Dr. Peter Rehling group for their inputs on this project.

I thank Dr. Ulrich Dürr and Dr. Lukasz Skora for introducing to me to the NMR data acquisition and analysis during the first year of this thesis. I am also grateful to Dr. Saskia Villinger and Sheng Qi Xiang for their discussions on NMR theory and constantly helping me in troubleshooting NMR spectrometer even at odd hours. Special thanks are due for Dr. Saskia for proof-reading thesis.

I thank Dr. Dirk Bockelmann and Jürgen Arve for maintaining computer network administration and programs. I also extend my thanks to Frederik Klama for

introducing me to the GWDG computer resources and help in troubleshooting programming scripts.

I am thankful to all members of NMR-II for cordial and scientific environment in department. I thank Jens, Guowei, JP, Vinesh, Ashutosh, Donghan, Hai Young, Min-Kyu, Vineta, Edward, Stefan, Korvin, Hari, Raghav, Yao, Aldo, Martin, Francesca, Hessam, Luis, Nilomani, Suresh, Yvonne, Sebastian and Claudia for their constant help and support during the course of this thesis.

I thank GGNB and Max Planck Stipend for providing financial support for my studies in Germany.

I sincerely express my gratitude towards my teachers who have constantly encouraged me to work hard and motivated me to achieve this goal. I thank my parents for their constant support and their faith in my abilities and my brother to share the load of my responsibilities so that I can successfully accomplish this thesis.

Table of Contents

| | |
|---|----|
| ABBREVIATIONS | IX |
| 1 INTRODUCTION..... | 11 |
| 1.1 MITOCHONDRIA | 11 |
| 1.2 MITOCHONDRIAL PREPROTEIN TRANSLOCATION | 12 |
| 1.3 PRESEQUENCE TRANSLOCASE PATHWAY | 17 |
| 1.3.1 Presequence and its receptor sites along the import pathway | 17 |
| 1.3.2 Functions and components of the presequence translocase | 19 |
| 1.3.3 An overview for transport of preprotein using the presequence translocase..... | 27 |
| 1.3.4 Dynamic changes and multiple interactions drives preprotein translocation | 28 |
| 1.4 INTERACTION NETWORK OF TIM23..... | 32 |
| 1.5 INTERMEMBRANE SPACE DOMAINS (IMS) OF TWO MITOCHONDRIAL TRANSLOCASES..... | 33 |
| 1.6 INTRINSICALLY DISORDERED PROTEINS (IDPs) AS HUB SITES IN INTERACTION NETWORKS..... | 35 |
| 1.7 INTRINSICALLY DISORDERED PROTEINS AND NMR SPECTROSCOPY..... | 36 |
| 1.8 FUZZY COMPLEXES..... | 39 |
| 1.9 AIMS AND OUTLINE OF THESIS..... | 41 |
| 2 MATERIALS AND METHODS | 45 |
| 2.1 MATERIALS | 45 |
| 2.1.1 Chemical reagents | 45 |
| 2.1.2 Bacterial strains..... | 46 |
| 2.1.3 Expression vectors for purification intermembrane space domains | 47 |
| 2.1.4 Growth media..... | 48 |
| 2.1.5 Stock solutions of antibiotics | 49 |
| 2.1.6 Scientific Software's | 52 |
| 2.2 MOLECULAR BIOLOGY METHODS | 53 |
| 2.2.1 Constructs design for intermembrane space domains | 53 |
| 2.2.2 Cloning of various intermembrane space domains: | 53 |
| 2.2.3 Site directed mutagenesis..... | 54 |
| 2.2.4 Expression of various intermembrane space domains | 54 |
| 2.2.5 Purification method for intermembrane space domains cloned into pET28a-His7-Z ₂ -TEV | 55 |
| 2.3 MITOCHONDRIAL LIPOSOME PREPARATIONS | 57 |
| 2.4 SPIN LABELLING WITH THE PARAMAGNETIC REAGENT | 57 |

| | | |
|-------|--|-----|
| 2.4.1 | MTSL..... | 57 |
| 2.4.2 | Cys-Ph-TAHA, a lanthanide tag | 58 |
| 2.5 | NMR SPECTROSCOPY | 58 |
| 2.5.1 | NMR assignment of intermembrane space domains | 58 |
| 2.5.2 | Secondary chemical shift analysis of Tim23..... | 59 |
| 2.5.3 | Estimation of T_2 and the global rotational correlation time | 60 |
| 2.5.4 | Steady state heteronuclear, ^{15}N -nuclear Overhauser effects (NOEs)..... | 60 |
| 2.5.5 | ^{15}N relaxation measurements..... | 60 |
| 2.5.6 | RDCs..... | 61 |
| 2.5.7 | ^{15}N - ^1H -HSQC based characterization of protein ligand interactions | 61 |
| 2.5.8 | Intermolecular paramagnetic relaxation enhancement (PRE) | 62 |
| 2.6 | ERROR ANALYSIS BASED ON S/N RATIO | 63 |
| 2.7 | PROTEIN PEPTIDE DOCKING USING FLEXPEPDOCK-ROSETTA 3.2 | 63 |
| 2.8 | DISTANCE RESTRAINTS AND STRUCTURE CALCULATION | 64 |
| 2.9 | OTHER BIOPHYSICAL TECHNIQUES | 64 |
| 2.9.1 | UV spectroscopy for determination of protein concentration | 64 |
| 2.9.2 | Circular dichorism (CD) spectroscopy | 65 |
| 2.9.3 | Size exclusion chromatography..... | 65 |
| 3 | RESULTS..... | 67 |
| 3.1 | INTERMEMBRANE SPACE DOMAINS (IMS) OF MITOCHONDRIAL TRANSLOCASES | 67 |
| 3.1.1 | Purification of intermembrane space domains | 67 |
| 3.1.2 | Structural characterization of IMS domains | 68 |
| 3.1.3 | Tim23ims as intrinsically disordered monomer | 70 |
| 3.1.4 | Solution properties of Tim21ims | 75 |
| 3.1.5 | Tim50 ¹⁶⁴⁻³⁶¹ ims is folded in solution | 83 |
| 3.1.6 | Tom22ims is largely unstructured in solution | 84 |
| 3.2 | PRESEQUENCE BINDING SITES IN MITOCHONDRIAL INTERMEMBRANE SPACE DOMAINS | 86 |
| 3.2.1 | Tim23ims interacts presequence through its C-terminus | 86 |
| 3.2.2 | Dynamics in Tim23ims-Presequence complex | 89 |
| 3.2.3 | Effect of key mutations in Tim23ims-presequence complex | 89 |
| 3.2.4 | Tom22ims interacts weakly with presequence | 91 |
| 3.2.5 | Tim21ims lacks presequence binding site | 92 |
| 3.2.6 | Mapping presequence binding site on Tim50ims164-361..... | 93 |
| 3.3 | INTERACTIONS AT THE TRANSLOCATION CONTACT SITE | 95 |
| 3.3.1 | Tom22ims does not bind Tim21ims..... | 95 |
| 3.3.2 | Tim23ims weakly interacts with Tom22ims | 98 |
| 3.3.3 | Tim23ims binds to Tom40ims | 100 |
| 3.4 | INTERACTION STUDIES OF TIM23 WITH MITOCHONDRIAL MEMBRANE MIMICS..... | 101 |

| | | |
|-------|---|-----|
| 3.4.1 | Tim23 interacts with mitochondrial membrane mimicking liposomes | 101 |
| 3.4.2 | Interaction of Tim23 to micelles (DHPC)..... | 105 |
| 3.4.3 | Presequence binding site is accessible in liposome bound Tim23ims | 111 |
| 3.5 | STRUCTURAL PROPERTIES OF INTERMEMBRANE SPACE DOMAINS OF TIM21-TIM23 | 114 |
| 3.5.1 | Localization of Binding sites of Tim21ims with Tim23ims | 114 |
| 3.5.2 | Linear motifs of Tim23ims involved in binding to Tim21ims..... | 122 |
| 3.5.3 | Model for Tim21-Tim23 complex | 126 |
| 3.6 | MAPPING TIM50IMS INTERACTION SITES ON TIM23IMS..... | 130 |
| 4 | DISCUSSION | 133 |
| 4.1 | INTERMEMBRANE SPACE DOMAINS OF MITOCHONDRIAL TRANSLOCASES ARE HIGHLY DYNAMIC WITH DIVERSE STRUCTURAL PROPERTIES..... | 133 |
| 4.1.1 | Tom22ims is largely disordered and possess a transient helix in solution | 134 |
| 4.1.2 | Tim21ims exhibits monomer-dimer equilibrium in solution and monomeric Tim21 ims is highly dynamic in solution | 135 |
| 4.1.3 | Tim50 ¹⁶⁴⁻³⁶¹ ims is well folded with distinct dynamics | 136 |
| 4.1.1 | Tim23ims is intrinsically disordered in solution | 136 |
| 4.2 | PRESEQUENCE RECOGNITION AND TRANSLOCATION THROUGH INTERMEMBRANE SPACE FOR MITOCHONDRIAL MATRIX-TARGETED PREPROTEIN | 138 |
| 4.2.1 | Presequence binding sites in intermembrane space domains of mitochondrial translocase | 138 |
| 4.2.2 | Presequence binding and preprotein import..... | 142 |
| 4.3 | INTERACTIONS AT THE TRANSLOCATION CONTACT SITE..... | 143 |
| 4.4 | TIM23IMS ACTS AS A HUB PROTEIN IN INTERMEMBRANE SPACE | 146 |
| 4.4.1 | Conserved linear motifs of Tim23ims in yeast | 146 |
| 4.4.2 | Tim23ims interact with the mitochondrial membrane at its N-terminus via two binding motifs. 148 | |
| 4.4.3 | Fuzzy complex involving Tim21ims-Tim23ims..... | 150 |
| 4.4.4 | Tim50ims interacts with Tim23ims using multiple interaction motifs | 153 |
| 4.5 | CONSERVED LINEAR MOTIFS OF TIM23IMS ACTIVELY MEDIATE ITS INTERACTION WITH TOM22, TOM40IMS, TIM21IMS, TIM50IMS AND PRESEQUENCE | 154 |
| 4.6 | MULTIVALENCY AND TUNED HYDROPHOBICITY GOVERNS THE WORKING OF TIM23 | 156 |
| 5 | SUMMARY AND OUTLOOK | 158 |
| 5.1 | INTERMEMBRANE SPACE DOMAINS OF MITOCHONDRIAL TRANSLOCASES | 158 |
| 5.2 | PRESEQUENCE-INTERMEMBRANE SPACE DOMAIN INTERACTIONS..... | 158 |
| 5.3 | INTERACTIONS BETWEEN IMS DOMAINS OF THE MITOCHONDRIAL TRANSLOCASES..... | 159 |
| 6 | REFERENCES..... | 163 |

| | | |
|-------|---|-----|
| 7 | APPENDIX | 175 |
| 7.1 | PURIFICATION OF INTERMEMBRANE SPACE DOMAINS | 175 |
| 7.1.1 | Tom22ims | 175 |
| 7.1.2 | Tim21ims | 176 |
| 7.1.3 | Tim50ims (164-361): | 178 |
| 7.2 | TIM21 | 179 |
| 7.2.1 | Effect of pH on backbone resonances of Tim21ims..... | 179 |
| 7.2.2 | Chemical shift changes in observed in backbone resonances of Tim21ims upon addition of different constructs of Tim23ims..... | 181 |
| 7.2.3 | Sequence alignment of Tim23 binding site in Tim21ims..... | 183 |
| 7.3 | PRIMARY SEQUENCE OF PRESEQUENCE RALDH AND ITS VARIANTS USED IN SECTION 3.2.3..... | 184 |
| 7.4 | PRIMARY SEQUENCE AND SECONDARY STRUCTURE PREDICTION OF TOM22IMS | 184 |
| 7.5 | PRIMARY SEQUENCE OF PRESEQUENCE TOM40 IMS | 185 |
| | CURRICULUM VITAE | 175 |

Abbreviations

TOM- Translocase of Outer mitochondrial Membrane

TIM- Translocases of Inner mitochondrial Membrane

IMS- Inner Membrane Space

OMM- Outer Mitochondrial Membrane

IMM- Inner Mitochondrial Membrane

NMR– Nuclear Magnetic Resonance

MORF- Molecular Recognition Fragments

RDC- Residual Dipolar Coupling

PRE- Paramagnetic Relaxation Enhancement

NOE- Nuclear Overhauser Effect

CSP- Chemical Shift Perturbation

IDP- Intrinsically Disordered Proteins

APSY- Automated Projection Spectroscopy

NOESY- Nuclear Overhauser Effect Spectroscopy

DHPC- 1, 2-diheptanoyl-*sn*-glycero-3-phosphocholine

CD- Circular Dichroism

DOSY- Diffusion Ordered Spectroscopy

TROSY- Transverse Relaxation Optimized Spectroscopy

ATP- Adenosine tri phosphate

SDS-PAGE- Sodium Dodecyl Sulfate-PolyAcrylamide Gel Electrophoresis

Ni-NTA- Nitrilo Triacetic Acid

IPTG- Isopropyl β -D-1-thiogalactopyranoside

GST- glutathione S-transferase

TEV- Tobacco Etch Virus

DTT- Dithiothreitol

PMSF- Phenylmethylsulfonyl fluoride

EDTA- Ethylenediaminetetraacetic acid

E.coli– Escherichia coli

IMAC- Immobilized Metal Affinity Chromatography

RP-HPLC- Reverse phased High-Performance Liquid Chromatography

ESI-MS Electrospray Ionization Mass Spectrometry

HSQC- Heteronuclear Single Quantum Coherence Spectroscopy

CSI- Chemical Shift Index

RMSD- Root Mean Square Deviation

CMC- Critical Micellar Concentration

SEC - Size Exclusion Chromatography

rALDH- retinal Aldehyde Dehydrogenase

1 Introduction

1.1 Mitochondria

Mitochondria were discovered by Richard Altmann in 1894 as “bio-blasts”. The term “mitochondria” were coined later by Carl Benda in 1898. Mitochondria are double membrane organelles that vary in number and size among various eukaryotic cells. Most commonly, mitochondria are known as the “powerhouse of the cell”. In addition to the well-known function of ATP production, they are also involved in a spectrum of other vital cellular functions including calcium regulation, metabolism of carbohydrates and fats, heme and steroid synthesis, regulation of membrane potential and apoptosis (Attardi and Schatz 1988; Mignotte and Vayssiere 1998; McBride et al. 2006).

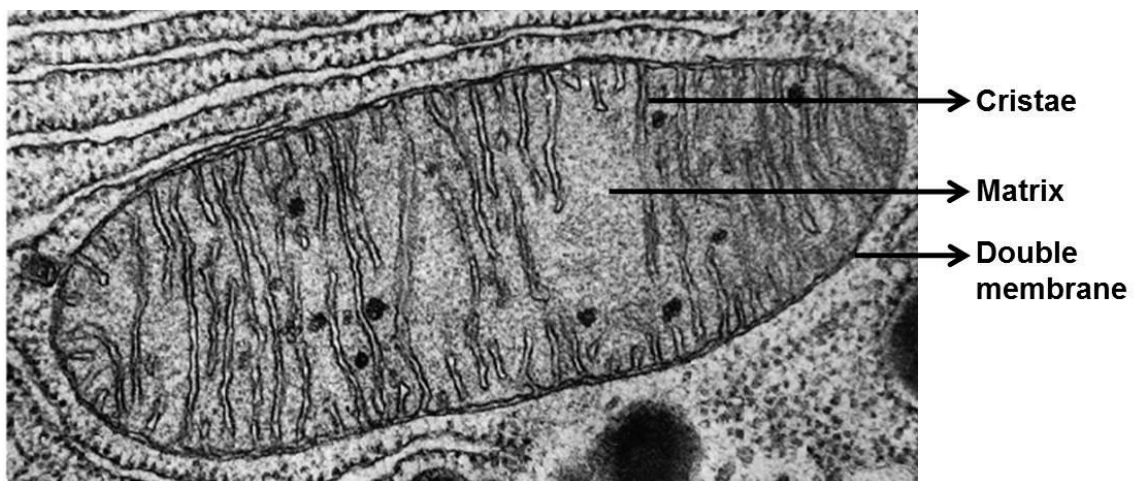


Figure 1: An electron micrograph of a mitochondrion highlighting its cellular compartments: Double membrane comprised of outer and inner membrane, cristae and mitochondrial matrix. The image is obtained from the Don W. Fawcett's 1981 atlas of electron micrographs available at <http://bioeducate.ascb.org/FawcettTheCell.html>

Topologically, mitochondria are composed of four well-defined compartments namely 1) the outer membrane 2) the inner membrane 3) the intermembrane space and 4) the mitochondrial matrix. One of the distinguishable features of mitochondria is the structure and lipid-protein composition of its inner membrane. Morphologically, the inner membrane invaginates into the matrix to form cristae (Figure 1). Compositionally, the inner mitochondrial membrane is specifically rich in proteins involved in the electron transport chain and the negatively charged lipid called cardiolipin.

1.2 Mitochondrial preprotein translocation

Mitochondria are semi-autonomous organelles carrying their own circular DNA and translational machinery. However, approximately 99% of yeast mitochondrial proteins are encoded by nuclear DNA and post-translationally they are translocated to their functional site in mitochondria (Van Der Laan et al. 2006). The translocation of proteins into mitochondria has been a major subject of interest for the past 30 years and is not yet fully explored as evident from the fact that a new component associated with the inner mitochondrial translocase has been discovered recently (Gebert et al. 2012) .

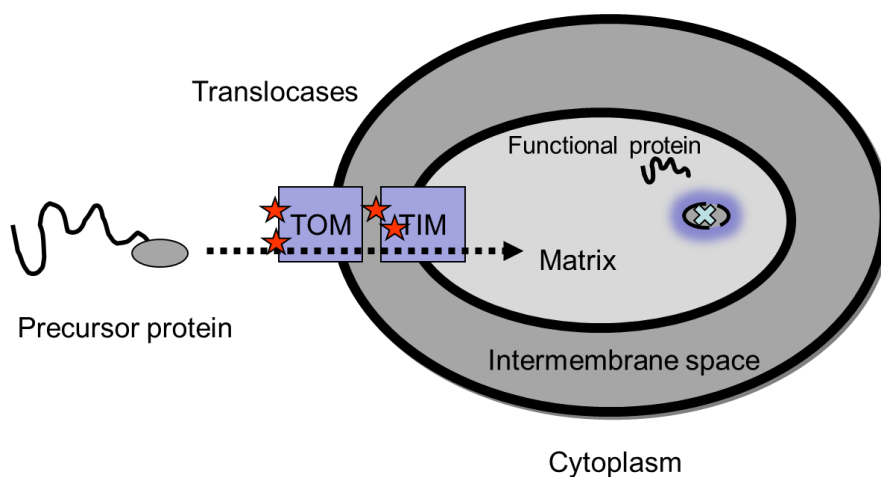


Figure 2: Schematic representation of a general mechanism for translocation of mitochondrial proteins is indicated by dashed arrows. Post-translationally, the translocation tag (represented by a gray oval) carrying preprotein interacts with the receptor sites (red stars) along the translocases (blue boxes) to reach the destined location (as matrix in this case) where the tag (usually an N-terminal presequence) is removed to yield the functional protein.

The most general mechanism of protein translocation into mitochondria includes the synthesis of a precursor protein that carries a translocation tag (also known as presequence) using the nuclear-cytoplasmic translational machinery. This is followed by binding of the preprotein to the surface receptors on the outer mitochondrial membrane translocase, (TOM) and the preprotein is subsequently imported to the various compartments in mitochondria. The machineries involved in preprotein translocation are called Translocases (TOM, Translocase of Outer mitochondrial Membrane, and TIM Translocases of Inner mitochondrial Membrane), which are hetero-oligomeric assemblies of various protein subunits. Various subunits of the TOM and TIM complex form the receptor sites and pore across the membrane to guide the translocation of incoming mitochondrial preprotein to their destined functional site in the mitochondrion. The final processing of preprotein involves the removal of the translocation tag (also known as presequence) to yield mature functional mitochondrial protein.

The quest to understand the molecular mechanism involved in translocation and the emergence of the mitochondrial proteome enlightened the field of protein translocation in mitochondria. Currently, there are five major protein import pathways in yeast that describe the mechanism of translocation of the mitochondrial preproteins to various sites in mitochondria (Becker et al. 2012). Two of these pathways are dedicated for translocating the membrane proteins to the outer membrane depending on their secondary content as α -helical or β -barrel (α -helical and β -barrel pathway) while two other pathways are involved in sorting inner membrane proteins depending on their translocation tag (with or without cleavable presequence called TIM23 and TIM22, respectively). The last pathway MIA (Mitochondrial Intermembrane space Import and Assembly) pathway is exclusively dedicated to import of the cysteine rich proteins into the intermembrane space. On the other hand, the TIM23 pathway shares the load of translocating the presequence carrying matrix preproteins either into the mitochondrial matrix or to the inner mitochondrial membrane.

Figure 3, Figure 4 and Table 1 highlight the essential features of these five pathways and exemplify their functional mechanisms and component machinery.

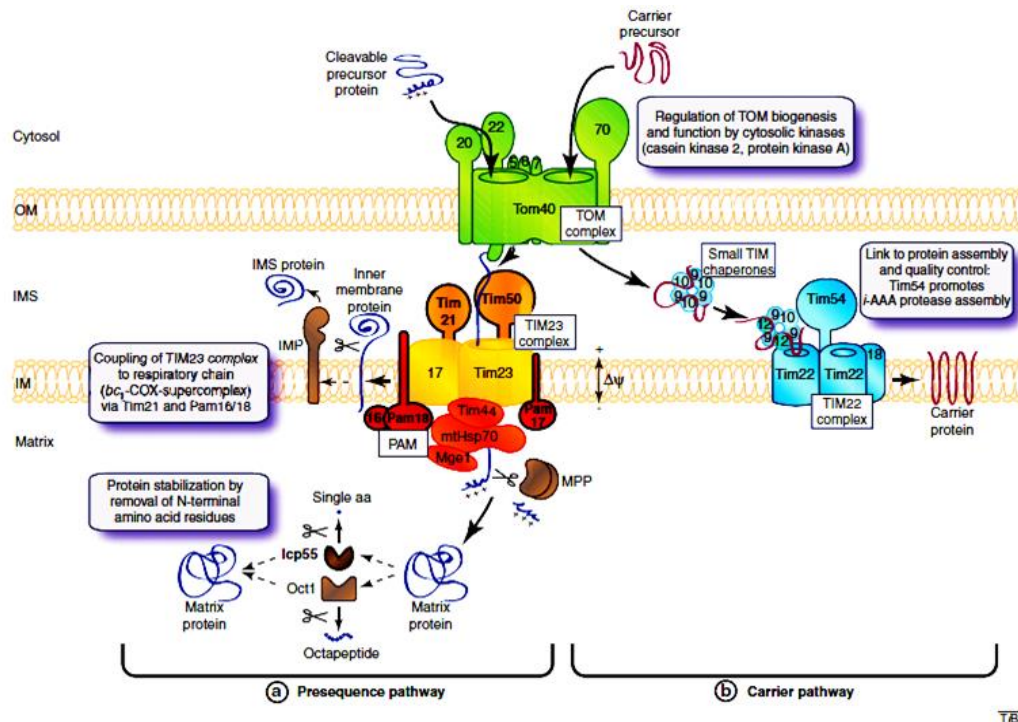


Figure 3: The two pathways for mitochondrial preprotein import. (a) Presequence translocase pathway directs an N-terminal presequence carrying preprotein to the mitochondrial matrix. The cytosolic parts of the Tom22 and Tom20 acts as the first receptors site at the outer mitochondrial membrane and guide the incoming preprotein to the channel across the outer membrane formed by Tom40. Intermembrane space domain (ims) of subunits of the TIM23 complex mainly Tim50 and Tim23 further acts as the receptor for the preprotein to guide it through the inner membrane channel formed by Tim23 and Tim17 in the presence of membrane potential. The translocation of the incoming preprotein into the matrix is accomplished by the motor part of TIM23 and mtHsp70. In the matrix, the imported preproteins are firstly processed by MPP (mitochondrial processing peptidase) that removes the presequence and can be further stabilized with the removal of certain destabilizing residues by special peptidases in the matrix such as Icp 55 or Oct 1. If the incoming preprotein contains the hydrophobic sorting signal, then it is inserted into the inner membrane either by lateral release using the TIM23 complex or can first be imported into the matrix and then can be re-inserted into the inner membrane using another Oxa1 import complex (not shown in this figure). (b) The second pathway shown here is called the Carrier Pathway, which imports the non-cleavable precursor proteins carrying the internal targeting signals. Tom70 acts as receptor for these proteins, and they are also translocated across the outer membrane using the channel formed by Tom40. The small chaperones (Tim9-Tim10 complex) present in the intermembrane space domain guides these hydrophobic precursors to the TIM22 complex that drives membrane potential dependent insertion of these hydrophobic precursors into the inner mitochondrial membrane. The transport pathways in mitochondria are not simple rather complex at multiple levels with the outer membrane translocase subunits being regulated by a number of cytosolic kinases. The inner membrane TIM23 complex (Tim21) cooperates with the respirator chain complexes to sort the preproteins to the inner membrane and Tim54 of the TIM22 complex interacts with i-AAA protease to maintain the

mitochondrial turnover. Reprinted from (Becker et al. 2012), Copyright(2012),with permission from Elsevier.

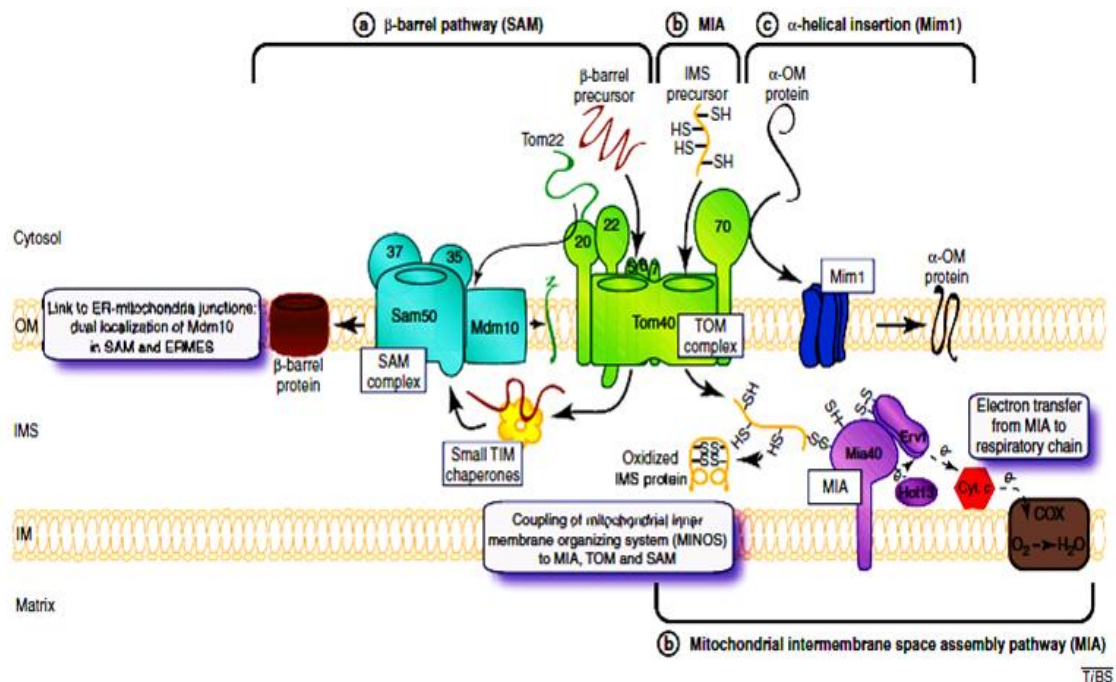


Figure 4: Recently discovered preprotein import pathways in mitochondria. (a) The β -barrel pathway directs the imports the precursors of β -barrel proteins to the outer mitochondrial membrane. The precursors transverse the outer membrane (OM) with the help of the TOM complex, followed by interaction with the small TIM chaperone complexes of the intermembrane space (IMS), which transfer these precursors to the Sorting and Assembly machinery (SAM) complex that mediates insertion of these proteins into the outer membrane. (b) The mitochondrial intermembrane space assembly pathway (MIA) imports the cysteine-rich IMS proteins through TOM complex. After emerging out from the TOM complex, the precursor protein interacts with Mia40 via disulfide bonds. Mia40 catalyzes the oxidative folding of these cysteine rich proteins. The electrons flow from Mia40 via Erv1 to cytochrome c and finally to the cytochrome c oxidase (COX) according to the redox potential gradient to assist in substrate folding. (c) α -Helical insertion (Mim1) pathway imports the outer membrane proteins containing multiple or single α -helical transmembrane segments, which are recognized by Tom70 and transferred to the Mim1 complex for membrane insertion. The above listed three pathways are also complex and links protein import and mitochondrial membrane morphology via coupling of Mia40, TOM and SAM. Mim1 can transiently interact with the SAM complex to assist the assembly of small (α -helical) Tom proteins with Tom40. Moreover, Mdm10 is found to be part of both the complexes, SAM and EPIMES (ER-mitochondria encounter structures) and thus links mitochondrial protein biogenesis to ER-mitochondria junctions. Reprinted from (Becker et al. 2012), Copyright(2012), with permission from Elsevier.

Table 1: Mitochondrial protein import pathways describing its components and site of translocation of preprotein in yeast.

| Pathway | Protein components | Functional destination | Comments |
|--|---------------------------------------|------------------------------|--|
| Beta barrel pathway | TOM40, SAM, small TIM chaperones | Outer mitochondrial membrane | Translocate β barrel membrane proteins |
| Alpha helical pathway | TOM40, SAM, Mdm10, Mim1 | Outer mitochondrial membrane | Translocate α helical membrane proteins |
| MIA pathway (Mitochondrial Intermembrane space Assembly pathway) | TOM40, Mia40, Erv1 | Intermembrane space | Cysteine rich intermembrane space proteins are translocated |
| Presequence pathway (TIM 23 pathway) | TOM40, TIM23, motor PAM, mtHsp70, MPP | Mitochondrial matrix | Membrane potential ($\Delta\psi$) and ATP are required for translocation of matrix targeted preprotein across inner membrane. N-terminal presequence is present. |
| Carrier pathway (TIM 22 pathway) | TOM40, small TIM chaperones, TIM 22 | Inner mitochondrial membrane | Internal targeting hydrophobic signal sequence is present. |

1.3 Presequence translocase pathway

The presequence translocase pathway is the oldest and most studied protein import pathway of yeast mitochondria. It is called presequence pathway as most of the preproteins imported through this pathway possess an N-terminal presequence.

1.3.1 Presequence and its receptor sites along the import pathway

Presequences are composed of 20-45 amino acids (varying in length and composition) and are self-sufficient to target the accompanying protein to the mitochondrial matrix (van Loon et al. 1986). The presequences are positively charged with a higher content of basic amino acids. They lack sequence conservation, and are believed to form amphipathic helices (Roise et al. 1986; Vonheijne 1986). They are proteolytically removed after reaching the mitochondrial matrix (as reviewed by Verner and Schatz, 1988; Pfanner and Neupert, 1990).

The presequence carrying preprotein interacts with various receptor sites along the preprotein import pathway. Two hypotheses are proposed for the movement of preprotein from one receptor to the other depending on their mode of interaction with the receptor. According to the acid chain hypothesis (Komiya et al. 1998; Rimmer et al. 2011), a series of acidic receptors, guides the positively charged presequence containing preprotein into the mitochondrial matrix and electrostatic interactions among the preprotein and receptor are important to facilitate the transfer of the preprotein towards the high affinity receptor site. Others suggest that additional noncovalent forces such as hydrophobic interactions are equally important for the interaction of a preprotein with its receptor (Abe 2001). These additional interactions direct the preprotein import to the mitochondrial matrix by sequentially binding preprotein to higher affinity receptor site i.e. binding chain hypothesis (Pfanner 2000).

At the outer mitochondrial membrane, the cytosolic domain of Tom20 cooperates with Tom22 to act as a presequence receptor site (Bolliger et al. 1995; Yamano et al. 2008; Shiota et al. 2011). Furthermore, Tom22 possesses an intermembrane space that also acts as (*trans*) site for binding to preproteins in

cooperation with intermembrane space domains of Tom40, Tom5 and Tom7 (Moczko 1997; Nargang et al. 1998).

Beyond the outer membrane translocase, the components of the inner membrane translocase also have a direct presequence binding site at Tim50, Tim23 and Tim44 (de la Cruz et al. 2010; Marom et al. 2011).

In vivo chemical crosslinking, surface plasmon resonance (SPR) and nuclear magnetic resonance (NMR) spectroscopy were used to identify the presequence-receptor components and to deduce the molecular basis of the presequence receptor interaction (Becker et al. 2012). Although the presequence receptor sites are scattered along the import pathway, until now there is only one known structure of a presequence receptor complex involving the twenty two residue long matrix targeted presequence from rALDH (retinal aldehyde dehydrogenase) with the cytosolic domain of Tom20 lacking first 50 residues (*cys*Δ50 Tom20) from rat.

The *cys*Δ50Tom20-presequence structure highlights the binding of an amphiphilic part of rALDH (¹⁵LSRLL¹⁹) to a hydrophobic groove of (*cys*Δ50) Tom20 (Abe et al. 2000). The side chains of the leucine residues, Leu¹⁸ and Leu¹⁹ are aligned on one side of the amphiphilic helix and are oriented toward the binding groove, that are in close contact with the side chains of residues located in the hydrophobic patch of *cys*Δ50Tom20 (Figure 5). The number of observed intermolecular NOEs between the rALDH and *cys*Δ50Tom20 are too few (5 NOEs) due to the dynamic nature of rALDH peptide in its Tom20 bound form. Additionally, X-ray structures of the same *cys*Δ50Tom20-rALDH complex were obtained via intermolecular disulphide tethering of a designed cysteine mutant of *cys*Δ50Tom20 with a cysteine at the C-terminus of presequence containing different linker sequences (Igura et al. 2005). For two different linkers used, this intermolecular disulphide tethering between the presequence and *cys*Δ50Tom20 resulted in different orientations of the presequence bound at the hydrophobic groove of Tom20. It was suggested that these structures could possibly highlight the snapshots of the solution conformation of the *cys*Δ50Tom20-presequence complex where the presequence can bind in different conformations highlighting the highly dynamic nature of the presequence-receptor complex.

Vectorially, moving preprotein from the outer membrane and inner membrane to the mitochondrial matrix requires a series of weak interactions among the presequence

containing preprotein and its receptors such as presequence-Tom20 (cytosolic). The presequence-Tom20 (cytosolic) interaction has a low affinity in the micro molar range (Abe 2001). More such weak presequence-receptor interactions along the preprotein import pathway that are transient in nature need to be characterized.

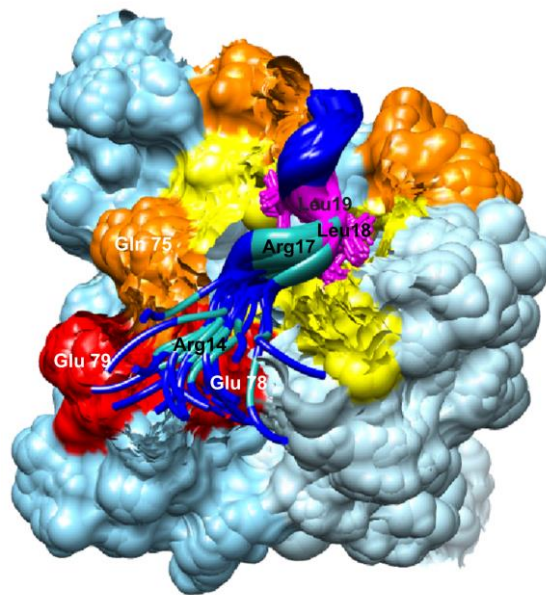


Figure 5: Representation of the binding groove of $\Delta 50$ Tom20 and the bound presequence peptide (PDB id 1OM2). (A) Surface representation of $\Delta 50$ Tom20 (residues 57–124) in cyan. Hydrophobic residues comprising the hydrophobic patch (Phe⁷⁰, Leu⁷¹, Ile⁷⁴, Leu¹⁰⁶, Val¹⁰⁹, Leu¹¹⁰, Thr¹¹³), and Gln (Gln⁶⁷, Gln⁷⁵, Gln¹⁰², Gln¹⁰⁴, Gln¹⁰⁵, Gln¹⁰⁸, Gln¹¹¹, and Gln¹¹²) and Glu (Glu⁷⁸ and Glu⁷⁹) in the peripheral region are colored in yellow, orange, and red, respectively. The bound peptide is drawn as a ribbon with magenta for Leu, sea green for Arg, and blue for others. The side chains of Leucines in peptide points into the binding groove and are highlighted.

1.3.2 Functions and components of the presequence translocase

The presequence translocase can be dissected into three parts:

- A) The core subunits comprising Tim23, Tim17 and Tim50 (also collectively called Tim23^{core}),
- B) The accessory subunits (such as Tim21, Pam17, Mgr2) that are regulatory in nature and are involved in the dynamic functioning of the presequence translocase,
- C) Motor associated subunits such as Tim44, Mge1, Pam18, and Pam 16.

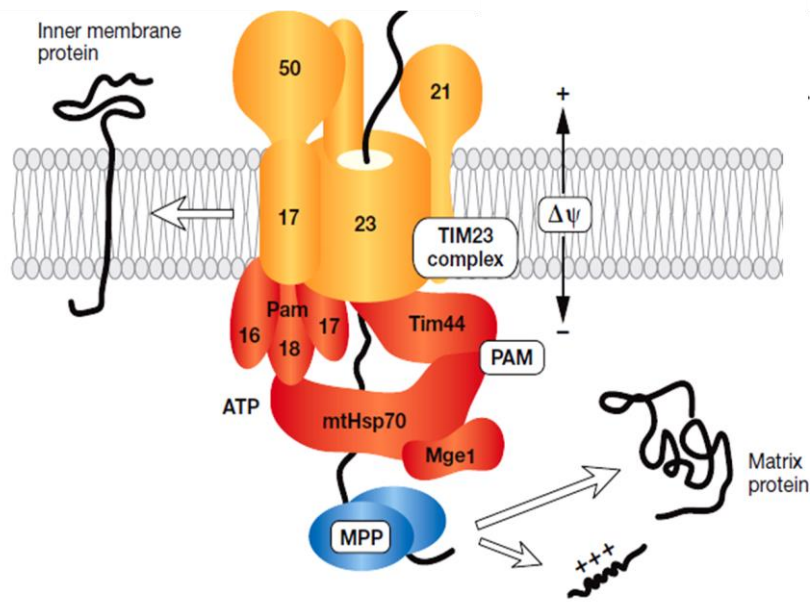


Figure 6 : Components of the presequence translocase. The presequence translocase of the inner membrane (TIM23 complex) consist of the core components Tim23, Tim17 and Tim50. The accessory component Tim21 permits association/dissociation of the presequence translocase-associated motor (PAM complex) with the TIM23 complex. The PAM complex contains of its central player mtHsp70, which is transiently anchored at the translocase by Tim44 and requires the additional components Pam18, Pam17, Pam16 and Mge1 for promotion and maintenance of several ATP-dependent cycles. The presequence translocase complex is capable of dual sorting of mitochondrial preproteins either to the inner mitochondrial membrane or into the mitochondrial matrix depending upon the translocation signal (presequence). Figure adapted from (Bolender et al. 2008), Copyright (2008), with permission from Nature Publishing Group.

Tim23 is the first core subunit of the presequence translocase complex that was discovered and thus the presequence translocase complex is also known as TIM23 complex. The presequence translocase is involved in sorting of preproteins either into the mitochondrial matrix or laterally into the inner mitochondrial membrane. It has diverse functions that include

- a) Receiving preproteins from the outer mitochondrial membrane translocase by forming a translocation contact site with components of the outer mitochondrial membrane translocase and by providing direct receptor sites for presequence recognition for incoming matrix-targeted preprotein,

- b) Sorting a presequence carrying preprotein to the matrix by formation of a translocation pore in the inner mitochondrial membrane and by assembling a mitochondrial import motor on the matrix side of TIM23,
- c) Lateral sorting of the inner mitochondrial membrane proteins,
- d) Interaction with respiratory chain complexes to form super-complexes.

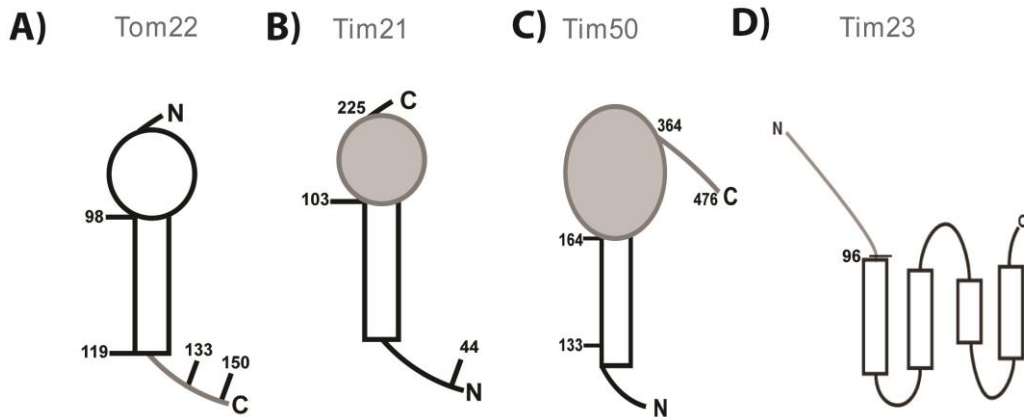


Figure 7: Topology of the components of the outer mitochondrial translocase (Tom22) and inner mitochondrial translocase (Tim21, Tim23, Tim50) studied in this thesis. For each of the component, numbers denote the length of the construct studied, N and C as N- and C-terminus, respectively. The membrane part of each component is represented as a rectangle and the soluble globular domain as a circle. The intermembrane space domains are colored grey. Details of each component is described in the text.

1.3.2.1 Tim23

Tim23 along with Tim17 and Tim50 forms the core of presequence translocase. It is topologically composed of an N-terminal soluble intermembrane part, and a C-terminal membrane part (comprising of four predicted helical transmembrane segments) that together with Tim17 forms the translocation pore across the inner mitochondrial membrane (Martinez-Caballero et al. 2007) (Figure 7D). The translocation pore formed by reconstitution of recombinant Tim23p in mitochondrial vesicles showed that it can form a channel that is sensitive to substrate (presequence) and a membrane potential across the inner mitochondrial membrane (Truscott et al. 2001). Electrophysiological studies and other biochemical assays indicate that the channel formed by Tim23 has a diameter of 13–24 Å and is voltage gated (Lohret et al. 1997; Schwartz and Matouschek

1999; Truscott et al. 2001). It has been proposed that dimerization of the intermembrane space domain of Tim23 and its association with Tim50 regulate the activity of the translocation pore (Meinecke et al. 2006). Electrophysiological experiments have shown that the translocation pore of inner mitochondrial membrane was further studied using a fluorescence mapping approach in which Tim23 was cross-linked with solvent sensitive fluorophore 7-nitrobenz-2-oxa-1,3-diazolyl (NBD) at different sites. The changes in the fluorophore property at each site in Tim23 were monitored as a function of presequence import. This study revealed that TM2 (transmembrane segment 2) lines the aqueous part of the channel and interacts with the incoming preprotein (Alder et al. 2008).

To date, structural details for the membrane part of Tim23 are absent. Tim23 gets dynamically associated with various subunits of its own translocase complex and of the outer membrane translocase complex to facilitate preprotein import (Alder et al. 2008; Gevorkyan-Airapetov et al. 2009; Tamura et al. 2009).

1.3.2.1.1 *N-terminal Tim23*¹⁻⁹⁶

The intermembrane space domain (ims) of Tim23 consists of the N-terminal 96 residues of full length Tim23 in *S.cerevisiae*. The first twenty amino acids of Tim23ims are sensitive to protease cleavage in intact mitochondria and have been proposed to traverse the outer mitochondrial membrane. This led to an unusual two membrane topology of Tim23 in yeast (Donzeau et al. 2000). It was initially proposed to facilitate the apposition of the outer membrane translocase and inner membrane translocase before the discovery of Tim21 subunit. The membrane insertion of Tim23 into the outer membrane has been shown to depend on the presequence load and association with Tim50 (Popov-Celeketic et al. 2008). Additionally, removal of the first 50 residues of Tim23 does not significantly affect the mitochondrial protein import (Chacinska et al. 2003) but cause morphological defects in mitochondria in yeast (Donzeau et al. 2000). However, deletion of the first 24 residues of Tim23 causes a lethal phenotype in yeast (Davis et al. 2000).

N-terminal residues (50-96) of Tim23 have also been proposed to dimerize and regulate channel activity (Bauer et al. 1996). They also help Tim23 to form the active part of the dynamic complex of TIM23 by associating with various other subunits e.g. Tim50, Tim21 and TOM40 (*in organelle* chemical crosslinking and SPR data) (Gevorkyan-Airapetov et al. 2009; Tamura et al. 2009).

1.3.2.2 *Tim21*

Tim21 is an accessory subunit of the presequence complex that is anchored in the inner mitochondrial membrane with a single transmembrane domain and its C-terminal domain protruding into the intermembrane space (ims) (Chacinska et al. 2005; Mokranjac et al. 2005) (Figure 7, B). The crystal structure of the intermembrane space domain involving residues 103-225 has been reported (PDB id: 2CIU) (Albrecht et al. 2006). This structure is characterized by a novel fold involving mixed alpha-beta sheet with two helices and eight beta strands. The first α -helix points towards the antiparallel β -sheet forming three hydrogen bonds (Figure 8).

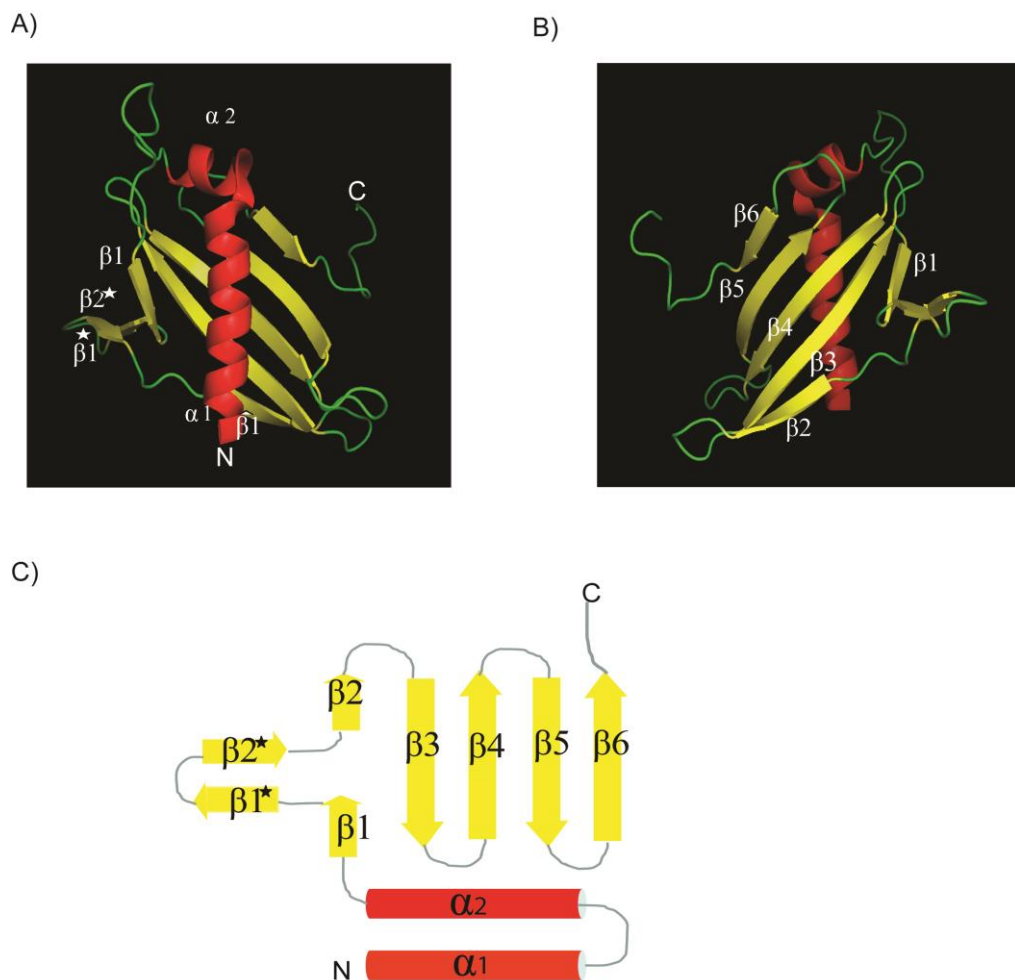


Figure 8: Crystal structure of *Tim21*_{IMS} (PDB id: 2CIU). (A) Ribbon representation of *Tim21*_{IMS} with labeled α -helices (red) and β -strands (yellow). (B) Schematic representation of the *Tim21*_{IMS} fold.

Moreover, the surface of the molecule harbors conserved positively charged patches. It was suggested that recombinant *Tim21*_{ims} can electrostatically bind to *Tom22*_{ims} (Albrecht et al. 2006). During the matrix-targeted preprotein import,

Tim21ims-Tom22ims might serve as a link between the two mitochondrial translocases by forming a translocation contact site (Mokranjac et al. 2005). Tim21ims can also cross-link to a subunit (Qcr 6) of respiratory chain complex. This interaction between Tim21ims and Qcr 6 links the TIM23 to the respiratory complexes that is important for sorting of the inner membrane proteins at low membrane potential and promotes membrane-potential-dependent protein sorting of preproteins into the inner mitochondrial membrane (van der Laan et al. 2006; Wiedemann et al. 2007).

1.3.2.3 *Tim17*

Tim17 is an integral membrane protein, and along with Tim23, it forms an integral part of pore across the inner mitochondrial membrane. The TIM23 pore morphology and its voltage sensitivity are affected on depleting Tim17 in yeast (Martinez-Caballero et al. 2007). Tim17 has shortest N-terminal intermembrane space domain of 17 residues that has two charged aspartates as the voltage sensors. The mutational analysis of these residues shows impaired protein import and are suggested to be important for gating of TIM23 pore (Meier et al. 2005). On the other hand, Tim17 was also shown to be essential for dual functioning of the presequence complex by sorting preprotein to either the inner mitochondrial membrane or mitochondrial matrix (Popov-Celeketic et al. 2008). Additionally, it was also found to be essential for linking the core presequence translocase to the motor part via its physical interactions with Pam18ims (D'Silva et al. 2008).

1.3.2.4 *Tim50*

Tim50 is essential for the viability of yeast cells. Topologically, it is composed of an N-terminal inner mitochondrial membrane-transversing segment and a C-terminal hydrophilic intermembrane space domain (Figure 7C). (Geissler et al. 2002; Yamamoto et al. 2002; Mokranjac et al. 2003). The C-terminal intermembrane space domain fosters the biological functions (Mokranjac et al. 2009). The intermembrane space domain of Tim50 (Tim50ims) is the first receptor among the various components of presequence complex (Mokranjac et al. 2009) for the incoming preprotein. It collaborates with Tom22 to receive the presequence containing preprotein (Shiota et al. 2011). Tim50ims has been shown to regulate the activity of the pore forming Tim23 in the inner mitochondrial membrane antagonistically to presequence (Meinecke et al. 2006). Tim23

also interacts with its intermembrane space and transmembrane helix 1 (Alder et al. 2008; Tamura et al. 2009) with Tim50ims. The presequence binding region in Tim50 is localized mainly to the C-terminal 100 residues (Schulz et al. 2011). Recently, a crystal structure was reported for intermembrane space domain of Tim50 involving residues 164-361 (PDB: 3QLE), which consists of five α -helices and nine β -strands (Qian et al. 2011). A striking feature of the structure is a protruding β -hairpin that consists of highly conserved residues and has been suggested to be the region involved in interaction with Tim23. Furthermore, the groove formed at the bottom of this protruding beta hairpin has been proposed to bind the presequence. In a nutshell, the structure of monomeric Tim50 was described to accommodate both prime functions as a presequence and a Tim23 receptor (Figure 9).

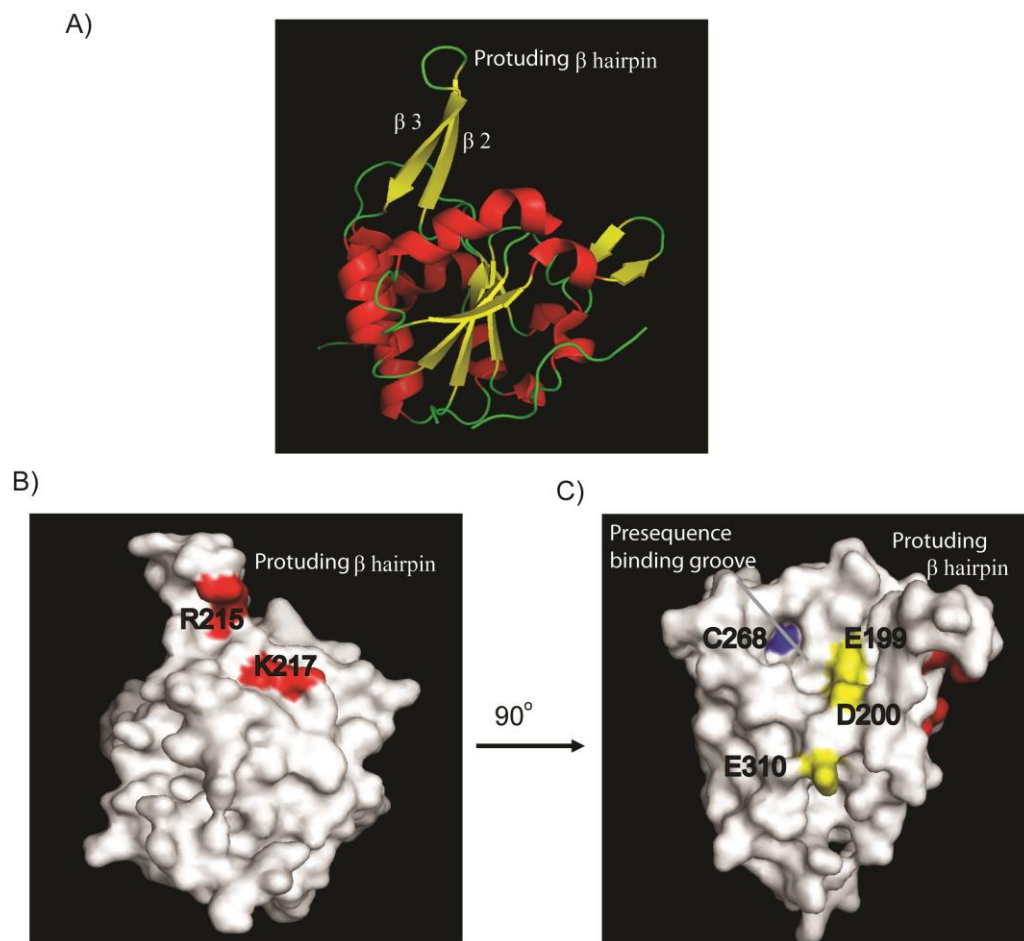


Figure 9: Structure of Tim50ims domain (PDB id: 3QLE). (a) Ribbon representation of the monomer structure of Tim50ims colored according to the secondary structure elements- α -helices (red) and β -strands (yellow). Surface representation of Tim50ims is highlighting the protruding β -hairpin and presequence binding groove important for binding to Tim23 (B) and presequences (C) respectively. The key residues are labeled.

1.3.2.5 Pam18 (Tim14)-Pam 16 (Tim16) complex

Pam 18 is composed of an N- terminal intermembrane space domain, followed by membrane part and a C-terminal soluble J like domain in mitochondrial matrix (Mokranjac et al. 2003). It binds to the core subunits of presequence translocase and forms a stable complex with Pam16 (Kozany et al. 2004). Pam 16 is a matrix protein peripherally attached to the inner mitochondrial membrane (Kozany et al. 2004). The structure of Pam18 (J protein) and Pam16 (J like protein) complex is known (Mokranjac et al. 2006). Both the proteins act as chaperones and Pam 18 stimulates the ATPase activity of Hsp70 whereas Pam 16 helps to recruit Pam 18 to the core subunits of the presequence translocase (D'Silva et al. 2003; D'Silva et al. 2008).

Besides the membrane potential across the inner membrane, the ATPase activity of Hsp70 is the energetic force for translocation of preprotein to the matrix.

1.3.2.6 Pam 17

Pam 17 is another accessory subunit anchored in the inner mitochondrial membrane and is exposed to the mitochondrial matrix (van der Laan et al. 2005). Pam 17 antagonistically modulates the binding of Tim21 to Tim23^{core} and is known to affect the stability of motor components Pam 18-16 to Tim23^{core} (van der Laan et al. 2005; Popov-Celeketic et al. 2008). Moreover, the lack of co-isolation of Pam17 with protein A-tagged Tim21 led van der Laan *et al.*, (2005) to suggest that Pam17 acts as the import motor component whereas Popov-Celeketic, Mapa et al. (2008) demonstrate that Tim21 and Pam17 bind directly to the Tim17–Tim23 core of the TIM23 complex (details in section 1.3.4).

1.3.2.7 Mgr2

Mgr2 is a small (~10.5 kDa), hydrophobic accessory subunit in the inner mitochondrial membrane. Mgr2 has been shown to link Tim21 to Tim23^{core} and aid in coupling the presequence translocase to respiratory chain complexes (Gebert et al. 2012). It is required for efficient import of matrix targeting proteins at elevated temperatures. Mgr2 is necessary for viability of cells lacking the mitochondrial genome (petite-negative phenotype) (Dunn et al. 2006).

1.3.2.8 *Tim44*

Tim44 is a matrix translocating peripheral subunit of the presequence translocase that interacts electrostatically with the inner mitochondrial membrane rich in negatively charged cardiolipin (Marom et al. 2009). It acts as a connector subunit for the TIM23 complex and PAM subunits by interacting with core subunits Tim17-Tim23 along with Pam 16-18 and Hsp70 as reviewed by (Mokranjac and Neupert 2010). Tim44 can be cross-linked *in vivo* and *in vitro* with Hsp70 via its N terminal domain (Bomer et al. 1997). The crystal structure of C terminal Tim44 is known (Weiss et al. 1999).

1.3.2.9 *Mt Hsp70 and Mge1*

MtHsp70 belongs to the molecular chaperone family which prototypically possesses an N terminal nucleotide binding site and substrate binding sites. It cycles between the ADP bound high affinity state and low affinity ATP bound state. Hsp70 has a low intrinsic ATPase activity and Mge1 act as a nucleotide release factor (Miao et al. 1997; Liu et al. 2001; Mokranjac and Neupert 2010).

1.3.3 An overview for transport of preprotein using the presequence translocase

Most of the matrix targeting preproteins and some of the inner membrane subunits are translocated with the help of general receptors at the outer membrane that includes Tom22 and Tom20 at the cytosolic site (Komiya et al. 1998; Ahting et al. 1999). Tom40 forms the channel required to transverse the outer membrane (Hill et al. 1998). The incoming preprotein from Tom40 can interact with the *trans* presequence binding site formed by Tom22 ims and/or Tom40ims (Court et al. 1996; Shiota et al. 2011) and can then be further transferred to Tim50, that has been proposed to be the first receptor of the presequence translocase for the incoming preprotein (Geissler et al. 2002; Mokranjac et al. 2003; Mokranjac et al. 2009). The incoming preproteins carrying only the N-terminal matrix targeting presequence are directed towards matrix through the inner membrane pore formed by Tim23 and Tim17 in the presence of membrane potential across the inner mitochondrial membrane. Moreover, the preproteins

containing an additional inner hydrophobic signal are laterally translocated to the inner membrane (Mokranjac and Neupert 2010; van der Laan et al. 2010).

In addition to a membrane potential, the complete translocation of an incoming preprotein into the matrix requires an additional energy source ATP, used by the motor part of the presequence translocase (mtHsp70). When the unfolded preparation reaches the exit site of inner membrane pore, the motor subunits interact with the core subunits of TIM23. This interaction between the core subunits and the motor part of the presequence translocase (mtHsp70, Pam16-Pam18) occur with the help of Pam17 and a scaffold like protein Tim44. The mtHsp70 chaperone binds to the incoming preprotein in an ATP dependent manner and works in coordination with chaperones, Pam16, Pam18 and Mge1 to vectorially translocate the preprotein into the matrix. Pam18 contains a J-domain that stimulates the ATPase activity of mtHsp70 and nucleotide release factor Mge1, facilitates the release of ADP from mtHsp70. The fully imported matrix-targeted preprotein is finally processed into a mature functional protein by mitochondrial matrix peptidase that cleaves off the presequence.

1.3.4 Dynamic changes and multiple interactions drives preprotein translocation

The translocation of the matrix-targeted preprotein translated by the nucleocytoplasmic machinery requires multiple interactions between the presequence of preprotein and the subunits of mitochondrial translocases. These interactions are well coordinated in and across various subunits of the outer and inner mitochondrial membrane translocases. The orientation of various subunits among translocase of outer and inner mitochondrial membrane must harmonize in response to the incoming preprotein to maximize the translocation process. TIM23 complex sorts the mitochondrial proteins into the inner mitochondrial membrane and mitochondrial matrix. Different models have been proposed to explain the functional mechanism of the dual sorting of the presequence translocase to translocate the incoming preprotein into either the mitochondrial matrix or the inner mitochondrial membrane. According to the modular model, two forms of TIM23 known as Tim23^{sort} and Tim23^{motor} exist in dynamic equilibrium with each other. Tim23^{sort} consist of core subunits Tim23, Tim50 and Tim17 along with Tim21 that is required for lateral sorting of inner membrane proteins and recruits the respiratory complexes. Tim23^{sort} receives the incoming preprotein from the TOM complex and helps in translocating preproteins to the inner

mitochondrial membrane. On the other hand, Tim21 antagonistic, Pam 17 subunits cause dissociation of Tim21 and thus lead to the recruitment of motor components (Pam 18-Pam16, Hsp70 etc.) to core subunits of presequence translocase resulting in Tim23^{motor}. Tim23^{motor} imports an N terminal carrying preprotein to the mitochondrial matrix (Chacinska et al. 2005; Chacinska et al. 2010; van der Laan et al. 2010; Marom et al. 2011) (Figure 10).

In contrast, the single entity model proposes that TIM23 complex always exists as a preassembled complex during the above said dual sorting function of presequence translocase. However, depending upon the incoming preprotein and sorting signals, conformational changes are required among the subunits of TIM23 either for incorporating preprotein to the inner membrane or importing it to the mitochondrial matrix (Popov-Celeketic et al. 2008) (Figure 11).

Besides these differences, both models agree that TIM23 is highly dynamic and these dynamic associations among subunits of TIM23 are essential for sorting preproteins to their destined compartments in mitochondria.

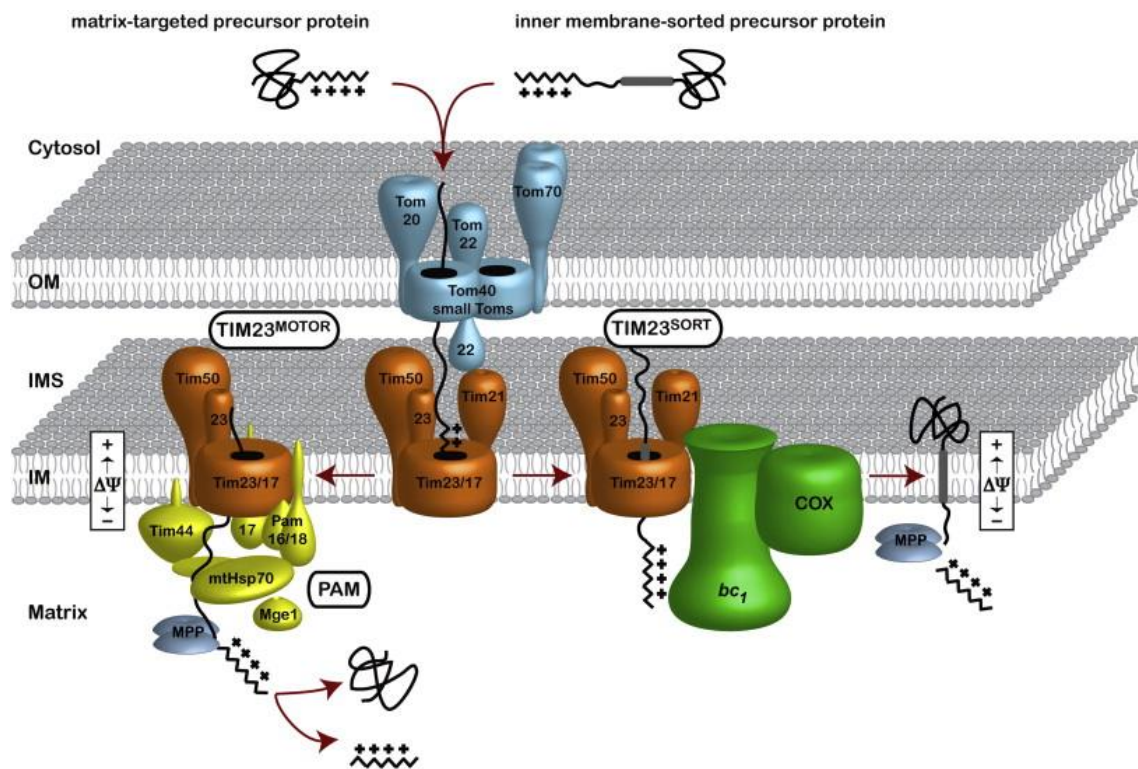


Figure 10: **Two distinct functional forms of the TIM23 complex** are involved in the sorting the preprotein either laterally into the inner membrane or into the mitochondrial matrix. The incoming preprotein is transferred from the TOM complex to the TIM23 complex with the help of Tim21, which interacts with the Tom22 and links the two translocases. The two forms of the TIM23 translocase are referred as 1) TIM23^{MOTOR} that lacks the Tim21 subunit and interacts with the motor part of TIM23 i.e. PAM (yellow color). TIM23^{MOTOR} translocates the matrix targeted N-terminal presequence to the matrix. The N-terminal presequence carrying preproteins after reaching the TIM23 translocase finally get translocated into the mitochondrial matrix with the help of mtHsp70 and co-chaperones in the presence of both the membrane potential and ATP. (2) TIM23^{SORT} interacts with the respiratory chain super complexes (complex III and IV subunits cytochrome bc_1 and COX respectively) to translocate additional hydrophobic stop transfer signal containing preproteins laterally into the inner mitochondrial membrane. MPP in the mitochondrial matrix removes the N-terminal presequence to yield the mature functional protein. Reprinted from (van der Laan et al. 2010), Copyright 2010, with permission from Elsevier.

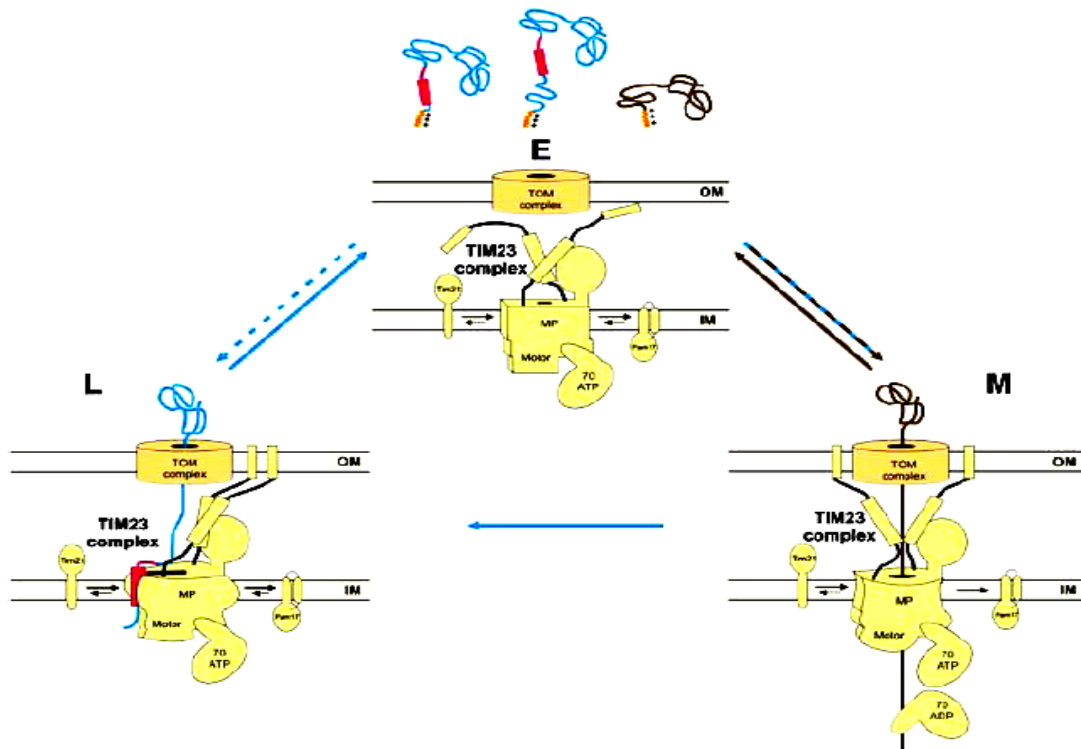


Figure 11: Remodeling of the TIM23 complex during preprotein translocation–Single entity model: The TIM23 exists in the three functional modes that dynamically interconvert into each other by a series of conformational changes in its subunits known as Empty (E) mode and Matrix translocation mode (M) and Lateral sorting mode (L). Subunits Tim21 and Pam17 bind to the complex in a dynamic manner (shown by bidirectional black arrows in Inner mitochondrial membrane) and affect the transitions between E, M and L modes in an antagonistic way. The incoming presequence induces a series of conformational changes until the membrane part adopts the conformation compatible with the import of the unfolded preprotein into the matrix (brown/cyan arrow in E to M direction). If no additional signal appears within the preprotein, matrix translocation will be completed by a number of cycles of the import motor. This mode of the translocase is defined as matrix translocation (M). In absence of preprotein at the completion of import cycle, the TIM23 complex exists in the E-mode (brown arrow in M to E direction). When additional sorting signal is recognized the translocase undergo a conformational change to attain the L-mode (cyan arrow in M to L direction) and initiates the lateral sorting of the preproteins. The translocase TIM23 can change from the L mode directly back to the E mode (cyan arrow in L to E direction) in case of (single) transmembrane containing preprotein. If the incoming laterally sorted preprotein lacks the presequence, then TIM23 complex can directly change from its E-mode to the L-mode (broken cyan line pointing from E to L). The essential components of the membrane part and the import motor of TIM23 complex act as a single entity during the entire reaction cycle. Reprinted from (Popov-Celeketić et al. 2008), Copyright(2008) with permission from NPG.

1.4 Interaction network of Tim23

The functional form of a protein in super-complex biological machinery can depend on its interactions with other subunits. The biological significance of a protein in complex machinery is defined by its function. This influences the biological outcome of the super-complex, which is fine-tuned with the interplay of interacting subunits (Pawson and Nash 2003). The past decade has explored the functional roles of one gene product (protein) with multiple interaction partners. This has led to the emergence of interesting interaction networks based on experimental data that can define the dynamic nature of biological super machineries (Waksman 2005).

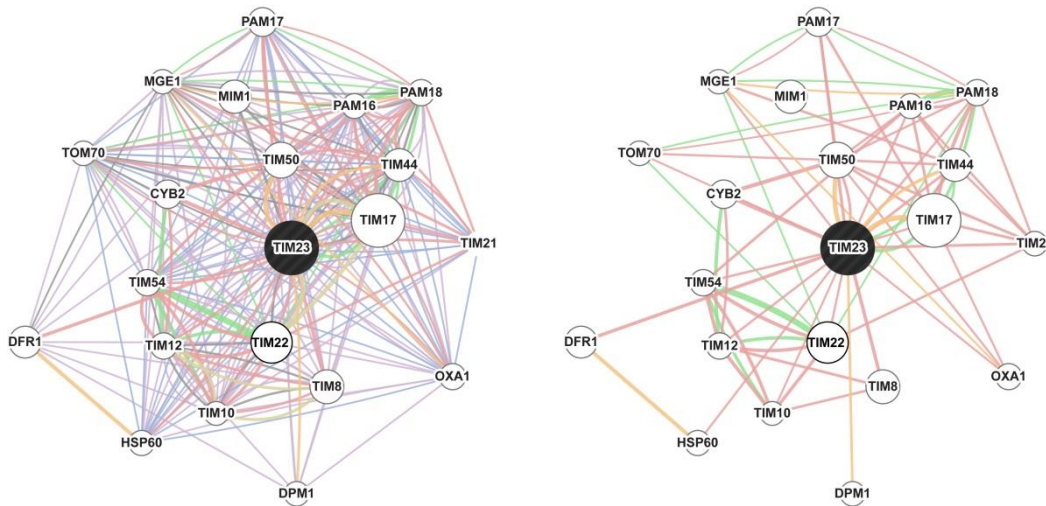


Figure 12: Information based Interaction network of Tim23 gene as deduced from GENEMANIA server (Warde-Farley et al. 2010), where each gene is represented as node and the line connecting two nodes represents an interaction between those genes. Different line color highlights different sources for interactions among two genes with purple as co-expression, pink as physical interaction, green as genetic interaction, blue as co-localization, orange as predicted and grey as other means. Right figure exclusively shows experimental interactions based on co-expression, genetic and physical interaction captured by western blot and affinity tag studies. TIM23 here in the picture represents Tim23 gene not the presequence translocase complex.

Tim23 is the central subunit of presequence translocase complex. It has been studied for almost two decades in yeast. Tim23 has been proposed to interact with the outer membrane and was cross-linked *in vivo/in vitro* with an incoming presequence, Tim50, Tim13-Tim8, Tom22, Tim17, Tim44, Tim21, mtHsp, Pam16 and Pam17. The

interactions among the various other subunits are present as depicted in Figure 12 A. The wide spectrum of protein complexes and modular states that presequence translocase (such as Tim23^{core}, Tim23^{motor}, TIM23-respiratory chain complex and TIM23-TOM22 super complex) adopts in order to perform dual sorting of different preproteins into either the inner mitochondrial membrane or mitochondrial matrix is efficiently described by the number and nature of interactions among its subunits. The molecular interactions that the Tim23 gene could possess in yeast cell are shown in Figure 12. The available interactions of Tim23 based on experimental data are highlighted on the right in Figure 12. It clearly highlights the need of understanding these interactions in details to unravel the molecular basis of the presequence translocase and its function.

1.5 Intermembrane space domains (ims) of two mitochondrial translocases

In addition to the membrane part, most subunits of the outer and inner mitochondrial translocases have a soluble domain in the intermembrane space known as intermembrane space domains. These intermembrane space domains are important sites for the reception of preproteins at the presequence binding sites like *trans* binding site at TOM formed by Tom22ims, Tom40ims *etc.* (exact composition of this site is not known) and presequence binding site at Tim50 and Tim23 (Moczko, Bomer et al. 1997; Esaki, Shimizu et al. 2004; Tamura, Harada et al. 2009; Shiota, Mabuchi et al. 2011). The intermembrane space domains also interact with each other to form translocation contact sites to vectorially direct the preprotein emerging from the outer mitochondrial membrane pore towards the inner mitochondrial membrane pore. The *in vivo* crosslinking experiments between Tim50-Tom22 and Tim23-Tom22 and with recombinant expressed Tim21ims and Tom22ims support this idea (Tamura et al. 2009; Shiota et al. 2011). The gating of Tim23-Tim17 channel in the inner mitochondrial membrane is sensitive to incoming presequence and interaction with Tim50ims (Truscott et al. 2001; Meinecke et al. 2006).

Intermembrane space domains of mitochondrial translocase such as Tim23ims and Tim50ims are essential and the mutations at important sites *i.e.* residues L71, 78, 64 in Tim23ims and L279, 282, 386 in Tim50ims affects *in vitro* protein import and the

growth of yeast (Tamura et al. 2009). On the other hand, Tim21ims and Tom22ims shows growth defects in yeast at only at elevated temperature indicative of their either accessory regulatory role or important role under certain conditions (Court et al. 1996; Chacinska et al. 2005; Mokranjac et al. 2005).

The intermembrane space domains studied in this thesis and their functional importance are listed in Table 2.

Table 2: Designed constructs length and functional importance of various intermembrane space domains of mitochondrial presequence translocase involved in this study.

| Subunit/ translocase name | IMS domain Residues | Functional Importance | Reference |
|--------------------------------------|------------------------------------|--|--|
| Tom22 | 133-153 | Constituent of trans presequence receptor site of outer membrane, Crosslinks with Tim50 , Tim21, Tim23 | (Court et al. 1996; Moczko et al. 1997; Tamura et al. 2009; Shiota et al. 2011) |
| Tom40 | 361-387 | Putative constituent of trans receptor site | (Pfanner 2000) |
| Tim23 | 1-96 | Dimerises to regulate the TIM23 channel activity in response to presequence. Interacts with Tim50. Can be cross-linked to Tim21, Tim17, TOM40 in organelle. Essential for maintaining morphology of mitochondria | (Donzeau, Kaldi et al. 2000; Tamura, Harada et al. 2009; Bauer, Sirrenberg et. al.1996; (Gevorkyan-Airapetov et al. 2009); |
| Tim21 | 103-227 | Role in formation of translocation contact site & interacts with respiratory chain &TIM23 subunits. | (Mokranjac et al. 2005; Albrecht et al. 2006) |
| Tim50 | 164-476 | Interacts with Tim23 presequence, and regulate channel activity. Can be cross-linked in organelle with Tom22 | (Meinecke et al. 2006; Tamura et al. 2009; Schulz et al. 2011; Shiota et al. 2011); |

1.6 Intrinsically disordered proteins (IDPs) as hub sites in interaction networks

The biological function of a protein is often associated with its well-defined secondary and tertiary structure. The discovery of proteins that lack stable secondary or tertiary structures and exist as linear polypeptide chains under physiological conditions, the so-called intrinsically disordered proteins (IDPs), has challenged the conventional protein structure-function-paradigm (Wright and Dyson 1999). These disordered proteins are rich in small hydrophilic residues (G, A, Q, S, P, K) and contain few sparsely placed bulky hydrophobic residues (V, L, I, M, F, W, Y), which prevent them to adopt a well-defined secondary structure. Disordered proteins often have a low sequence complexity with a large number of repeats and are highly abundant in nature. The IDPs can possess transient secondary structures and long range interactions. The IDPs are involved in the various physiological processes such as cell regulation and signaling wherein they bind to multiple partners with high specificity and low affinity. The IDPs exist in heterogeneous conformations that are represented by an ensemble (Rezaei-Ghaleh et al. 2012). These ensembles of conformations are highly dynamic and provide a huge challenge to characterize them under experimental conditions. NMR spectroscopy enables to study the proteins with larger conformational heterogeneity and is the method of choice to study these IDP's under physiological conditions. (Section 1.7).

The function of biological machineries can be visualized as an interaction network wherein each protein (gene product) represents a node and the interaction between two such nodes can be represented as a line (Figure 12) (Pawson and Nash 2003). The most interesting feature of such an interaction network is the presence of a hub node that has the maximum number of interactions originating from it. Interestingly, the hub nodes constitute mainly of amino acids with less sequence diversity, a characteristic of IDP's, and adopt various conformations contributing to their multivalent interactions (Dosztanyi et al. 2006).

Current post-genomic and proteome-wide studies are placing IDPs at the center of interaction networks as hub proteins. Disordered proteins are also shown to be a key determinant of scale free networks, which are severely affected by removal of hub

proteins but are resistant to the removal of other nodes. These scale free interaction networks can be used to represent biological systems. The removal of disordered hub protein in a scale free network can be linked to a pathological condition such as cancer. (Uversky 2011; Csermely et al. 2012).

The rugged conformational landscape with multiple local energy minima at the bottom of the protein energy funnel is one of the unique features of disordered proteins that help them to explore a variety of partners. Moreover, the absence of well-defined structure helps in favorable interactions due to the availability of excessive interaction surface and fine coordination interpolating specificity and binding promiscuity. (Uversky 2011; Uversky 2011; Uversky 2012). Thereby, disordered proteins are the most suitable candidates for hub proteins in biological interaction networks, as they possess the structural plasticity, i.e. they can adopt different structures depending on the partner protein. Some of them possess well separated linear motifs which can be used to bind different partners either at the same time (simultaneous interactions) or along the interaction pathway (sequential interaction) (Fuxreiter and Tompa 2012). Hence, they are often players in various regulatory pathways in cells

1.7 Intrinsically disordered proteins and NMR Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy exploits the magnetic moment, an intrinsic property of nuclear spins to probe its chemical environment. The method development for isotopic enrichment and labeling in *E.coli*, higher dimensional NMR (3D, 4D up to 7D) and development of hardware (cryoprobes, high magnetic fields, carbon detection probes) have made it feasible to study biomolecules at atomic resolution. Moreover, this technique can provide structural as well as dynamic details of biomolecules at physiological conditions. It is complementary to X-ray crystallography to study folded proteins but is currently, the only technique to study high-resolution atomistic details of disordered proteins.

The characterization of IDPs by NMR spectroscopy primarily requires the assignment of its observed resonances as in case of folded proteins. The proton chemical shift dispersion along the amide proton resonances is frequently used to characterize the state of a protein of interest. Unfolded proteins or disordered proteins have poor amide proton chemical shift dispersion in comparison to folded proteins. Due

to poor dispersion in the amide proton (H^N) dimension higher dimensional experiments (such as Automated Projection Spectroscopy, APSY) conjugated with higher dispersion nuclei (nitrogen and carbonyl based) are most prevalent. IDPs benefit from recording APSY due to their short effective rotational correlation time in solution as compared to globular proteins, which allows 5D to 7D NMR experiments with high sensitivity, by which the problem of chemical shift degeneracy of IDPs is circumvented. Additionally, to simplify and accelerate the basic assignment process, computational methods for assignment (such as MARS) are used (Narayanan et al. 2010).

The study of α -synuclein and 441-residue long-disordered protein tau revealed that disordered proteins are not the fully extended chain of amino acids rather are compact and possess a transient long range tertiary interactions and secondary structure propensity (Bertoncini et al. 2005; Mukrasch et al. 2009).

Disordered proteins exist as the ensembles of conformations that are in dynamic equilibrium and the structural properties of ensembles can be measured by various NMR parameters (Table 3). The most recent way of representing the structural characteristics of IDPs is to generate an ensemble of conformers. Various NMR parameters listed in Table 3 coupled with the computational methods and other biophysical methods like SAXS can be used to generate ensembles of IDPs that are in accordance with the experimental data (Mittag, Kay et al. 2010).

The quantitative description of an ensemble of conformations of a disordered protein is coded in its NMR parameters and can be decoded using computational methods like molecular dynamics (MD) simulation, and recently developed computational approaches for ensemble generation and selection such as ENSEMBLE, Flexible-Meccano and ASTEROIDS (as reviewed by Mittag and Forman-Kay 2007; Rezaei-Ghaleh et al. 2012; Schneider et al. 2012). Remarkably, various functional aspects of the disordered proteins such as their state upon interaction (folding upon binding or disorder in the bound state), regions involved in binding, and post-translational modifications can be monitored in real time by NMR spectroscopy. These features make NMR spectroscopy a pivotal biophysical technique to unravel structural and dynamic aspects of disordered proteins.

Table 3: List of common NMR parameters highlighting their source of origin and information provided to characterize the protein.

| <i>NMR Parameter</i> | <i>Source/Importance</i> | <i>Information</i> |
|---|---|--|
| Chemical shift(δ) | Resonance frequency of spin nuclei in magnetic field with respect to a standard For same nuclear spin it depends on the chemical environment i.e. electronic environment of nuclear spin | Used to assign resonances (backbone side chain atoms in protein sequence); Can also be used to characterize the different conformers and possible states of protein |
| J-coupling | Through bond interactions among spins | Used to calculate torsion angles; Characterize the presence of residue in a helix, loop or sheet |
| Relaxation measurements R_1 , R_2 , HetNOE | Spin-spin and spin lattice relaxation | Global and site-specific correlation time and dynamics at various time scales (picoseconds to nanosecond, exchange contribution from micro-milliseconds. |
| Residual dipolar coupling (RDCs) | Dipole–Dipole interaction among spins due to partial alignment in solution | Average orientation of internuclear vectors with respect to external magnetic field; global orientation of molecule in solution; Sign of RDC can interpret the content of secondary structure (PPII, β -sheet, α -helix) and tertiary contacts. |
| Secondary Chemical shifts | Difference in observed chemical shift and the random coil chemical shift | Content of secondary structure, to characterize residues involved in a α -helix, β -sheet, loops |
| Paramagnetic relaxation enhancement (PREs) | Difference in spin-spin relaxation rate in the presence and absence of a paramagnetic center | Long range distance between the spin label and resonance nuclei up to 25Å |
| Nuclear Overhauser effect (NOEs) | Cross relaxation induced spin polarization between neighboring spins (through space magnetization transfer) | Short-range NOEs are used to sequentially assign residues and define local geometry. Long range NOEs (upto6Å) defines topology of molecule. |

1.8 Fuzzy complexes

Although ordering of IDPs in the bound form is in line with the structure function paradigm, recent years have envisioned examples where IDPs retained disorder in the bound form. The magnitude of the disorder varies in the bound form and recently structural polymorphism has been appreciated in the bound form with multifunctional p53 that has a large number of interacting partners (Dunker et al. 2005; Oldfield et al. 2008). In order to generalize the importance of the disorder in the continuum of structures the phenomenon of fuzziness was described by P Tompa. Structural polymorphism in the bound form means that in complex (between a disordered protein and a folded protein); the same disordered segment upon binding to different interacting partners can form different secondary structures. This accounts for static disorder. However, when the disordered protein remains fully or partially flexible even in the bound form and is represented by rapidly interconverting conformers instead of one stable conformer it accounts for dynamic disorder (Fuxreiter 2012; Fuxreiter and Tompa 2012).

The structural description of the fuzzy complexes could be envisioned by the complex of Sic1p-Cdc4, in which Cdc4p has a single interaction site for multiple scattered suboptimum interaction motifs in disordered Sic1p (Mittag et al. 2008). These suboptimum motifs in the disordered Sic1p need to be phosphorylated prior to its interaction to Cdc4.

In Cdc4-Sic1p complex, interaction motifs of disordered Sic1 are in rapid exchange with a single binding site of Cdc4 (Mittag et al. 2008). The disordered Sic1p in the Cdc4-Sic1p complex was represented as an ensemble describing the average properties of conformational heterogeneity in the bound form. The fast exchanging weak complex of Sic1-Cdc4 requires at least six suboptimal motifs in Sic1p for its function. This Cdc4-Sic1p complex was further described by a polyelectrostatic model, in which long-range electrostatic interactions confer the overall affinity to the complex.

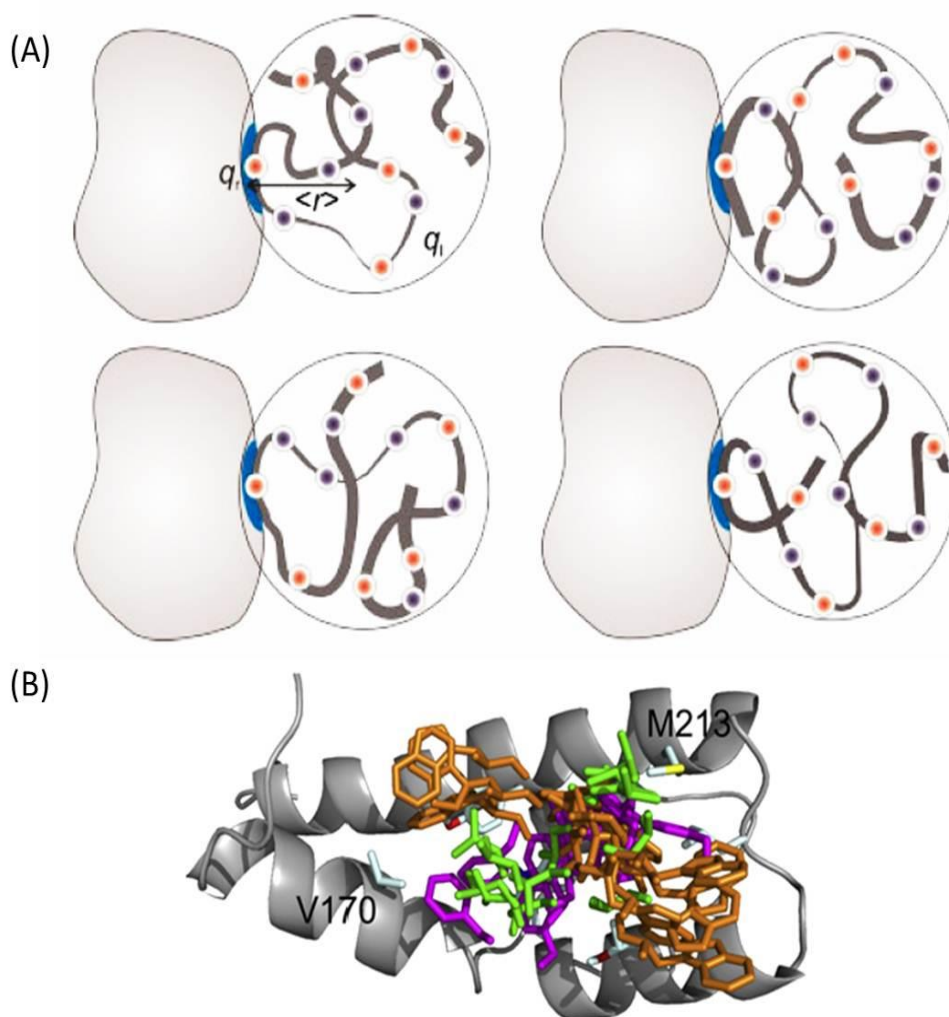


Figure 13: **Fuzzy complexes:** (A) Interaction of intrinsically disordered protein, Sic1p (ribbon) interacting with its folded receptor, Cdc4 (grey shape) can be explained by a Polyelectrostatic model. Sic1p has several distinct interaction motifs and exists in an ensemble of conformations (illustrated by four different ribbon representations of Sic1p). Sic1p, possesses positive and negative charges (illustrated by blue and red circles, respectively) that results in a net charge q_i , whereas the interaction site in the receptor Cdc4 (light blue) has a charge q_r . The $\langle r \rangle$ represents the effective distance between the interaction site and the center of mass of the intrinsically disordered protein Sic1p. Overall electrostatics governs the interaction between the Sic1-Cdc4 complexes. (B) Ribbon representations for the ensemble of HADDOCK generated structures for Gcn4-activation domain binding to the ABD1 (activation binding domain, gray). The important side chains of important residues of Gcn4 such as W120 (orange), L123 (green), and F124 (magenta) relative to ABD1 (gray ribbon) are shown from the ensemble highlighting the different modes of binding of residues W120, L123, and F124. Reprinted from (Mittag et al. 2010;), Copyright 2010, with permission John Wiley & Sons, Ltd. and (Brzovic et al. 2011), Copyright (2011) with permission from Elsevier.

The complex between the activation domain (AD) of the transcription activator Gcn4 and activator-binding domain (ABD) of its co activator Gal-11 is another interesting example of a fuzzy complex. Gcn4 has two activation domains whereas Gal-11 has three conserved activator binding domains. Each of the three-activator binding domains of Gal-11 binds to the activation domain of Gcn4 with a micro molar affinity. The binding of these three domains of Gal-11 contribute additively to the transcription activation of Gcn4 specific genes (Herbig et al. 2010).

The central AD domain of Gcn4 is disordered in its free form in solution. Gcn4 (AD) interacts via three hydrophobic residues (W, L and F) and forms an α -helix upon binding to ABD1 of Gal-11. This complex of Gcn4 (AD) and Gal-11 (ABD1) is governed by hydrophobic interactions (Brzovic et al. 2011). In this complex, Gcn4 exists in multiple and rapidly exchanging conformations bound to Gal-11. Thus, the protein interface is dynamic and allows the activator domain of Gcn4 to interact with various activator binding domains of Gal-11 etc. activators to regulate the gene transcription.

The previous examples structurally exemplify dynamic biological interactions involving strong binding as a result of a cumulative effect of several weak binding sub-sites present in transiently ordered or disordered proteins.

1.9 Aims and outline of thesis

The current post-genomic era blended with hybrid technologies of structure determination is aimed at understanding the molecular details of the fundamental cellular processes. Mitochondrial preprotein import is one of such fundamental processes in cell encoding the details of protein trafficking. TIM23 or presequence translocase complex is an important machinery involved in the sorting of presequence carrying preproteins either to the mitochondrial matrix or to the inner mitochondrial membrane. It has been exclusively chosen to study in detail as the multifunctional biochemical role of Tim23, the central component of TIM23 is known. However,

- 1) Residue specific information and the structural details about interaction of Tim23 with the mitochondrial membranes, Tom22, Tim50 and presequence are not available.

- 2) Different models have been proposed for the functional working of TIM23 to understand the dual sorting of incoming preprotein either into the inner mitochondrial membrane or into the mitochondrial matrix. The two models differs in terms of the dynamics of association of the various TIM23 core subunits with motor associated subunits either as two forms as Tim23^{motor} and Tim23^{sort} or as a single entity. The definitive interaction site between the various subunits involved in Tim23^{sort} and Tim23^{motor} are unknown. Tim21 is a key player in both the models. Additionally, it has been shown biochemically that the molecular basis of function of the presequence translocase gets modulated by the incoming preprotein but structural information about subunits involved in binding to the presequence is not available.
- 3) Despite the availability of X-ray structure of Tim21ims, an accessory component of TIM23, its complex with other subunits is not available.
- 4) Translocation contact sites are known to be formed by subunits of outer and inner mitochondrial translocase in the intermembrane space of mitochondria. However, the direct interaction between the subunits involved in formation of translocation site has not been characterized.

To address the above mentioned issues at the atomic level and to establish a residue specific interaction network of Tim23, we have employed NMR spectroscopy as the main technique to study protein- protein (Subunit1-Subunit2 of TIM23) and protein peptide (Subunit of TIM23-presequence) interactions.

We have selectively aimed to map out the interactions among the subunits of TIM23 [i.e. Tim21, Tim23, and Tim50] and with TOM complex [i.e. Tom22 and Tom40] in the intermembrane space. Consequently, the first part of the results section 3.1 describes purification and solution properties of intermembrane space domains of Tim23, Tim21, Tim50 and Tom22.

The potential presequence binding sites for Tim21, Tim50, Tim23 and Tom22 that constitute the second part of results of this study are illustrated in section 3.2. Tim23-presequence interaction is characterized in detail.

Third part of results includes the additional interactions at translocation contact site involving ims domains of Tom22, Tom40, Tim23 and Tim21 are described in section 3.3.

Interactions of Tim23ims with mitochondrial membrane mimetic are characterized in section 3.4. The detailed interaction study of intermembrane space domains of Tim21 and Tim23 is described in section 3.5 to gain insights into nature of the Tim21-Tim23 interaction

Consolidating the results from the binary interactions helped us to reveal a residue level interaction network of Tim23ims with its various biochemical ligands. Finally, the comprehensive network of interactions is used to elucidate the mechanism of TIM23 during the import of mitochondrial matrix targeting preprotein.

2 Materials and Methods

2.1 Materials

2.1.1 Chemical reagents

Isotopic reagents (^{15}N Ammonium chloride, ^{13}C -D Glucose, Deuterium oxide) were purchased from Sigma Aldrich and Cambridge Isotope Laboratories. Phospholipids and detergents including DHPC (1,2-diheptanoyl-*sn*-glycero-3-phosphocholine), DOPC (1,2-dioleoyl-*sn*-glycero-3-phosphocholine), DOPE (1,2-Dioleoyl-*sn*-Glycero-3-Phosphoethanolamine), DOPS (1,2-dioleoyl-*sn*-glycero-3-phospho-L-serine) and 1,1,2,2 tetraoleoyl cardiolipin (Cardiolipin or CL) were obtained from Avanti Polar Lipids, Inc. Paramagnetic reagents for PRE measurements-MTSL [(2, 2, 5, 5-Tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methylmethanesulfonylthioate] was purchased from Toronto Research Chemicals Inc. whereas Cys-Ph-TAHA (cysteinyl-phenyl-triaminohexaacetate) lanthanide binding tag was kindly provided by Prof. Christian Griesinger.

The Pf1 phages, alignment media for RDC measurements, were purchased from ASLA biotech.

The molecular weight standard marker for SDS-PAGE, DNA ladder, restriction enzymes, and CIP alkaline phosphatase were obtained from Fermentas. T4 DNA ligase and Phusion High-Fidelity PCR Kit used for cloning genes of interest were obtained from Thermo Scientific. The Quik change mutagenesis kit was obtained from Agilent Technologies, Inc. The following resins for protein purification were used. Glutathione Sepharose 4B (GE Healthcare), Ni-NTA Agarose (Qiagen), SP-Sepharose (Applied Biosystems) and a Superdex 75 (16/60) column was purchased from GE Healthcare.

The peptide sequence corresponding to rALDH presequence, Tom40ims, Tom22E1C were synthesized by K. Overkamp in our department using standard solid-phase peptide synthesis (FMOC).

2.1.2 Bacterial strains

XL-2 Blue MRF' (Minus Restriction) strain is deficient in all known restriction systems (McrA-, McrCB-, McrF-, Mrr-, HsdR-), preventing cleavage of cloned DNA that carries cytosine and/or adenine methylation, which is often present in eukaryotic DNA and cDNA, was obtained from Stratagene.

BL21 (DE3) contains the DE3 lysogen encoding the lac promoter and T7RNA polymerase. This strain is deficient of lon and omp-t proteases and is therefore suitable for expression of non-toxic genes.

BL21 Codon plus RIL chemically competent cells carry extra copies of the argU, ileY, and leuW tRNA genes. The tRNAs encoded by these genes recognize the AGA/AGG (arginine), AUA (isoleucine), and CUA (leucine) codons, respectively.

2.1.3 Expression vectors for purification intermembrane space domains

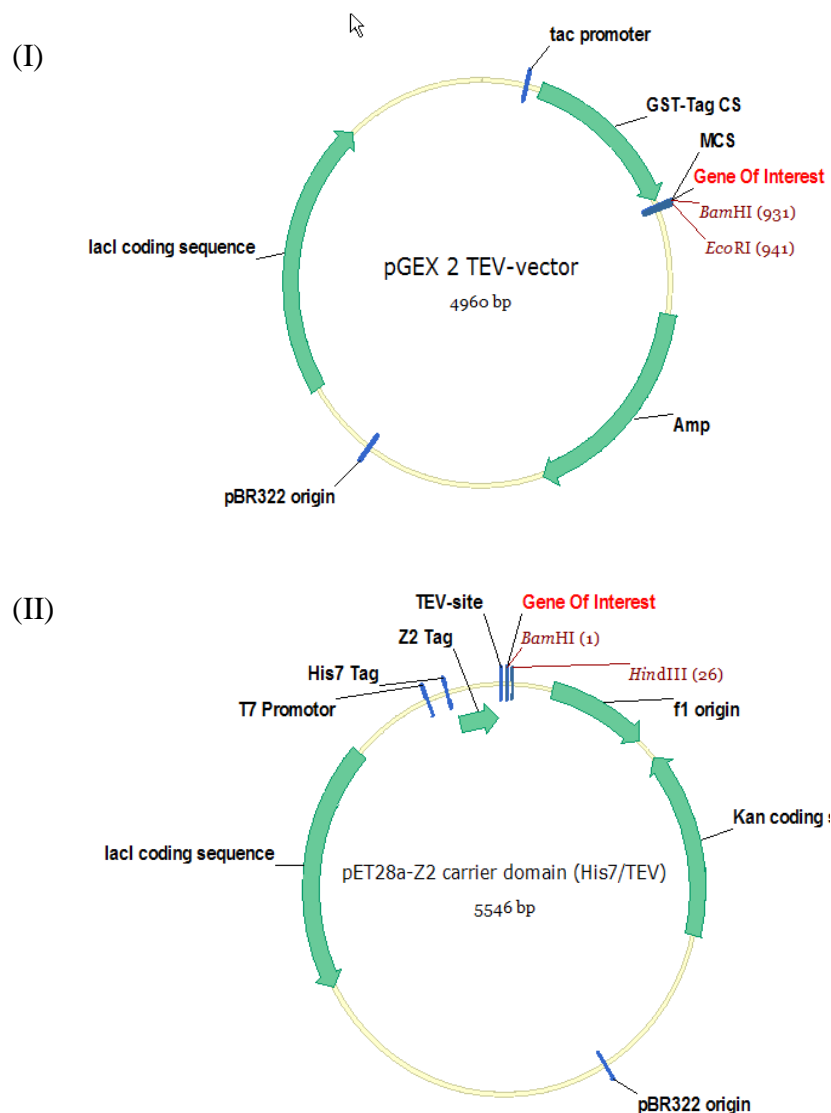


Figure 14: Expression Vector maps of pGEX2-TEV (I) and pET28a His7-Z₂-TEV (II) used in this study for expression of various constructs of intermembrane space domains of yeast mitochondria along with the antibiotic resistance marker and key restriction sites.

pGEX2-TEV encodes an N-terminal GST fusion protein followed by TEV cleavage site and the protein of interest. Tim23ims was cloned into this vector.

pET28a- His-7-Z₂-TEV encodes an N-terminal His tag , followed by Z₂ fusion domain, which is separated from protein of interest by TEV cleavage site. Tom22ims, Tim21ims and Tim50ims were cloned in this vector.

2.1.4 Growth media

Table 4: Composition of various culture media for protein expression in *E.coli*.

| Medium | Constituent | Amount |
|-------------------|---|--------------|
| M9 minimal medium | M9 salt (5x)* | 200 ml |
| | Trace elements# | 10 ml |
| | Thiamine HCl (5 mg/ml) | 6 ml |
| | CaCl ₂ (2 M) | 0.05 ml |
| | MgSO ₄ (1 M) | 2 ml |
| | D-Glucose, ¹³ C-D-glucose(20% w/v) | 20 ml |
| | ¹⁵ NH ₄ Cl (0.25 g/ml) | 4 ml |
| | H ₂ O, deionized or D ₂ O | 758 ml |
| *M9 salt (5x) | Na ₂ HPO ₄ | 33.9 g |
| | KH ₂ PO ₄ | 15 g |
| | NaCl | 2.5 g |
| | H ₂ O, deionized or D ₂ O | add 1,000 ml |
| #Trace elements | FeSO ₄ .7 H ₂ O | 0.6 g |
| | MnCl ₂ .4 H ₂ O | 0.115 g |
| | CoCl ₂ .6 H ₂ O | 0.08 g |
| | ZnSO ₄ .7 H ₂ O | 0.07 g |
| | CuCl ₂ .2 H ₂ O | 0.03 g |
| | H ₃ BO ₃ | 0.002 g |
| | (NH ₄) ₆ Mo ₇ O ₂₄ .4 H ₂ O | 0.025 g |
| | EDTA | 0.5 g |
| | H ₂ O, deionized or D ₂ O | add 100 ml |

| | | |
|----------|---|----------------------------------|
| LB Media | Tryptone, Yeast Extract, NaCl H ₂ O | 10g 5g 10g add 1,000 ml |
|----------|---|----------------------------------|

2.1.5 Stock solutions of antibiotics

Table 5: List of stock solutions of antibiotics used in this study.

| Stock Name | Concentration |
|-----------------|--------------------------------------|
| Ampicillin | 25 mg /ml in H ₂ O (500X) |
| Kanamycin | 70 mg/ ml in H ₂ O(1000X) |
| Chloramphenicol | 36mg/ml in ethanol(1000X) |

2.6 Buffers for purification using pET28-HIS-7-Z₂TEV vector

Table 6: Composition of buffers used for the purification of protein cloned into pET28-HIS7- Z2-TEV vector. *means Elution Buffer either contains 100mM or 500mM imidazole.# means additional buffer components present in case of Tim50ims.

| Name | Chemicals | Concentration | pH |
|-----------------------|--|--|-----------|
| Lysis buffer | Tris HCl NaCl Imidazole PMSF Protease inhibitor tablet(EDTA free) DTT | 20mM 500mM 10 mM 1mM 1 tablet/100 ml 1mM | 8.0 |
| Wash Buffer | Tris HCl NaCl Imidazole PMSF Complete (EDTA free) DTT | 20 mM 500 mM 20 mM 1mM 1 tablet/100 ml 1mM | 8.0 |
| Elution buffer* | Tris HCl NaCl Imidazole PMSF DTT | 20 mM 500 mM 100 mM or 500mM 330 µl/100 ml 1mM | 8.0 |
| Gel filtration buffer | Hepes NaCl DTT# | 20 mM 150 mM 5 mM (optional) | 7.5# /7.2 |
| Tev Cleavage buffer | Tris HCl NaCl DTT | 20mM 100mM 1mM | 8.0 |
| NMR buffer | Hepes NaCl DTT# | 20mM 50mM 5mM(optional) | 7.2 |

Table 7: Composition of buffers used for the separation of Tim21ms from TEV digested mixture of Z₂ and using cation exchange chromatography (SP-Sepharose).

| Name | Chemicals | Concentration | pH |
|-------------------|-----------|---------------|-----|
| Wash Buffer | Tris HCl | 20 mM | 7.5 |
| | NaCl | 50 mM | |
| Elution buffer | Tris HCl | 20 mM | 7.5 |
| | NaCl | 500 mM | |

2.1.6 Scientific Software's

Table 8: List of software's used in this study.

| Software Name | Source | Purpose |
|------------------------|---|---|
| Topspin 2.1 | Bruker Bio spin, Karlsruhe, Germany | Acquisition, |
| NMRPipe/NMRDraw | spin.niddk.nih.gov/NMRPipe (Delaglio et al. 1995) | processing and |
| Sparky 3 | www.cgl.ucsf.edu/home/sparky (Goddard and Kneller) | visualization of NMR spectra. |
| PROSA | Güntert, P., Dötsch, V., Wider, G. and Wüthrich K,1992 | processing and visualization of NMR spectra. |
| Pymol | DeLano Scientific LLC, Palo Alto, CA, USA | Protein Structural visualization and analysis |
| ORIGIN 8.1 | Origin Lab Corporation. | Chart graphics, |
| IGOR Pro 6 and | Wave Metrics, Inc. | mathematical |
| Microsoft Excel | Microsoft Office | analysis and calculations |
| Rosetta 3.2.1 | http://www.rosettacommons.org/ | |
| FlexPepDock | London N, Raveh B, Cohen E, Fathi G, Schueler-Furman O | Protein peptide Docking |
| Rosetta | (Raveh et al. 2011) | |
| GAPRO | (S. Hiller et al,2005) | Peak picking for APSY spectra |
| MARS | Jung, YS and Zweckstetter, M | Automated backbone assignment |

2.2 Molecular Biology Methods

2.2.1 Constructs design for intermembrane space domains

The various constructs used in this study were designed based on available literature for Tom22, Tim21, Tim23 and Tim50 whereas for Tom40, I-Tasser was used for predicting the C-terminal intermembrane space domain guided by its topology (Appendix Figure 10). These intermembrane space domains possess diversity in size ranging from 30 residues (Tom22) to 312 residues (Tim50) and are listed with their functional importance Table 9.

Table 9: Designed construct length of various intermembrane space domains of mitochondrial presequence translocase involved in this study.

| IMS domain name | Residues | Expression Vector |
|-----------------|----------|--------------------------------|
| Tom22 | 133-153 | pET28a-HIS-Z ₂ -TEV |
| Tom40 | 361-387 | chemical synthesis |
| Tim23 | 1-96 | pET28a-HIS-Z ₂ -TEV |
| Tim21 | 103-227 | pET28a-HIS-Z ₂ -TEV |
| Tim50 | 164-361 | pET28a-HIS-Z ₂ -TEV |

The ensemble view of solution properties of isolated intermembrane space domains of mitochondrial translocases in *S.cerevisiae* is described in section 3.1.

2.2.2 Cloning of various intermembrane space domains:

The constructs involving intermembrane space domains of Tim21, Tom22, and Tim50 were amplified using plasmids as template purchased from Harvard plasmid repository with primers listed in Appendix Table 1 and ligated into the pET28a-7His-Z₂-TEV vector using standard protocols described in *Sam brook et al*. Tim23ims has been cloned earlier in our group into the pGEX2-TEV vector by Laura de Cruz.

2.2.3 Site directed mutagenesis

Point mutations were introduced with the Quik-change® site-directed mutagenesis kit (Stratagene), following the instructions of the supplier. Successful mutagenesis was verified by DNA sequence analysis. Primers used in this thesis are listed in Appendix Table 1.

2.2.4 Expression of various intermembrane space domains

The plasmid encoding the subunit of interest was transformed into the expression strains and grown under the conditions described in following Table 10.

Table 10: Expression parameters for various constructs involved in this study.

| Construct name | Expression host cell | Inducer(IPTG) concentration | Temperature/ | | Media tested |
|--|----------------------|-----------------------------|-----------------|----------------|-------------------------|
| | | | Harvesting time | post induction | |
| GST-Tim23 ¹⁻⁹⁶ | BL21(DE3)RIL | 1mM | 25°C/12 hrs. | | LB,M9,D ₂ O |
| Z ₂ -Tim21 ¹⁰³⁻²²⁷ | BL21(DE3) | 1mM | 25°C/12 hrs. | | LB,M9, D ₂ O |
| Z ₂ -Tom22 ¹²⁰⁻¹⁵² | BL21(DE3) | 1mM | 37°C/5 hrs. | | LB,M9 |
| Z ₂ -Tim50 ¹⁶⁴⁻³⁶¹ | BL21(DE3) | 1mM | 16°C/16 hrs. | | LB,M9,D ₂ O |

To enhance the expression of intermembrane space domains the N-terminal fusion protein domains GST and Z₂ were used (Smith and Johnson 1988; Bogomolovas et al. 2009). However, during cloning and purification the care was taken to effectively remove them and to isolate native intermembrane space domains for further studies. Depending upon the type of sample required the growth media was appropriately selected. The small scale (10ml) of primary culture was established in same media prior to large expression cultures.

Preadaptation of cells was required in case of deuterated samples for the Tim50ims. This involved the adaptation of transformed *E.coli* cells by growing them in 10 ml deuterated M9 media with increasing concentrations of D₂O ranging from 0%, 33%, 55% and 99%. The final pre-adapted 99% deuterated culture with cell density of 0.8AU at OD₆₀₀ was then used to inoculate the large expression culture.

2.2.5 Purification method for intermembrane space domains cloned into pET28a-His7-Z₂-TEV

The transformed *E.coli* cells were grown expressed and harvested as shown in Table 10 and were resuspended in pre-cooled lysis buffer (Table 6) (approximately 50ml per litre of culture). Thoroughly resuspended cells were incubated with 0.5mg/ml of lysozyme for 10 minutes on ice and then sonicated 8 to 9 times for 20 seconds (50 % of the maximal power output) with pause intervals of 1 minute. The efficiency of sonication was monitored visually for decrease in viscosity of cell lysate. The sonicated lysate was centrifuged at 20000 rcf for 45 minutes at 4°C. The supernatant was loaded onto a 5ml bed volume Ni⁺²-NTA column, pre-equilibrated with lysis buffer.

The soluble supernatant containing fusion protein was incubated with Ni⁺²-NTA beads for 1 hour at 4°C, after which the flow through was collected and the Ni⁺²-NTA beads were washed extensively with 3-5 column volume of pre-cooled wash buffer (Table 6) to remove weakly bound impurities. The bound fusion protein was eluted using 10 ml of elution buffer (Table 6) containing 100mM imidazole, followed by 10 ml of elution buffer containing 500mM imidazole. The eluted fractions were collected and analyzed by 15% SDS-PAGE.

The fractions containing the fusion protein were pooled and dialyzed against TEV cleavage buffer (Table 6) for ~12 hours and then digested with TEV at ratio of 1:1000 (TEV: protein) at room temperature for ~10 hours.

The TEV digestion mixture containing the undigested fusion protein, Z₂ domain, and TEV were separated from the protein of interest using different strategies (Figure 15, Route 1 and 2) were used for the separation of protein of interest from TEV digested mixture.

The TEV digested mixture was again loaded onto a Ni⁺²-NTA column (2ml bed volume) pre equilibrated with TEV cleavage buffer (Table 6). This step efficiently

removed the His tagged TEV and the Z₂ domain while the protein of interest (Route 1 in Figure 15, used for separation of Tom22ims and Tim50ims) was present in the flow through. However, whenever the efficiency of separation of the protein of interest from TEV digested mixture was not better than 80% other chromatographic purification steps for purification such as cation exchange chromatography were used (for instance, Route 2 in Figure 15 used for Tim21ims). Size exclusion chromatography (Route 4 and 5 in Figure 15 was used for Tim21ims and Tim50ims) and RP-HPLC (Route 3 in Figure 15 for Tom22ims) were further used to purify the respective proteins of interest to homogeneity. The overall purification strategy employed for the generation of various ims domain is summarized in Figure 15.

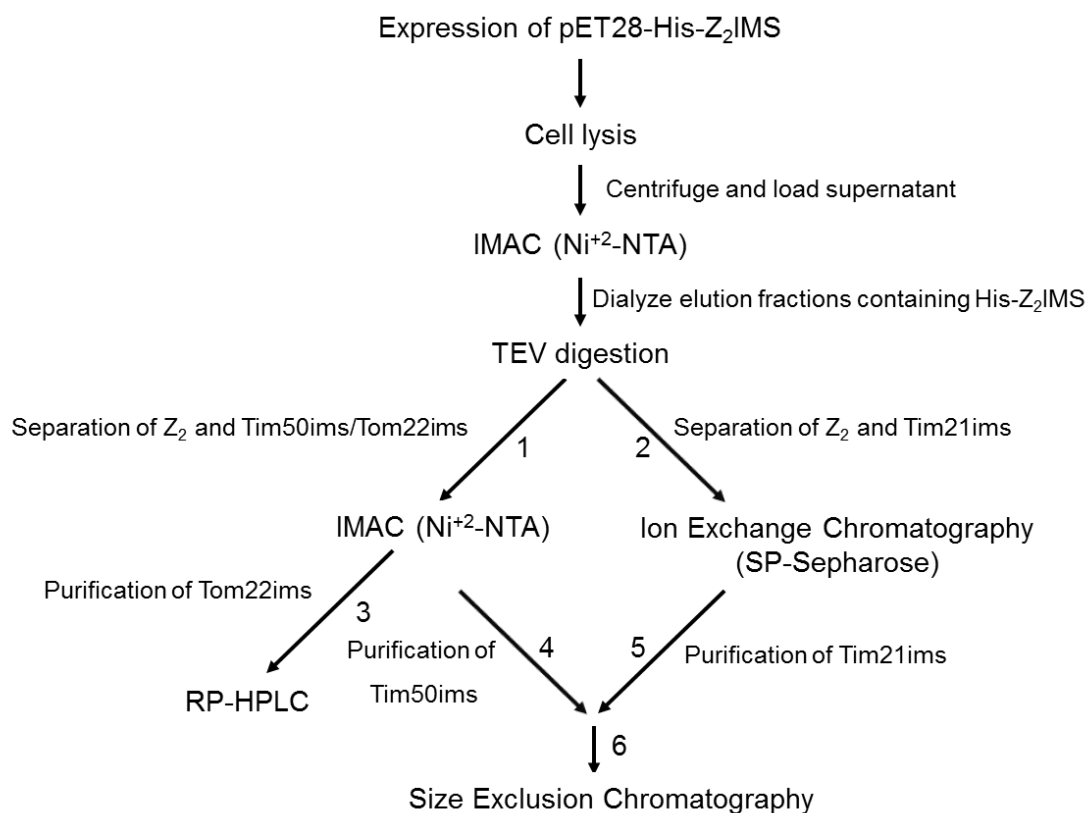


Figure 15: Summary of purification strategy for various IMS domains cloned into pET28-His7-Z₂-TEV. See appendix section 1.1 for details.

2.3 Mitochondrial liposome preparations

The synthetic lyophilised phospholipids DOPC, DOPE, DOPS and Cardiolipin corresponding to constituent molar ratios of outer mitochondrial membrane and molar inner mitochondrial membrane as 48:28:10:4 and 38:24:4:16 respectively, were each weighed and dissolved in 200µl of chloroform. The dissolved phospholipids were mixed and dried under N₂ gas for 30 minutes. The semidried mixture of phospholipids was additionally lyophilised for 12 hours to remove the traces of organic solvent. NMR buffer was added to the dried phospholipid film and vortexed slowly to homogeneously suspend the phospholipids in the buffer. The resuspended phospholipid mixture was sonicated in a precooled water bath at 4°C for 2 hours with 5 minutes pulses alternating with 15 minutes pauses resulting in multilamellar vesicles (MVs).

The phospholipid stock solution with total lipid content of either 10 mg or 25 mg were prepared as described in previous paragraph. These stock solutions were diluted appropriately in NMR buffer prior to NMR measurements.

2.4 Spin labelling with the paramagnetic reagent

2.4.1 MTSL

The protein of interest was modified to incorporate a cysteine residue either using site directed mutagenesis or chemical synthesis to attach the thiol-specific nitroxide spin label reagent MTSL (Table 11). DTT was removed prior to labeling from the buffer using 10ml Sephadex-G25 fast desalting columns (PD-10 columns, Amersham Biosciences). The 100µg/µl MTSL in cold acetone were added to the protein solution in a three-to five-fold molar excess and was incubated for ~1.5h at room temperature and loaded again onto a PD-10 column (pre-equilibrated with NMR buffer) to remove the excess MTSL. For Tom22EIC peptide, RP-HPLC was used to remove the excess MTSL. The efficiency of the paramagnetic labeling reaction was confirmed using ESI-MS which showed the expected increase of molecular weight by 186.3 Daltons.

2.4.2 Cys-Ph-TAHA, a lanthanide tag

The tag Cys-Ph-TAHA (Peters et al. 2011) was pre-loaded in deionized water with a twofold excess of LuCl_3 or TbCl_3 . The aqueous stock solutions of Cys-Ph-TAHA (5 eq., 10 mg mL^{-1}) and LuCl_3 or TbCl_3 (10 eq., 10 mg mL^{-1}) were combined and incubated for 1 hour at 20°C .

Tim21ims was labelled with Cys-Ph-TAHA at residue 128 with its native cysteine. Tim21ims was dialysed overnight into 20mM Tris Cl, pH=8.0 buffer containing 100mM NaCl. The preloaded lanthanide tag solution was incubated with the previously said dialysed Tim21ims at 5 fold molar excess for 4 hours at 4°C . The excess lanthanide ions were removed by centrifugation and the supernatant was buffer exchanged with the NMR buffer to remove excess of tag-bound lanthanides.

2.5 NMR spectroscopy

NMR spectra were recorded with ^{15}N or $^{13}\text{C}/^{15}\text{N}$ labelled samples containing 0.05-1.0 mM of intermembrane space domains of Tim21, Tom22 and Tim23. $^{15}\text{N} / ^2\text{H}$ (75%) labelled samples was used for Tim50ims. The standard NMR buffer for all measurements used was 20mM Hepes, pH=7.2 containing 50 mM NaCl and 5-10% D_2O . All spectra were measured at 288K for Tim23ims and Tom22ims, 298 K for Tim21 and 310 K for Tim50ims (unless stated otherwise) on Bruker 600, 700, 800 or 900 MHz spectrometers equipped with and without cryogenic probes. Spectra were processed using NMRPipe (Delaglio et al. 1995) and analysed using SPARKY (Goddard and Kneller 2006).

2.5.1 NMR assignment of intermembrane space domains

Backbone H^{N} , N, C^{α} resonance assignments of Tim23ims was obtained using combination of HNN, (HA) CANNH previously by Laura de la Cruz. Later on, C' resonances were correlated to H^{N} and N using HNCO (spectra were acquired by Laura de la Cruz). Additionally, the 6-dimensional (6D) APSY-seq-HNCOCANH NMR experiment which correlates two sequentially neighboring amide moieties in proteins via the C' and C^{α} nuclei, with efficient suppression of the back transfer from C^{α} to the

originating amide moiety was acquired with 25 set of projection angles. The automatic analysis of two-dimensional (2D) projections of this 6D experiment with the use of GAPRO (Hiller et al., 2005) provides a high-precision 6D peak list, which was used as the input for the automated sequential assignments of proteins with the assignment software MARS. MARS successfully assigned 80 out of 84 expected resonances with high and medium reliability (84 resonances are expected from 96 residue Tim23 as it contains 6 prolines).

^1H - ^1H -TOCSY and ^1H - ^1H -NOESY spectra were used to assign the resonance specific assignment of unlabeled micellar bound Tim23 $^{1-13}$ peptide.

Tim21ims: The backbone assignment and solution structure of Tim21ims was kindly provided by Lukasz Jaremko and Mariusz Jaremko in our group.

Tom22ims: The backbone assignment of Tom22ims was done by Raghavendran L. Narayanan.

2.5.2 Secondary chemical shift analysis of Tim23

Secondary chemical shift values are characteristic indicators for secondary elements in proteins. For Tim23ims, secondary chemical shift values were calculated as the differences between measured C^α (or C') chemical shifts and the empirical random coil value for the appropriate amino acid type. Random coil values were taken from (Schwarzinger et al. 2001) except for the charged residues. The random coil values for charged residues were taken from Wishart and Sykes as the chemical shifts of these residues are particularly sensitive to pH (Berjanskii and Wishart 2005). The positive, negative and near to zero average C^α and C' values were used for detecting residues involved in formation of α -helix, β -strand and random coil respectively.

For Tom22, random coil values were obtained from Poulsen IDP/IUP random coil chemical shifts using server (http://spin.niddk.nih.gov/bax/nmrserver/Poulsen_rc_CS/)

2.5.3 Estimation of T_2 and the global rotational correlation time

A ^1H (1-1) echo experiment (Sklenář and Bax 1987) with delays of 0.1, 3.1, 6.1, and 7.1 ms was recorded for *wt*-Tim21ims and Tim50ims. Transverse relaxation times T_2 were estimated from two values or from a single-exponential curve fit. The rotational correlation time was calculated as:

$$\tau_c \approx \frac{1}{5 \cdot T_2}.$$

Alternatively, the global correlation time of Tim21 was also estimated by recording a $^{15}\text{N}/^1\text{H}$ -TRACT experiment (Lee et al. 2006). Here, the intensity decay of the integrated signal from the region between 8.5 and 10 ppm was used to determine the relaxation rates $R_2^{N_x H_\alpha^N}$ and $R_2^{N_x H_\beta^N}$, from which τ_c was calculated.

2.5.4 Steady state heteronuclear, ^{15}N -nuclear Overhauser effects (NOEs)

Steady state heteronuclear ^1H - ^{15}N -NOEs were measured using standard Bruker pulse program (hsqcnoef3gpsi). Heteronuclear NOE values are reported as the ratio of peak intensities in paired interleaved spectra collected with (referred as I_{sat} or I) and without (referred as I_0) initial proton saturation period (4s) during the 5s recycle delay. This spectrum was acquired for the free Tim23ims, Tim23ims saturated with 64 fold molar excess of rALDH, Tim23ims saturated with 32 fold molar excess of DHPC and Tim23ims saturated with 16 fold excess of Tim21ims using 128 x 2024 complex points (in t_1 and t_2) and 128 scans at 600MHz.

2.5.5 ^{15}N relaxation measurements

^{15}N relaxation data were acquired at 700MHz for Tim21ims with modified versions of pulse sequences incorporating pulsed field gradients and a WATERGATE for solvent suppression. T_2 times were sampled using different relaxation delays ^{15}N (Δ): 5, 10, 20, 80, 100, and 160 milli-seconds. ^{15}N T_2 relaxation times were extracted by measuring the peak height (I) for all spectra recorded with different relaxation delays

and fitting the single-exponential function using SPARKY relaxation fitting module (rh.py)

$$I=I_0\exp(-\Delta T_2)$$

$$R_2=1/T_2$$

The errors for R_2 value are obtained from the error in fit of R_2 values relaxation fitting module of SPARKY.

2.5.6 RDCs

Residual dipolar couplings (RDCs) $^1\text{D}(\text{HN})$ were measured for Tim23ims by weakly aligning it using Pf1 bacterial phage (7mg/ml) using IPAP-HSQC (Ottiger et al. 1998) in 20mM Hepes buffer containing 50mM NaCl, pH=7.2 at 288K respectively.

2.5.7 ^{15}N - ^1H -HSQC based characterization of protein ligand interactions

The ligand binding to protein could lead to the changes in chemical shift (in fast exchange regime), which is either directly correlated to the changes in chemical environment of the residues involved in interaction or indirectly to the changes in chemical environment of some residues that could arise due to conformation changes in protein, induced upon binding the ligand. These changes are termed as the chemical shift perturbations, which are obtained by subtracting the chemical shift position of resonances in spectrum with a ligand from the reference spectrum lacking the ligand.

The ^{15}N (or ^{13}C)- ^1H HSQC spectra of ^{15}N (or ^{13}C)-labeled sample are recorded in the presence and absence of the ligand (binding partners in our case) at different molar ratios with identical acquisition parameters and constant amount of protein.

Average chemical shift perturbation (CSP), Δ_{HN} , were calculated as

$$\Delta_{HN} = \sqrt{\frac{(\Delta\delta_N / 5)^2 + (\Delta\delta_H)^2}{2}}$$

The changes in an average chemical shift perturbations as a function of concentration of the binding partners was fitted to a single-site binding model using the equation (Cavanagh et al. 2007).

$$\Delta\delta = \Delta\delta_{\max} \left[\frac{([L]_T + [P]_T + K_d - \{([L]_T + [P]_T + K_d)^2 - 4[L]_T[P]_T\}^{1/2})}{2[P]_T} \right]$$

where, $\Delta\delta_{\max}$ is the maximal chemical shift perturbation value at saturation, $[P]_T$ and $[L]_T$ are the total concentration of protein and ligand to K_d (mostly for $K_d > 10^{-6}$ M i.e. in fast exchange regime).

For the intermediate exchange, small changes in chemical shift and intensity changes are observed. However, in slow exchange regime at NMR time scale, only changes in intensity for the resonances are monitored for free protein (reference) with the emergence of a new set of resonances for the ligand bound form. For such cases, the changes in intensity was plotted as intensity ratios (I/I_0) where, I represents the intensity of resonance in molar excess of ligand and I_0 represents the intensity of resonances in free protein.

2.5.8 Intermolecular paramagnetic relaxation enhancement (PRE)

The ^1H - ^{15}N -HSQC spectra for samples (as listed in Table 11) with 1:1 ratio of labeled and unlabeled protein was measured in the presence and absence of the paramagnetic MTSL-tag at indicated cysteine position using identical acquisition parameters. The removal of the spin label for measuring the diamagnetic state was accomplished by addition of (10 molar equivalents to protein) ascorbic acid to the protein solution followed by incubation at room temperature for 30 minutes.

*Table 11 List of samples used for obtaining intermolecular PRE's. Tim21wtC128 highlighted by * is also tagged with another lanthanide based paramagnetic tag Cys-Ph-TAHA. The cysteine was engineered using site directed mutagenesis*

| Labeled Sample | Unlabeled MTSL tagged sample |
|--------------------------|------------------------------|
| ¹⁵ N Tim23ims | Tim21wt C128* |
| ¹⁵ N Tim23ims | Tim21 C128A/S114C |
| ¹⁵ N Tim21ims | Tim23 T11C |
| ¹⁵ N Tim21ims | Tim23 G67C |
| ¹⁵ N Tim21ims | Tom22 E1C |

The intensity ratios ($I_{\text{para}}/I_{\text{dia}}$) were obtained from the intensity of resonances in ¹H, ¹⁵N-HSQC spectra with paramagnetic and diamagnetic ligand. The intensity error in each spectrum was obtained from the S/N ratio. The cut off of 0.8-0.85 was used for identifying non-interacting residues (Battiste and Wagner 2000).

2.6 Error analysis based on S/N ratio

The error in intensity ratios based NMR experiments such as ¹H, ¹⁵N-HetNOEs, PREs or ¹H, ¹⁵N-HSQC titration plots were based on S/N ratio. The intensity value reported by SPARKY for the resonance of interest was used to estimate the noise in spectra by S/N ratio followed by error propagation as per following equation,

$$Error = \frac{I}{I_0} \sqrt{\left(\frac{1}{(S/N)_I}\right)^2 + \left(\frac{1}{(S/N)_{I_0}}\right)^2}$$

2.7 Protein peptide docking using FlexPepDock-Rosetta 3.2

The peptides corresponding to Tim23 (²SWLFGD⁷, ⁶⁷GVEYLDL⁷³, and ⁹⁰SRGWTDD⁹⁶) were manually placed 12Å away from the binding site in Tim21ims (lowest energy NMR structure) based on the chemical shift perturbation data, using editing mode of Pymol. These models of protein peptide complex were refined using FlexPepDock refinement protocol in Rosetta3.2.

25000 decoys were created for each coarse model and ranked according to reweighted rosetta energy score and top 500 models were then clustered using backbone RMSD of 2 Å.

2.8 Distance restraints and structure calculation

The distance restraints for micelle bound Tim23¹⁻¹³ peptide were obtained from the data height of the cross peaks assigned in ¹H-¹H-NOESY spectra. Automated NOESY based structure calculation mode of structure calculation software CYANA (with standard script CALC.cya) was used. This script automatically converts data heights into distance bounds with maximum upper bound limit of 6Å.

The assigned sequence and the conformational restraints were used to calculate 200 conformers are calculated using the *CYANA version 3.0* (Guntert 2004).

2.9 Other biophysical techniques

2.9.1 UV spectroscopy for determination of protein concentration

The concentration of proteins in solution was determined, according to the Lambert-Beer law, by measuring the absorption at 280 nm (A_{280}).

$$A_{\lambda} = \epsilon_{\lambda} \cdot c \cdot l$$

where, A_{λ} : absorption at wavelength λ

ϵ_{λ} : molar extinction coefficient at wavelength λ in $\text{M}^{-1}\text{cm}^{-1}$

c : protein concentration (M)

l : path length of cuvette in cm

ϵ_{λ} was derived from the protein sequence using the EXPASY web server (<http://web.expasy.org/protparam/>). ϵ_{λ} used for Tim23¹⁻⁹⁶, Tim21¹⁰³⁻³³⁵ and Tim50¹⁶⁴⁻³⁶¹ are 12666, 14440 and 34380 $\text{M}^{-1}\text{cm}^{-1}$ respectively.

2.9.2 Circular dichorism (CD) spectroscopy

The secondary structure of intermembrane space domains were determined by CD spectroscopy using a Photophysics Chirascan spectrometer. The Far UV-CD spectra (190-250 nm) were collected at room temperature with various protein concentrations ranging from 5 μ M to 25 μ M in 20mM Na-phosphate buffer pH 7.2 and 50mM NaCl. Data were expressed as difference in absorbance (ϵ) or circular dichorism ($\Delta\epsilon$) per molar concentration of protein. The data was plotted along the measured wavelength to qualitatively deduce the secondary structural elements.

2.9.3 Size exclusion chromatography

An Äkta chromatography systems (Äkta purifier; GE Healthcare) equipped with a Superdex 75 (16/60) preparative grade gel filtration column (with column volume of 120ml) at a flow rate of 1ml/min was used for establishing the oligomeric state of Tim21ims. An analytical grade HR-Superdex 75, (column volume 24ml) gel filtration column with flow rate 0.5ml/min was used for monitoring the purity of samples for Tim50ims.

3 Results

3.1 Intermembrane space domains (ims) of mitochondrial translocases

With a great wealth of available biochemical data and the aim to deduce the interaction sites among intermembrane space domains of mitochondrial translocases, we cloned, expressed and purified the intermembrane space domains of Tom22, Tim23, Tim21 and Tim50 to reconstitute the *in vivo* interactions in our studies. The *in vitro* direct interaction studies among the aforementioned intermembrane space domains is the first step towards the understanding of the complex *in vivo* scenario, where dynamic subunit-subunit interactions are modulated as a function of incoming presequence, membrane potential and additional cofactors that determines the molecular mechanism for mitochondrial preprotein import .

3.1.1 Purification of intermembrane space domains

One of the prerequisite for the NMR studies of these interactions is the production of pure and sufficient quantities of isotopically enriched proteins. In order to achieve this, we have expressed the various intermembrane space domains as soluble proteins in *E.coli* using fusion tags such as GST and Z₂ (Smith and Johnson 1988; Bogomolovas et al. 2009). The expression and purification of the intermembrane space domains of Tom22, Tim21, Tim23 and Tim50 have been described in detail in Appendix (section1.1). The purity for all the intermembrane space domains was found to be higher than 95 % (Figure 16). Expression of each intermembrane space domain yielded at least one 0.5ml sample at 0.5mM concentration per litre of M9 culture.

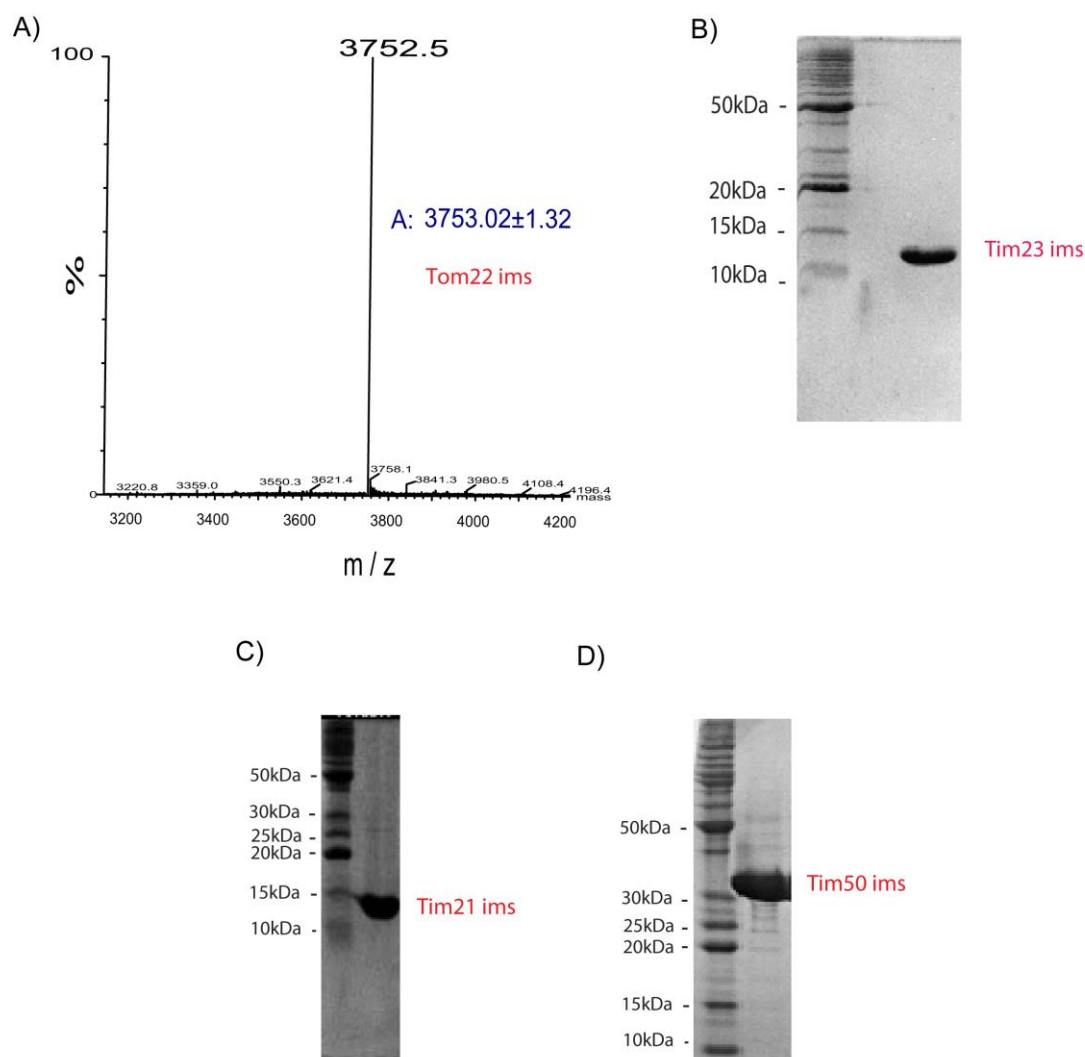


Figure 16: **Homogeneously purified intermembrane space domains used in this study:** (A) ESI MS of Tom22ims showing expected mass. 15% SDS-PAGE gel corresponds to (B) Tim23ims, (C) Tim21ims and (D) Tim50ims highlighting the purity of each sample.

Further characterization of isolated intermembrane space domains is described below.

3.1.2 Structural characterization of IMS domains

The solution structural properties along with the physiochemical properties determine how the interactions are mediated in solution. Hence, it is important to understand the solution structural properties of individual proteins prior to their

interaction studies. Upon successful purification of five IMS domains of two mitochondrial translocases, their solution properties were analyzed using a range of biophysical techniques that includes circular dichorism (CD), dynamic light scattering, size exclusion chromatography and ^1H NMR spectroscopy.

Circular Dichorism and ^1H NMR spectroscopy were used to determine the secondary structure content and folded state of the protein whereas SEC aided to investigate the oligomeric status of protein. Correlation times was additionally determined by $\text{H}^{\text{N}} \text{T}_2$ obtained from ^1H echo type 1D one-one experiment for globular proteins (Tim21ims and Tim50ims) whereas for disordered proteins hydrodynamic radius was calculated using DOSY-NMR (Tim23). The findings obtained using these techniques for the various intermembrane space domains are summarized in Table 12.

Table 12: Observed parameters for characterization of intermembrane space domains of mitochondrial translocases in solution using size exclusion chromatography and 1D- ^1H -NMR. The correlation time for Tim21ims and Tim50ims was deduced using ^1H 1D one-one echo at 1mM and 200 μM , respectively. The hydrodynamic radius of Tim23ims was measured by DOSY-NMR. (ND means not determined).

| Construct (ims) | pI | M.W. (kDa) | TauC (ns) /R _h (nm) | No. of peaks in SEC | Most Up field signal(ppm) | H ^N Range (ppm) |
|--------------------|------|---------------|-----------------------------------|------------------------|------------------------------|-------------------------------|
| Tom22 | 3.95 | 3.6 | ND | ND | 0.7 | 7.8-8.4 |
| Tim21 | 9.59 | 14.4 | 8.5ns | 2 | 0.07 | 6.5-9.9 |
| Tim23 | 4.35 | 10.28 | 2.6nm | 1 | 0.4 | 7.6-8.4 |
| Tim50 | 5.69 | 23.28 | 14ns | 1 | -0.5 | 6.4-9.8 |
| Tom40 | 3.49 | 2.849 | ND | ND | 0.48 | 7.8-8.6 |

The folded state of the intermembrane space domains under investigation were assessed by the analyses of two main regions i.e. side chain aliphatic region and amide proton of 1D- ^1H NMR spectrum. One dimensional ^1H spectrum of Tom22ims and Tim23ims exhibited an amide-proton amide-nitrogen chemical shift dispersion of 0.6 and 0.8 ppm respectively, whereas Tim21ims and Tim50ims, each exhibited chemical shift dispersion range of more than 3 ppm. Additionally, most up field aliphatic resonance for the Tom22ims and Tim23ims were close to 0.4 ppm in contrast to Tim21ims and Tim50ims that have the most up field signal close to 0 ppm and up field -0.5 ppm respectively.

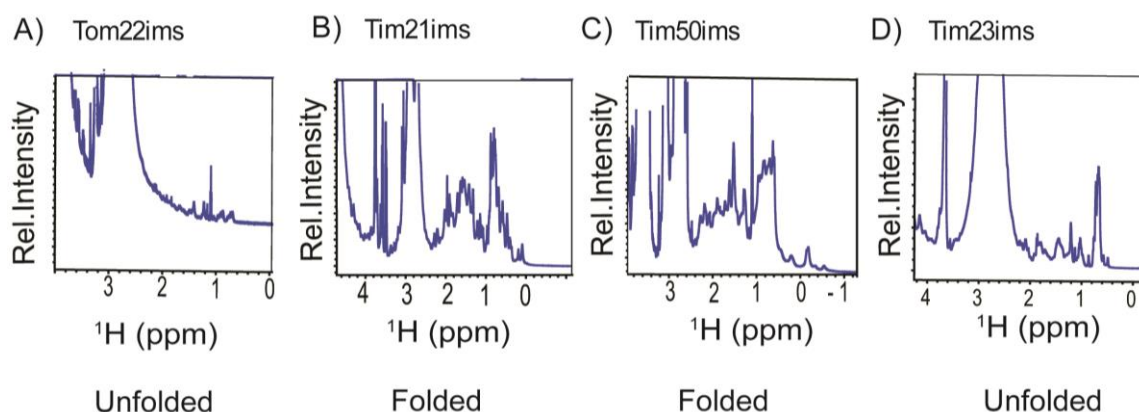


Figure 17: Aliphatic region of 1D- ^1H NMR spectra for intermembrane space domains of mitochondrial translocases. (A) Tom22ims, (B) Tim21ims, (C) Tim50ims and (D) Tim23ims with most up field (below 0.4ppm) resonance indicating the presence or absence of protein core, a characteristic for folded and unfolded protein respectively.

The low amide chemical shift dispersion range of Tom22ims and Tim23ims indicate that the amide protons of most their residues experience similar chemical environment (solvent exposed), which is characteristic of disordered proteins. The lack of protein core in Tim23ims and Tom22ims is also reflected by high up field (0.4ppm) aliphatic signals. On the other hand, large chemical shift dispersion was observed for Tim21ims and Tim50ims as expected for a folded protein wherein the amide protons of the core residues will differ from the surface exposed residues. Furthermore, the pattern of H-bond (secondary structure) also provides diversity in chemical shift for folded proteins.

The line width of the resonances in 1D- ^1H NMR spectra was as expected proportional to the size of construct with the Tom22ims and Tim50ims exhibiting narrowest and broadest line widths, respectively. Additional processes like chemical exchange and oligomerisation might also contribute to the observed line broadening effects which could not be accessed accurately with 1D- ^1H NMR and required further characterization of these domains.

3.1.3 Tim23ims as intrinsically disordered monomer

Tim23¹⁻⁹⁶ (or Tim23ims) eluted as a single peak from the SEC column and possess constant hydrodynamic radii as a function of increasing concentration (data not shown from DOSY-NMR) indicating that it exists as a monomer in solution. The CD

spectroscopy was used to measure the secondary structure content of Tim23ims. The obtained CD spectrum of Tim23ims showed a sharp negative band around 196 nm and very low ellipticity above 210 nm, which is often seen in case of disordered proteins (Figure 18, A). Furthermore, ^1H - ^{15}N HSQC of Tim23ims exhibited poor amide proton chemical shift dispersion and sharp resonances which are characteristic of rapidly interconverting conformers of the disordered proteins (Figure 18, B).

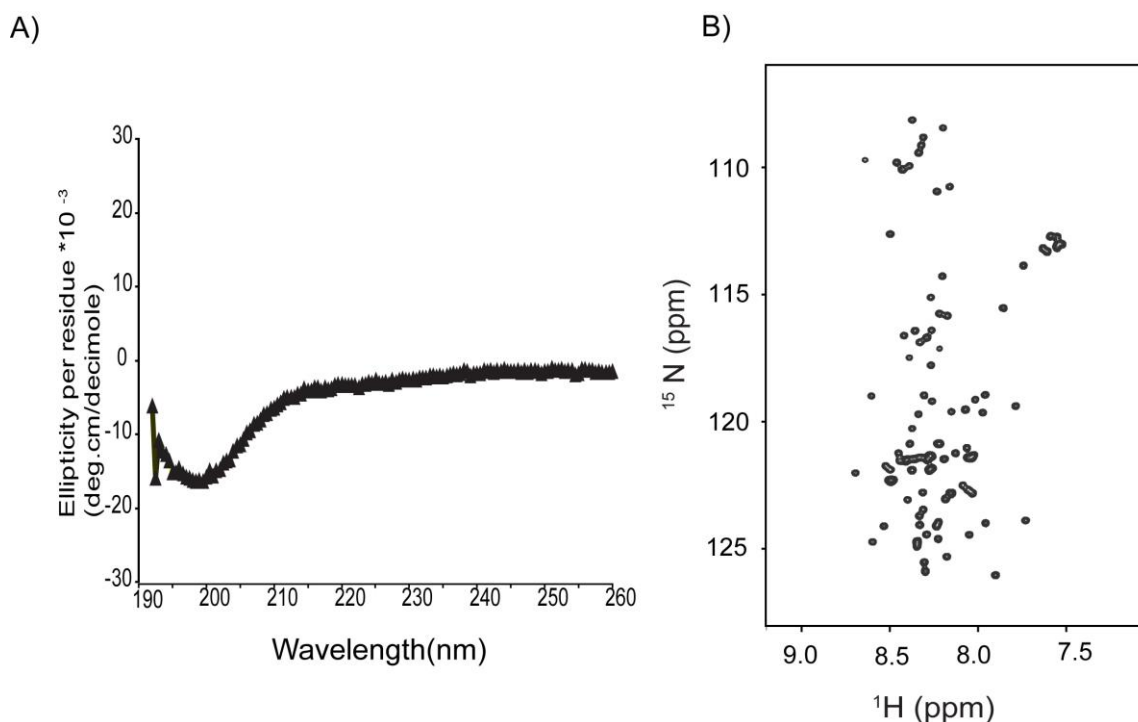


Figure 18: **Tim23ims is intrinsically disordered.** (A) Near-UV circular dichroism spectrum in phosphate buffer at 10 μM concentration. (B) ^1H - ^{15}N -HSQC spectra of Tim23ims show narrow chemical shift dispersion and sharp resonances in amide proton region.

Intrinsically disordered proteins are known to possess transient long range interactions and residual secondary structural elements that may populate differentially under different physiological conditions (Bertoncini et al. 2005; Mukrasch et al. 2009). To investigate the presence of such residual secondary structural elements at residue level for Tim23 $^{1-96}$ secondary carbon chemical shifts, ^1H - ^{15}N Het NOE and ^1H - ^{15}N RDCs were measured.

The residue specific secondary structure propensity of Tim23ims as exhibited from average secondary C^α and C' chemical shift showed that most of the residues possess values within $\pm 0.5\text{ppm}$ than that of the random coil chemical shifts (Figure 19, A). Smaller deviations in secondary chemical shift values in Tim23ims from that of

random coil, indicate the lack of rigid secondary structural elements *i.e.* α -helices or β -sheets. The four or more consecutive positive values indicate the propensity for the helices. From Figure 19A, it is evident that residues 73-87, which have stretch of positive average values could transiently populate the helix.

The transient nature of secondary elements in terms of timescales of backbone motion could be understood by ^1H - ^{15}N HetNOE, which measures the amplitude of motion of HN bond vector at picosecond to nanosecond timescale. ^1H - ^{15}N HetNOE values are directly proportional to rigidity of backbone motion at previous said timescale *i.e.* higher positive value means more rigidity that would be expected for residues engaged in forming secondary structural elements. However, most of residues in Tim23imspossess average values ^1H - ^{15}N HetNOE below ± 0.25 implying that backbone of Tim23 contain high degree of motion at picosecond to nanosecond timescale (Figure 19, B). The higher negative ^1H - ^{15}N HetNOE values are expected for highly flexible residues, which is interestingly observed for region A17-D22 in addition to expected sites of N and C- terminus.

^1D (HN) RDCs that provide information about the average orientation of NH vector with respect to the external magnetic field showed small values along the sequence of Tim23ims (Figure 19,C). Thus, there exists small, yet, noteworthy conformational preference and dynamics in Tim23ims. Qualitatively, the sign of RDC for first 12 residues is opposite to the rest of polypeptide chain in Tim23ims. In addition, isolated positive values are seen for residues 7, 40, 53, 59 and 82 that are characteristic for turn formation (Mukrasch et al. 2009).

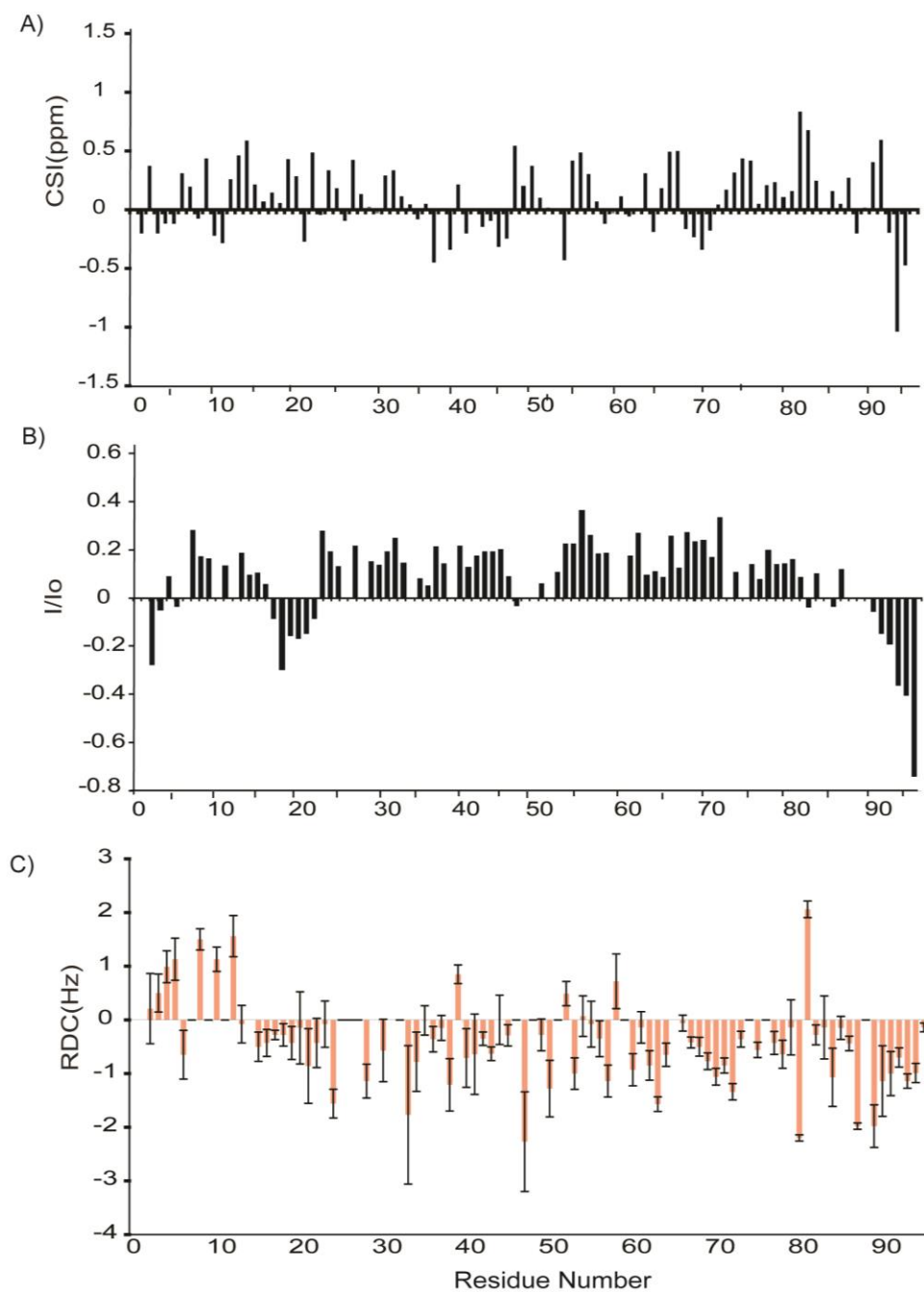


Figure 19: Structural properties of Tim23ims. (A) Averaged C^α/C' secondary chemical shifts (CSI) observed in Tim23ims as a function of residue number. Negative (positive) secondary chemical shifts spanning several residues indicate propensity for beta-strand (alpha-helix). (B) Steady-state heteronuclear ^{15}N - ^1H -NOEs and (C) Profile of $^1\text{D}(\text{HN})$ dipolar couplings in aligned in Pf2 bacteriophage of Tim23ims at 15°C . Different sign of residual dipolar coupling indicate different alignment of HN internuclear vector with respect to magnetic field. The error bars for dipolar couplings are based on Line width/Signal to Noise ratio.

3.1.3.1 Long Range interactions in disordered Tim23ims

Paramagnetic relaxation enhancement (PREs) provides important information about the transient tertiary interactions in unfolded and partially folded proteins (Lietzow et al. 2002). To probe the transient long range interactions in Tim23ims, a cysteine was engineered at position 11 using site directed mutagenesis (Tim23ims T11C). Tim23ims T11C mutant was covalently attached to MTSL (a nitroxide containing paramagnetic tag). The nitroxide label causes paramagnetic dipolar relaxation of nuclear spins typically, amide protons and thus, the residues close to the site of paramagnetic tag will experience selective line broadening in the NMR spectrum. The broadened H^N can be correlated to the average distance of residue from the site of paramagnetic tag up to 25 Å. The paramagnetic relaxation enhancement (PRE) profile of Tim23-T11C tagged with MTSL is obtained from the ratio of intensity cross peaks in 1H - ^{15}N -HSQC spectra of Tim23ims T11C in paramagnetic (with MTSL) and diamagnetic state (with MTSL+DTT) (I_p/I_d). The PRE profile for residues (± 10 residues on either side of MTSL tag) close to T11 as expected shows the severe amide proton (H^N) broadening (Figure 20). The residues 25-40 and 40-56 have the average intensity broadening of 0.35 and 0.6 respectively, indicating presence of transient interaction in this region. Additionally, intensity broadening was observed for the C-terminal residues 90-94 suggesting that the Tim23ims in solution is not completely an extended polypeptide chain but possess the intramolecular long range contacts between its N and C-terminus. The residues such as N40, D54, L58 and L71 are well resolved in the spectra and show an average intensity attenuation value of 0.2. The residues 74-88 show intensity ratio (I_p/I_d) value close to unity. This implies that these residues lack transient interactions with N-terminus T11C.

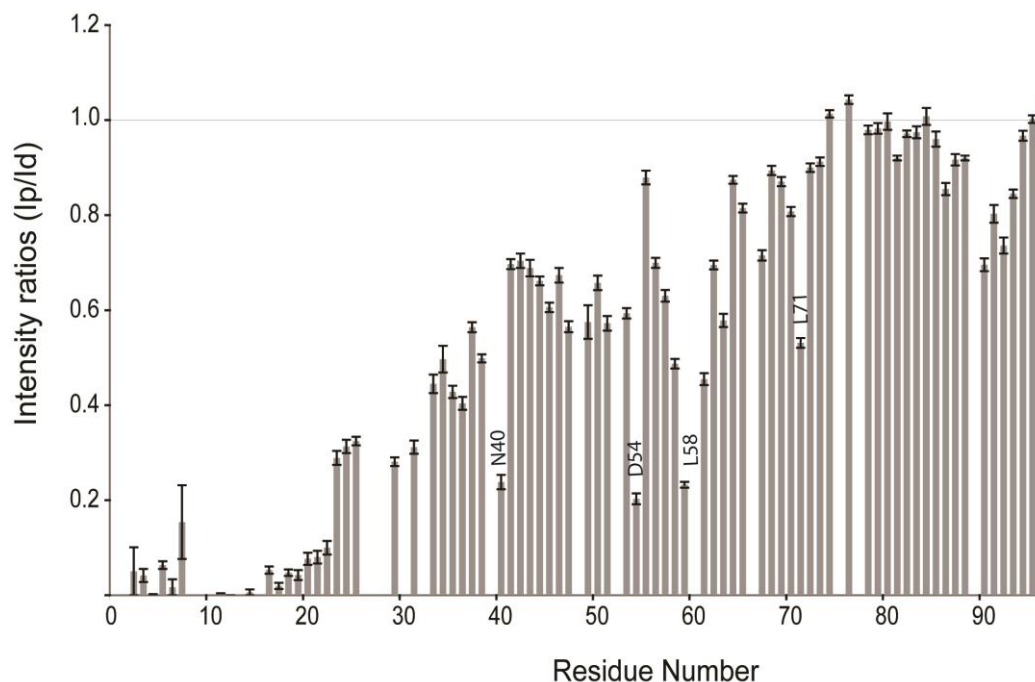


Figure 20: **Intramolecular long range interactions in Tim23ims.** PRE profile of Tim23ims T11C obtained with paramagnetic tag (MTSL) conjugated at position 11 (T11C) of Tim23ims. The intensity ratios are derived from signal intensity ^1H - ^{15}N HSQC spectra of Tim23ims T11C with MTSL paramagnetic (oxidized) and DTT added (reduced). The error bars are obtained based on the S/N values. Deviations from intensity ratio of 1 beyond residue 25 are indicative of presence long range interactions. Some resonances such as N40, D54, L58 and L71 that shows more severe broadening in comparison to its neighboring resonances are labeled.

3.1.4 Solution properties of Tim21ims

3.1.4.1 Association state of Tim21ims in solution

Size exclusion chromatography (SEC) was used to characterize the solution state of Tim21ims in terms of its shape and molecular weight. Notably, Tim21ims¹⁰³⁻²²⁵ contains one cysteine residue and in order to prevent intermolecular disulphide interactions size exclusion chromatography was performed under reducing conditions (5mM DTT). The SEC chromatogram of purified Tim21ims shows two peaks at elution volume maximum of 75ml (peak 1) and 85ml (peak2) respectively, which roughly correspond to the retention time of globular folded protein of 30kDa and 15kDa respectively (in comparison to S-75 standards) (Figure21, A). The SEC chromatogram of the reinjection (of pooled elution fractions) of peak 1 again showed a similar two

peak pattern with identical retention time as seen before (Figure 21, B). The identity of each eluted peak was further confirmed using ESI-MS (and 15% SDS-PAGE), which showed identical mass corresponding to monomeric Tim21ims. These results implied that Tim21ims exists in equilibrium as monomer and dimer in solution.

Furthermore, in order to characterize the structural differences between the monomeric and dimeric fractions of Tim21ims, the fingerprint spectrum of proteins *i.e.* ^1H - ^{15}N HSQC spectrum for each fraction was acquired. The overlay of ^1H - ^{15}N HSQC of Tim21ims for both the peaks (*i.e.* with elution volume 75ml, dimeric-Tim21ims and 85ml, monomeric-Tim21ims) at two different concentrations of 10 μM and 1mM respectively, were nearly identical in chemical shift position for most of resonances expect for the small chemical shift changes in resonances belonging to β -strand 2 and 3. The three histidine residues in β -strand 3 and a histidine residue (H221) at C-terminal end was also perturbed (Figure 21, C). The pH was kept constant for 10 μM and 1mM samples, additional control pH titration of Tim21ims was also performed to clarify the reason for the changes in CSP at two different concentrations of Tim21ims. Upon pH titration of Tim21ims, the β -strand 2, 3 and 4 were perturbed (Appendix Figure 5). However, the magnitude and direction of the chemical shift changes in Tim21ims as a function of pH are different. These results suggest that the concentration dependent changes in the chemical shift position of resonances of Tim21ims could be a characteristic of chemical exchange between different states of Tim21ims.

Additionally, correlation time of Tim21ims at concentration 1mM was found to be ~8.6ns that roughly corresponds to the size of the monomeric Tim21ims (15kDa). The two results for expected dimeric fraction from SEC, (a) the similar chemical shift position of resonances in ^1H - ^{15}N HSQC spectra and (b) molecular weight similar to monomeric Tim21ims indicate that the SEC eluted dimeric fraction of Tim21ims is not stable and rapidly dissociates to monomeric Tim21ims. Thus, effectively low populations of dimeric-Tim21ims are present in both the samples (10 μM and 1mM) and could not be observed by NMR. The NMR observable (chemical shift and the intensity of each residue) of Tim21ims in ^1H - ^{15}N HSQC spectra for both the SEC eluted fractions is an average weighted population of two states of Tim21ims with major contribution from a monomeric state up to the concentration of 1mM in Tim21ims.

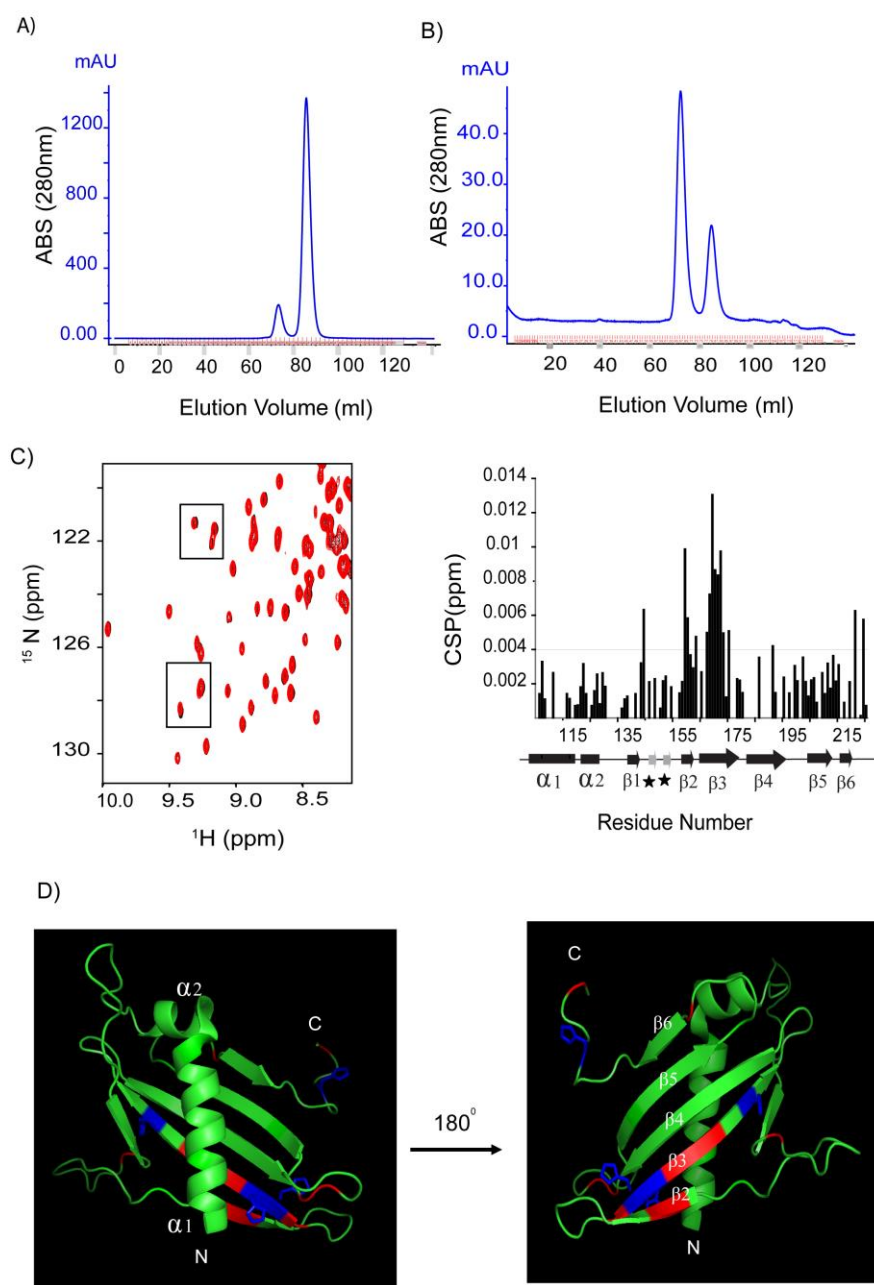


Figure 21: Tim21ims exhibits monomer-dimer equilibrium in solution. (A) SEC chromatogram of purified Tim21ims showing two peaks at elution volume of 75 and 85 ml. (B) SEC chromatogram of elution peak at 75ml obtained upon its reinjection. (C) Left: Overlay of a section of ^1H - ^{15}N HSQC of Tim21ims at different concentration 10 μM (red) from the peak eluted at 75ml and 1mM (black) from the peak eluted maximum at 85ml in SEC, acquired with 128 and 16 transients respectively at 298K. The resonances that showed chemical shift perturbations are highlighted by black boxes. Right: Average chemical shift perturbation (CSP) plot for these two ^1H - ^{15}N HSQC's as a function of the sequence and secondary structural elements of Tim21ims. Two short β -strands represented by asterisk (\square) are absent in NMR structure, details in Figure 24. (D) The cartoon representation of Tim21ims (NMR model 1) is mapped with the residues showing a magnitude of CSP greater than 0.004 ppm. All perturbed histidine residues are colored as blue while other residues are colored in red. The secondary structural elements and N and C terminus of Tim21im are labeled as N and C respectively.

3.1.4.2 Solution structure of monomeric Tim21ims

Despite the observance of significant amide-proton chemical shift dispersion (Figure 21C), the normalized intensities (w.r.t last residue) of the peaks in the ^1H - ^{15}N HSQC spectrum of Tim21ims were not uniform (Figure 22). Some of the residues (such as 130, 131, 135) show higher intensity ratio than average intensity ratio, whereas others show low intensity ratio (such as 141, 142, 144, 149, 151) indicating the presence of variable motion at different time scale in NMR. Loops and the highly flexible parts of proteins show higher intensity (dynamics at ps-ns time scale) whereas the regions undergoing chemical exchange are less intense (dynamics at μs -ms time scale). The residues in the two short β^* strands are less intense. This interesting observation of non-uniform intensity in Tim21ims led us to probe the chemical exchange and local dynamics at the various sites in Tim21ims.

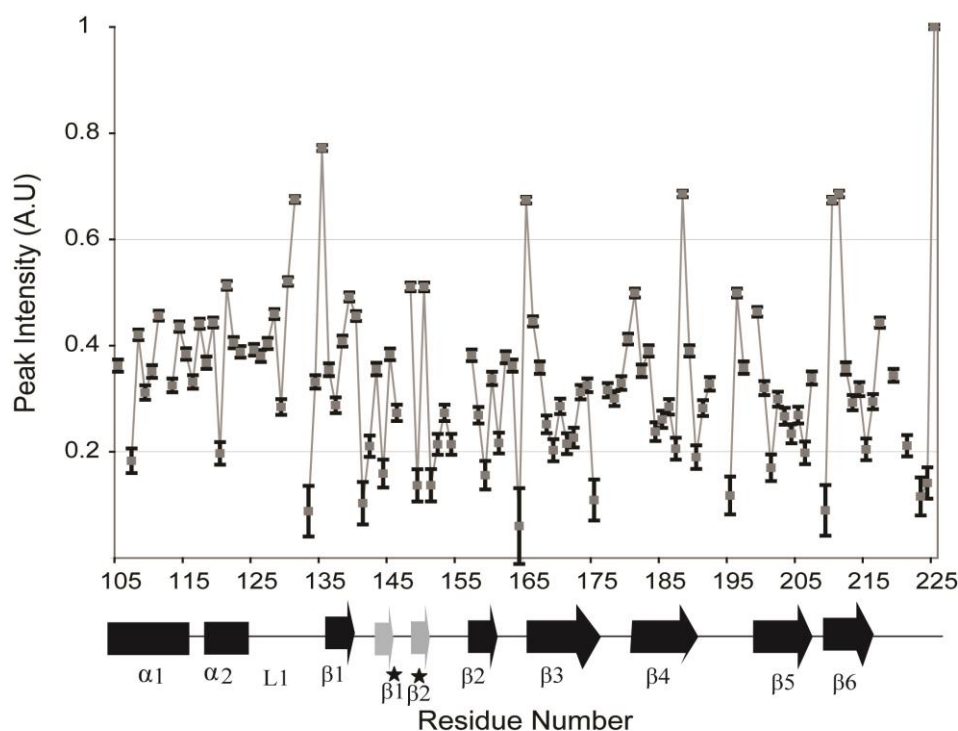


Figure 22: Signal intensity variation in ^1H - ^{15}N HSQC of Tim21ims as a function its residue number. The residues belonging to region 130-133 and 144-155 shows the significant variation along with isolated residues in loops connecting strands in β -sheet. Last residue was used to normalize the intensity for all resonances, the error bars are based on the S/N ratio as described in materials and methods. The secondary structural elements of Tim21ims are shown at the bottom. The two β^* -strands in grey are present in X-ray structure and are absent in NMR structure of Tim21ims (described in section 3.1.4.2).

Hence, ^{15}N - R_2 values and ^1H - ^{15}N Het NOE were measured for Tim21ims in solution. The ^1H - ^{15}N HetNOE intensity ratio of residue is inversely proportional to the degree of fluctuations of (NH) amide bond vector at picosecond to nanosecond time scale. In Tim21ims, most of the residues have an intensity ratio above 0.65, which is expected for an ordered part of protein i.e. secondary structures and rigid loops. Some of residues corresponding to loop1 in Tim21 (residues 126-138) have ^1H - ^{15}N -HetNOE values lower than 0.4 indicating it to be flexible at this time scale. Additionally, residues (145-155) belonging to $\beta 1^*$ and $\beta 2^*$ -strands are mostly missing in the analysis due to resonance overlap and extremely weak resonance in reference spectrum. However, this region has two well isolated residues 148 and 153 that have low ^1H - ^{15}N -HetNOE value, suggesting $\beta 1^*$ and $\beta 2^*$ -strands are flexible (Figure 23, A).

Thereafter, ^{15}N R_2 values of Tim21ims were measured to probe the microsecond to millisecond chemical exchange. The measured ^{15}N R_2 values are cumulative sum of the intrinsic ^{15}N R_2 and the R_{ex} (contribution due to chemical exchange). The residues corresponding to lower ^{15}N - ^1H -Het NOE values in first loop (residues 126-138) also possess more than 40 % lower value in comparison to average R_2 confirming it to be highly flexible region. On the other hand, the second region of Tim21ims deviating from average R_2 values have almost double or 50 % higher R_2 value corresponds to residues β -strands 2, 3 and 4 in NMR structure of Tim21ims. The residues 159, 160, 168-170, 172, 174 and 186-190 are most prominent with higher R_2 values in these strands. Such higher R_2 values are indicative of slow exchange at millisecond to microsecond time scale that could arise either due to internal breathing of protein or possibly the exchange between the monomeric and dimeric Tim21ims (Figure 23, B).

Taken together, these results suggested that in solution, the conformation of Tim21ims could adopt different structure under solution conditions with the major changes localized to aforementioned regions. To this end, the monomer structure of Tim21ims was determined using NMR spectroscopy by Lukasz Jaremko and Mariusz Jaremko in our group.

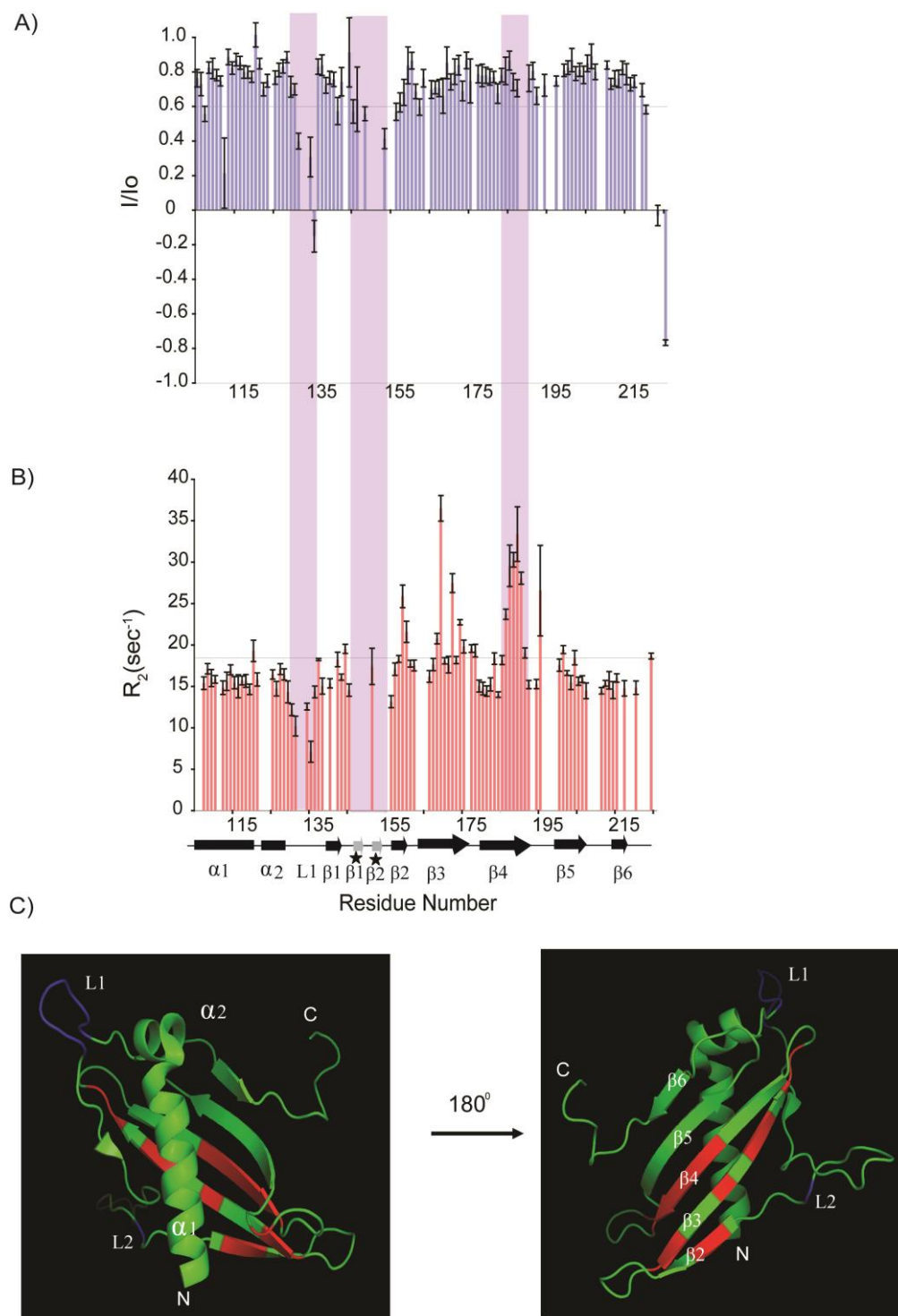


Figure 23: ^{15}N relaxation of apo Tim21ims in solution recorded in 700MHz at 298K. The values of ^{15}N - ^1H Het NOE (A) and ^{15}N R_2 (B) are plotted as a function of residue number. The non-proline residues with either overlapping peaks or peaks with low S/N were omitted from the analysis and they appear as blank in plot. The lower panel shows the schematic representation of secondary structural elements in Tim21ims. (C) Two views of the cartoon representation Tim21ims (NMR model 1) mapped with the residues exhibiting ^{15}N R_2 value higher than 19 sec^{-1} in red and lower than 12 sec^{-1} in blue color.

The backbone RMSD for the structured regions of 20 low energy conformers of Tim21_{ims} determined by NMR with respect to mean structure was found to be 0.47 ± 0.09 Å indicating the NMR structures converge well. The comparison between NMR structure and X-ray structure of Tim21_{ims} revealed that two α -helices are conserved in both the structures. However, two small β -strands (out of 8 observed in X-ray structure i.e. $\beta 1^*$ and $\beta 2^*$) involving residues 144-146 and 151-153 are not well defined in NMR structure (Figure 24). This region has a higher RMSD of value of 2.40 ± 0.88 Å (marked with label 2 in Figure 25, B and C) in 20 low energy conformers in comparison to mean NMR structure. In addition, 20 low energy conformers have higher backbone RMSD as expected for the N-terminal (additional residues due to cloning exigencies) and C-terminal residues (marked with label 4 in Figure 25, B and C) as 2.01 ± 0.66 Å. Indeed, residues 144-146 and 151-153 ($\beta 1^*$ and $\beta 2^*$ strand) exhibits low signal intensity in ^1H - ^{15}N -HSQC spectra as well as low ^1H - ^{15}N HetNOE values in comparison to other parts of protein. Additionally, residues in region 146-156 ($\beta 1^*$ and $\beta 2^*$) show higher B-factors and possess crystal contacts with the C-terminal residues that lead to multiple conformations of this region in crystal and thus the uncertainty in their position i.e. static disorder in B factor.

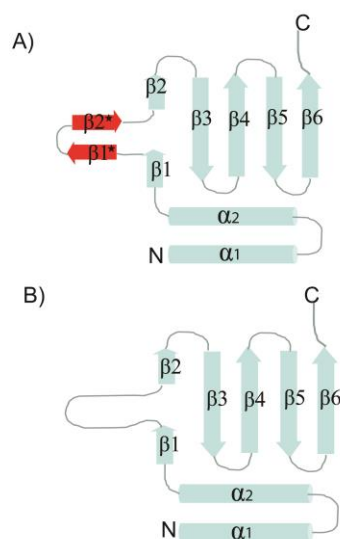


Figure 24: Comparison of secondary structure elements present in X-ray structure (A) and NMR structure (B) of Tim21_{ims} by the schematic representation of Tim21_{ims} fold. The secondary structure elements for X-ray and NMR structure of Tim21_{ims} were defined using DSSP program (<http://www.cmbi.ru.nl/dssp.html>). Two β -strands ($\beta 1^*$ and $\beta 2^*$) highlighted in red are not well defined in ensemble of NMR structure.

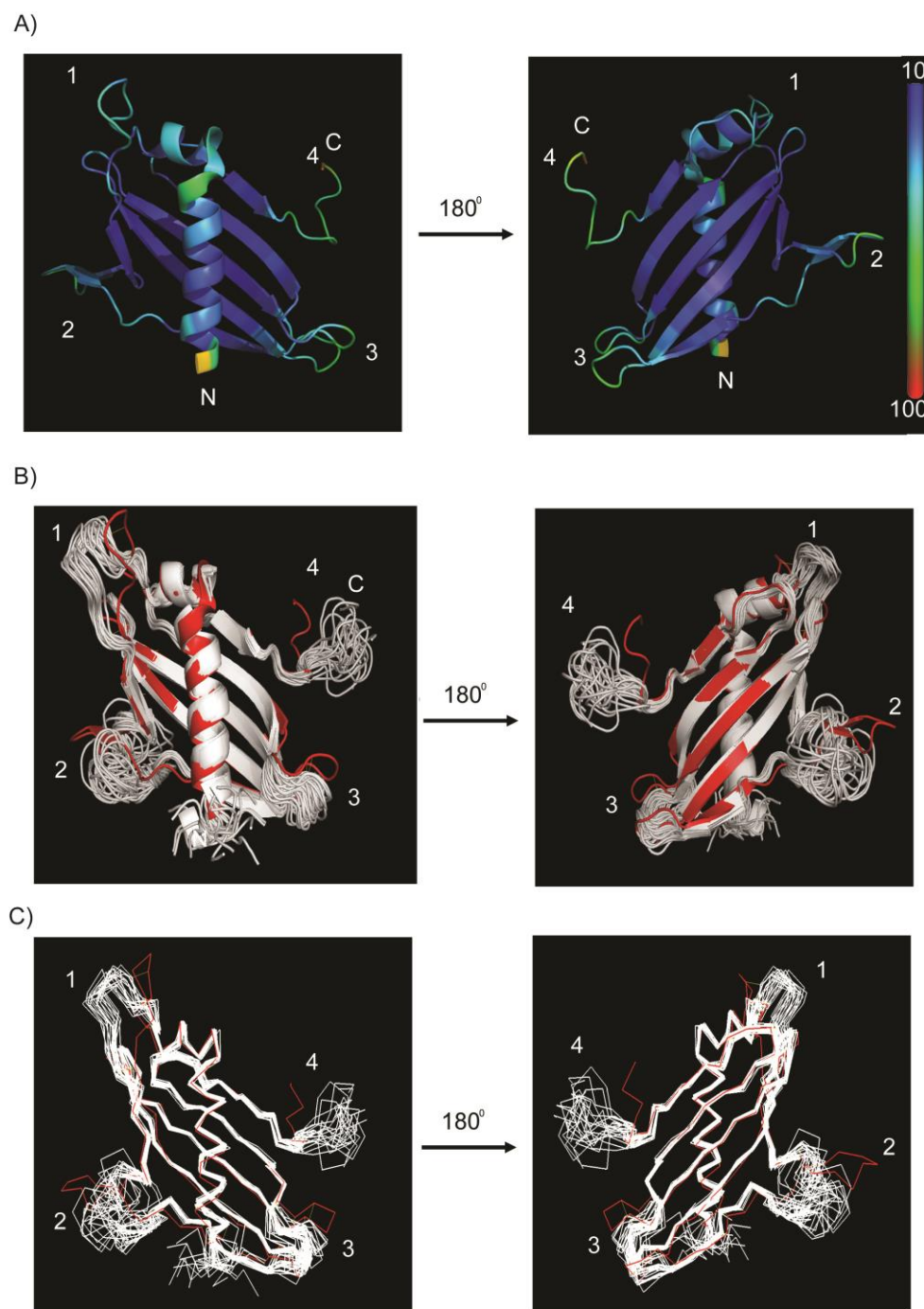


Figure 25: Structure of Tim21ims differs in solution:(A) Two views of cartoon representation of Tim21ims generated from crystal structure (PDB id:2CIU) and is color coded according to the B-factors using PYMOL with blue, green and red exhibiting gradient of B-factor from 0-95. Cartoon (B) and Ribbon (C) representation of superposition of 20 lowest energy NMR conformers (white) with X-ray structure of Tim21ims (red). The differences in NMR conformers and X-ray structure involve regions (1) Loop involving residues C128-L138, (2) residues 145-154 involving $\beta 1^*$ and $\beta 2^*$ strand, (3) Loop connecting β -strand 2 and 3, (4) C-terminal residues 221-225.

3.1.5 Tim50¹⁶⁴⁻³⁶¹ ims is folded in solution

The SEC chromatogram of purified Tim50¹⁶⁴⁻³⁶¹ ims showed a single peak (Appendix Figure 4, C) indicating a homogeneously purified sample free from higher order aggregates. The CD spectrum of Tim50¹⁶⁴⁻³⁶¹ ims (Figure 26, A), showed a negative signal at 207 and 222nm. This is characteristic for the presence of helical content in proteins. The lack of prominent CD signal at 195nm and 215nm indicates that disordered parts are dominantly absent. However, it is often difficult to predict accurately the secondary structure content by CD spectrum, which is mainly dominated by α -helices.

The ¹H-¹⁵N TROSY-HSQC spectra of Tim50¹⁶⁴⁻³⁶¹ ims displayed approximately 127 well dispersed resonances in contrast to the expected 178 from this 198 residue construct with 19 prolines (Figure 26, B). As expected for well folded protein, ¹H-¹⁵N TROSY-HSQC display broad amide dispersion encompassing characteristic regions of beta strands and alpha helices. The lack of 40 peaks in, ¹H-¹⁵N TROSY-HSQC spectrum could be attributed to chemical exchange occurring at intermediate NMR timescale for these peaks. This chemical exchange could be conformational or exchange with water for solvent accessible regions corresponding to loops or other flexible parts of Tim50ims.

The detailed inspection of X-Ray structure (PDB id: 3QLE) of Tim50¹⁷⁶⁻³⁶¹ ims corresponding to Tim50¹⁶⁴⁻³⁶¹ ims construct (that lacks an N-terminal 12 residues), highlights nearly thirty residues with higher B factors indicating that these resonances might be missing ¹H-¹⁵N TROSY-HSQC spectra of Tim50¹⁶⁴⁻³⁶¹ if they exhibits chemical exchange at the slower NMR timescale (μ s-millisecond) (Qian et al. 2011). Interestingly, an alpha helix (residues 241-251) and beta hairpin loop (residues 205-215) showed higher B factors in addition to expected loop region (Figure 26C). The absence of concentration dependent changes in the ¹H-¹⁵N TROSY-HSQC spectrum of Tim50¹⁶⁴⁻³⁶¹ ims (concentration ranging from 50 μ M to 350 μ M), indicate the absence of oligomerisation of Tim50¹⁶⁴⁻³⁶¹ims.

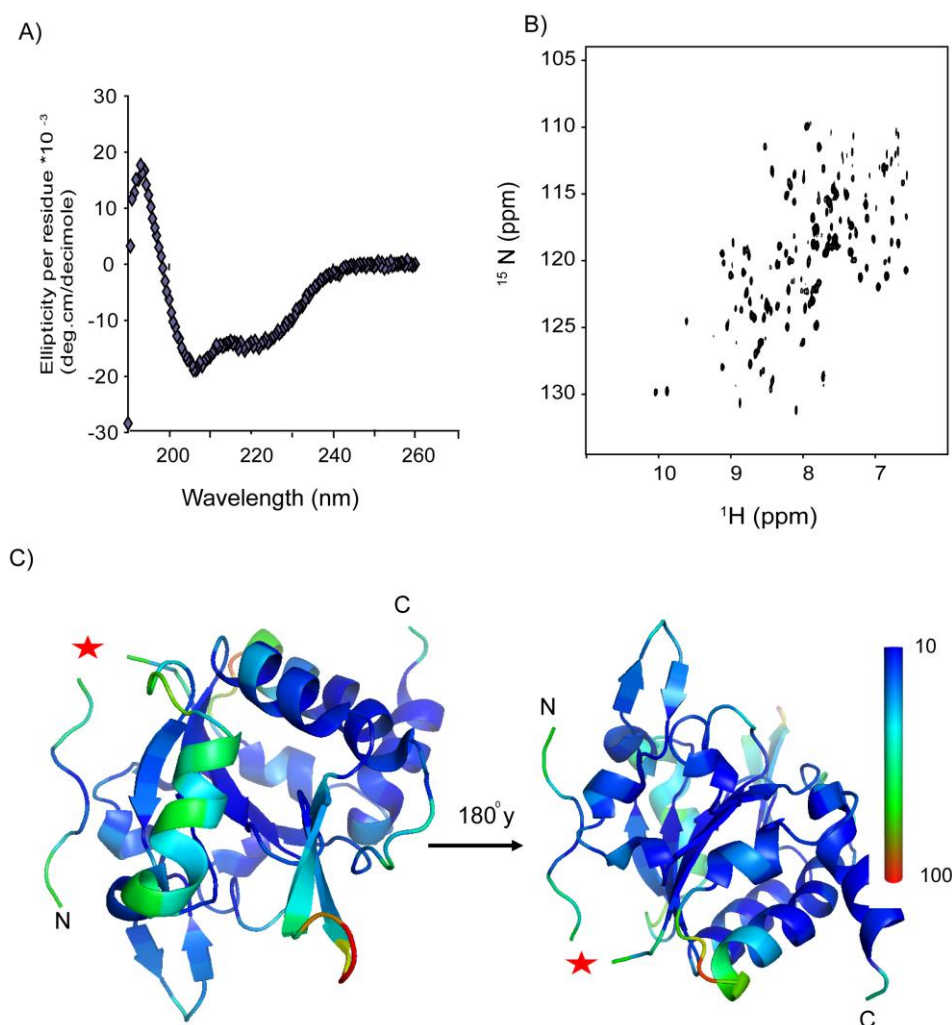


Figure 26: **Tim50ims¹⁶⁴⁻³⁶¹ is folded in solution:** (A) Near-UV circular dichroism spectrum in phosphate buffer at 5 μ M concentration. (B) ^1H - ^{15}N TROSY-HSQC of 80 μ M Tim50ims (164-361) at 310 K acquired at 900 MHz ^1H frequency. (C) Two views of cartoon representation of Tim50ims (164-361) of crystal structure (PDB id: 3QLE). The cartoon representation is color coded by the gradient of the B factor of 10-100 with blue, green and red highlighting the abundance of regions with higher B factors. N- and C-terminus are indicated. The red star represents the fitting artifact in model building using the molecular replacement method. At left, an α -helix (241-251) and β -hairpin (205-215) face towards the reader.

3.1.6 Tom22ims is largely unstructured in solution

The CD spectroscopy was used to measure the secondary structure content of Tom22ims. The obtained CD spectrum of Tom22ims showed a sharp negative band around 198 nm and very low ellipticity above 210 nm, indicating it to lack stable secondary structure elements (Figure 27A). Furthermore, ^1H - ^{15}N HSQC of Tom22ims

exhibited poor amide proton chemical shift dispersion and sharp resonances indicative of its disordered state in solution. Similar to Tim23ims, the secondary C^α chemical shift index was calculated to probe the presence of transient secondary elements in Tom22ims. The low positive secondary C^α chemical shift values for residues 120-127 are indicative of transient helix in this region (Figure 27C).

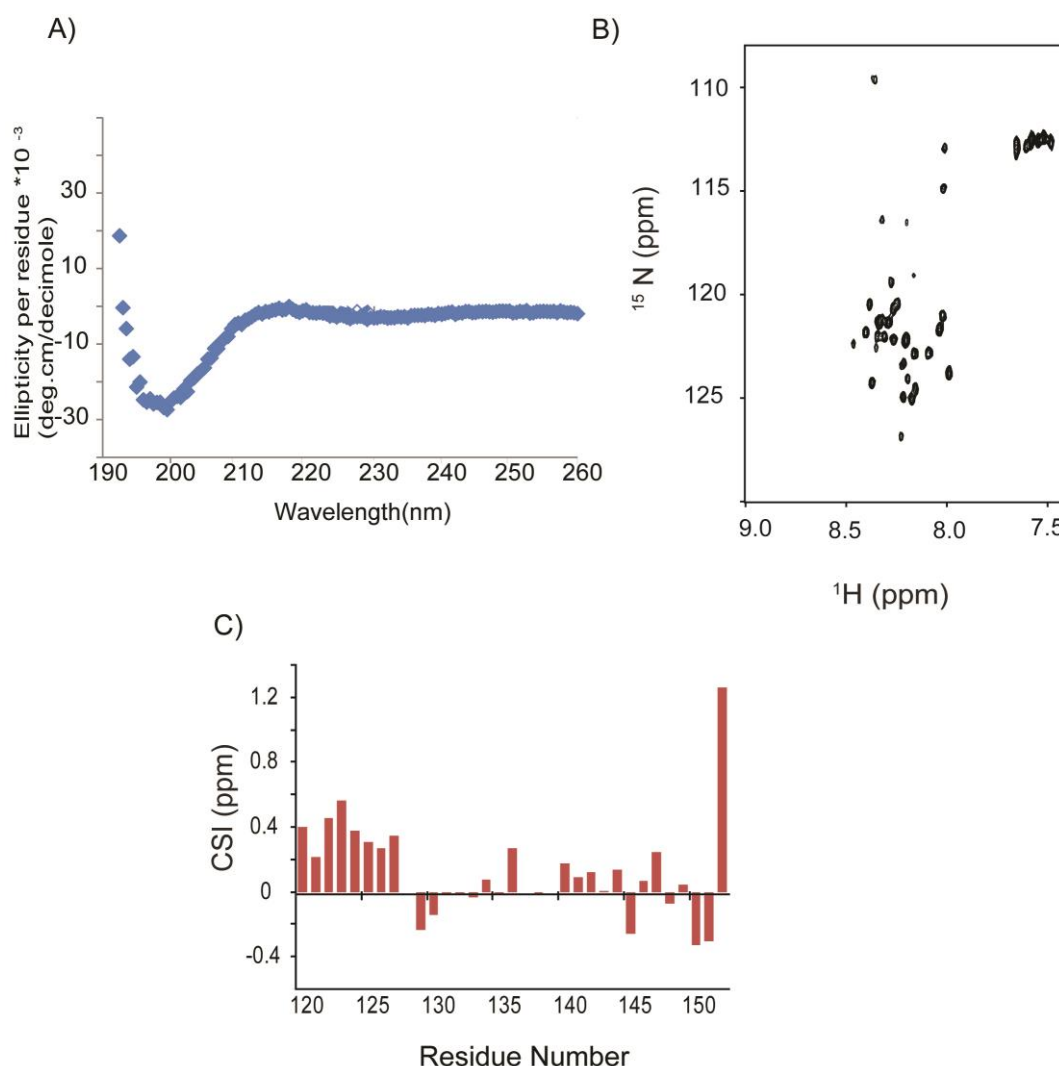


Figure 27: Tom22ims is intrinsically disordered. (A) Near-UV circular dichroism spectrum in phosphate buffer at concentration of 0.3mg/ml. (B) ^1H - ^{15}N -HSQC spectrum showing narrow amide chemical shift dispersion and sharp resonances. (C) C^α secondary chemical shifts (CSI) observed in Tom22ims as a function of residue number. Positive secondary chemical shifts spanning several residues indicate propensity for α -helix.

3.2 Presequence binding sites in mitochondrial intermembrane space domains

Preproteins targeted to mitochondrial matrix possess N-terminal presequence that interacts with series of receptors along its way to the mitochondrial matrix (Pfanner 2000). Many of the subunits of the outer mitochondrial membrane translocase such as Tom40, Tom22, Tom5, Tom7 and the inner mitochondrial membrane translocase such as Tim23, Tim50, Tim17 and Tim21 possess intermembrane space domains.

Of these, *in vivo* crosslinks of the presequence with the intermembrane space domains of Tom22, Tim23, Tim50 has been well studied (Bauer et al. 1996; Moczko et al. 1997; Geissler et al. 2002; Yamamoto et al. 2002; Marom et al. 2011). The direct binding of the presequence with the purified intermembrane space domain of Tim23 and Tim50 has been shown recently during the course of this thesis (de la Cruz et al. 2010; Marom et al. 2011). However, information about the structure of presequence-ims complex explaining the nature of interaction and critical residues involved in the binding presequence with the intermembrane space domains (i.e. Tom22ims, Tim23ims and Tim50ims) still remains elusive. The direct interaction among the presequence and soluble intermembrane space domain of Tom22 and Tim21 is not established, so far.

With the aim of elucidating these interactions, we have studied the interactions involving a matrix targeting N-terminal presequence (rALDH) with the intermembrane space domains of Tim21, Tom22, Tim23 and Tim50 at atomic resolution to identify the potential receptor sites of presequence and to understand the molecular mechanism of presequence interaction with their receptor during mitochondrial matrix preprotein import.

3.2.1 Tim23ims interacts presequence through its C-terminus

Recombinant Tim23ims forms a presequence sensitive voltage channel in isolation and along with Tim50; Tim23 has been established as a putative receptor for the presequence (Bauer et al. 1996; Truscott et al. 2001; Jensen and Dunn 2002). However, the definitive receptor presequence binding site was not clear. In order to determine the direct binding of the Tim23ims with presequence rALDH, presequence of

CoxIV and Syn B2 were titrated and binding to Tim23ims was followed with ^{15}N - ^1H -HSQC spectra's by Laura de la Cruz.

^{15}N - ^1H -HSQC based titration of rALDH with Tim23ims was performed with an aim to saturate the Tim23ims with presequence and to accurately determine the binding affinity of presequence rALDH with Tim23ims. Upon stepwise addition of presequence, the residues of Tim23ims such as residues L71 and L78 undergoes progressive chemical shift changes and hence are involved in the binding with the presequence. On the other hand, residues such as A14 and E38 remained unaffected (Figure 28 A). The absence of line broadening effects and the non-observance of new peaks even at 100 fold molar excess of presequence rALDH indicates the complex exhibits fast exchange regime at NMR time scale and involves rapid association and dissociation of Tim23-presequence complex under these conditions. The chemical shift perturbation along the sequence of Tim23ims highlighted the residues experiencing changes in chemical shift and could be directly or indirectly involved in interaction as 58-64, 71-84 and 91-95 (Figure 28 C)

The main binding site for presequence was achieved by the saturating Tim23 with presequence till no further changes in ^{15}N - ^1H HSQC spectra were observed (at 100fold excess). The dissociation constant (K_d) and nature of fits were determined by plotting the observed chemical shift changes as a function of presequence concentration. The residues 71-84 experiencing largest chemical shift perturbation during titration fitted well to one site binding model with average K_d of $0.47 \pm 0.02\text{mM}$ (de la Cruz et al. 2010).

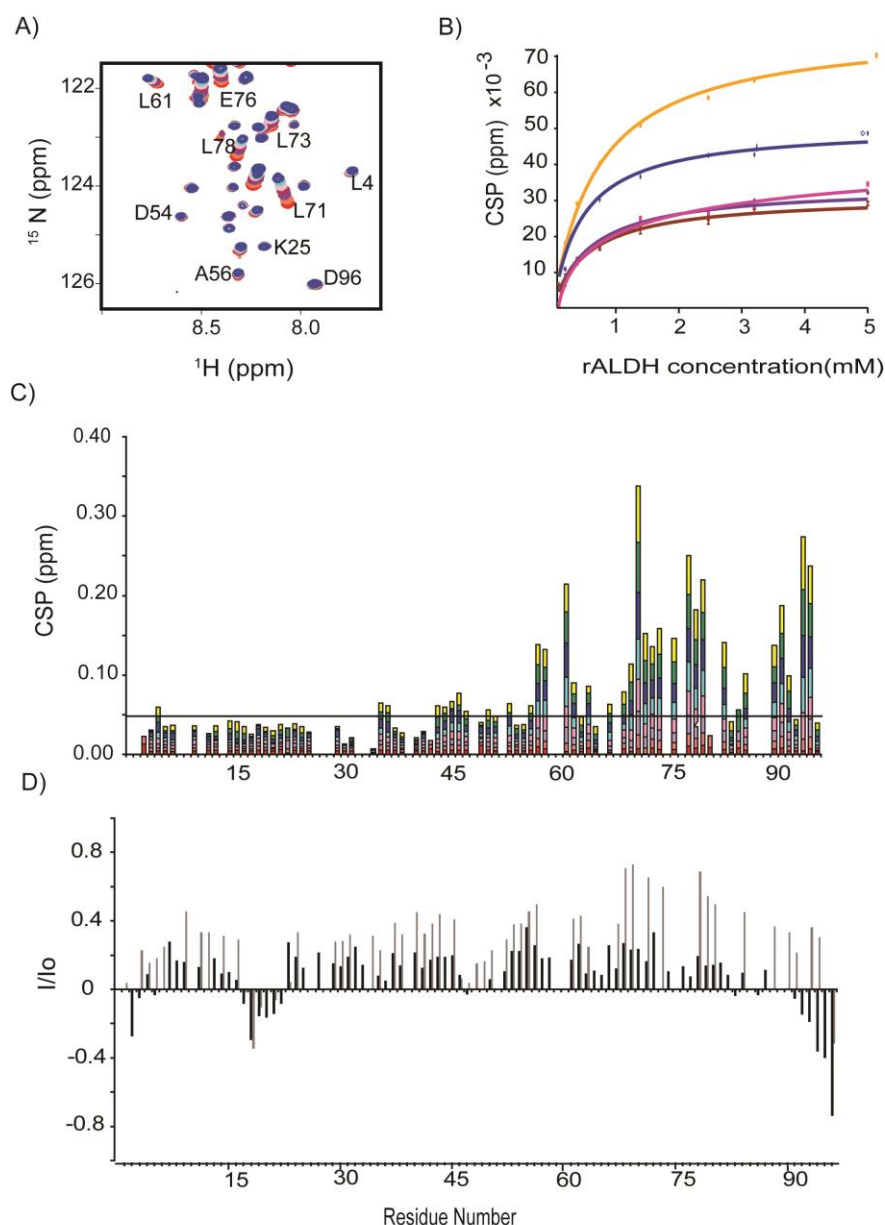


Figure 28: Binding of the presequence (rALDH) to Tim23ims. (A) Superposition of 2D ^1H - ^{15}N HSQC spectra of Tim23ims without rALDH (red) and at 32-fold excess of rALDH (blue). (B) Chemical shift perturbation (CSP) as a function of rALDH concentration for residues 71 (orange), 72 (pink), 74 (purple), 76 (blue), and 78 (brown). Solid lines show the binding curves obtained by fitting the chemical shift changes to a single binding-site model. A weighted average of the dissociation constant of 0.47 ± 0.02 mM was obtained. (C) Averaged ^1H , ^{15}N chemical shift deviation observed for backbone resonances of Tim23N in 2D ^1H - ^{15}N HSQC NMR spectra upon addition of rALDH at molar ratios (Tim23ims:rALDH) of 1:1 (red), 1:2 (orange), 1:4 (purple), 1:8 (pink), 1:16 (cyan), 1:32 (blue), 1:64 (green), and 1:100 (yellow). Black line represents average CSP of the least-shifting residues. (D) Steady state heteronuclear ^{15}N - ^1H NOEs observed in Tim23ims without (black) and with 64-fold excess of rALDH presequence (grey) coding the transient nature of Tim23ims-presequence complex.

3.2.2 Dynamics in Tim23ims-Presequence complex

To directly probe the dynamics of Tim23ims-presequence complex, steady state ^1H - ^{15}N HetNOE that report on backbone dynamics at picosecond to nanosecond time scale was measured for Tim23ims in the presence of 64 fold molar excess of rALDH. Even in the presence of presequence, the residues V17-Q21 of Tim23ims remains flexible to the same extent as in that of free form, whereas the region 71-78 shows maximum increase in ^1H - ^{15}N HetNOE to an average value of 0.6-0.7. Along with it, C-terminal residues 90-95 also get rigidified with an average ^1H - ^{15}N HetNOE value of 0.3 in comparison to their negative ^1H - ^{15}N HetNOE in the free form of Tim23 ims (Figure 28 D). These results indicate the rigidity in the binding site residues of Tim23 upon addition of presequence extend from main presequence binding region of 71-78 to 90-95 at C-terminus.

3.2.3 Effect of key mutations in Tim23ims-presequence complex

The most perturbed residue in Tim23ims upon presequence binding was found to be L71 and L78. The double mutant of Tim23 L71A/L78A shows significant growth and import defects in vivo (Tamura et al. 2009). Hence, we titrated presequence with L71A/L78A mutant to address whether the *in vivo* import defects arose from the reduced binding of Tim23ims with the incoming presequence. Similar to that of wild type Tim23, presequence titration was performed with Tim23 L71A/L78A. The Tim23 L71A/L78A shows similar chemical shift perturbations upon presequence binding as that of wild type. Thus, mutation at sites L71 and L78 of Tim23 ims does not completely abolish binding of presequence rather resulted in the reduction of binding affinity (Figure 29, B).

We further investigated important residues in presequence by mutating L17-L18 to Q17-Q18 and this mutant presequence also showed two fold reduced affinity similar to that of L71A/78A mutant of Tim23ims (Figure 29, B). Tim23ims has three charged residues in presequence binding site E74, E75 and E76. Instead of mutating these three residues, we mutated single charged residue R17 in the presequence rALDH. This mutation resulted in four-fold reduction in affinity (Figure 29, A).

Conjointly, these mutation studies suggest that the binding of presequence to Tim23ims is mediated by a combination of electrostatic and hydrophobic interactions.

Table 13 Tabular representation of binding affinities of various constructs of presequence-Tim23ims complex. wt means wild type. The peptide sequences of rALDH variants are listed in Appendix 1.3

| | Tim23 | Presequence | K _d |
|---|----------------|----------------|----------------|
| 1 | wt (1-96) | wt (1-22) | 0.45mM |
| 2 | L71AL78A(1-96) | wt (1-22) | 1.1mM |
| 3 | wt (1-96) | wt (1-11) | 3.5mM |
| 4 | wt (1-96) | wt (12-22) | 2mM |
| 5 | wt (1-96) | L18QL19Q(1-22) | 1mM |
| 6 | wt (1-96) | R17Q(1-22) | 2mM |

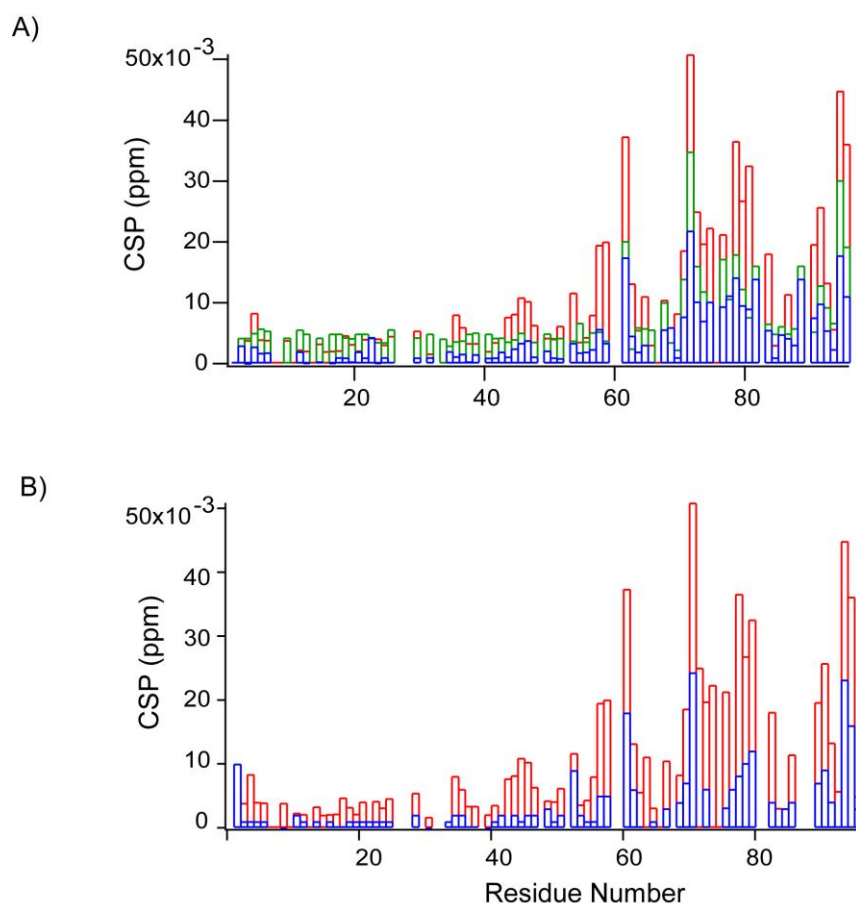


Figure 29: Mutation of important residues in presequence (rALDH)-Tim23ims complex (A) Comparison of averaged ¹H, ¹⁵N chemical shift deviation observed for backbone resonances of Tim23N in 2D ¹H-¹⁵N HSQC NMR spectra upon addition of 16 fold excess of wt rALDH (red) L18L19Q rALDH (green) and R17Q rALDH (blue). (B) Comparison of averaged ¹H, ¹⁵N chemical shift deviation observed for backbone resonances of wt (red) and L71A/78A Tim23N (blue) in 2D ¹H-¹⁵N HSQC NMR spectra upon addition of 16 fold excess of wt rALDH.

3.2.4 Tom22ims interacts weakly with presequence

The cytosolic part of Tom22 has been shown to be a receptor site for the presequence (van Wilpe et al. 1999; Rimmer et al. 2011). Additionally, Tom22ims along with Tom40ims and Tom7ims have been proposed to be involved in binding preprotein at *trans* site of outer mitochondrial translocase (Moczko, Bomer et al. 1997; Esaki, Shimizu et al. 2004). In order to deduce the binding site of presequence in Tom22ims, we titrated presequence (rALDH) to Tom22ims and followed the changes using ^1H - ^{15}N HSQC spectra. The addition of presequence (rALDH) to Tom22ims caused progressive changes in chemical shifts in ^1H - ^{15}N HSQC spectra of Tom22ims for the residues 123-133, without any significant changes in intensity (Figure 30, A). This is indicative of the fast exchange at the NMR timescale in the binding site of Tom22 involving residues 123-133 similar to that of Tim23-presequence complex. The residues (123-133) experiencing chemical shift perturbation upon the addition of presequence, were fitted to a single site binding model with an average K_d value of $\sim 0.9\text{mM}$ (Figure 30, D). The higher K_d is indicative of transient interaction i.e. higher dissociation rate of presequence with Tom22ims.

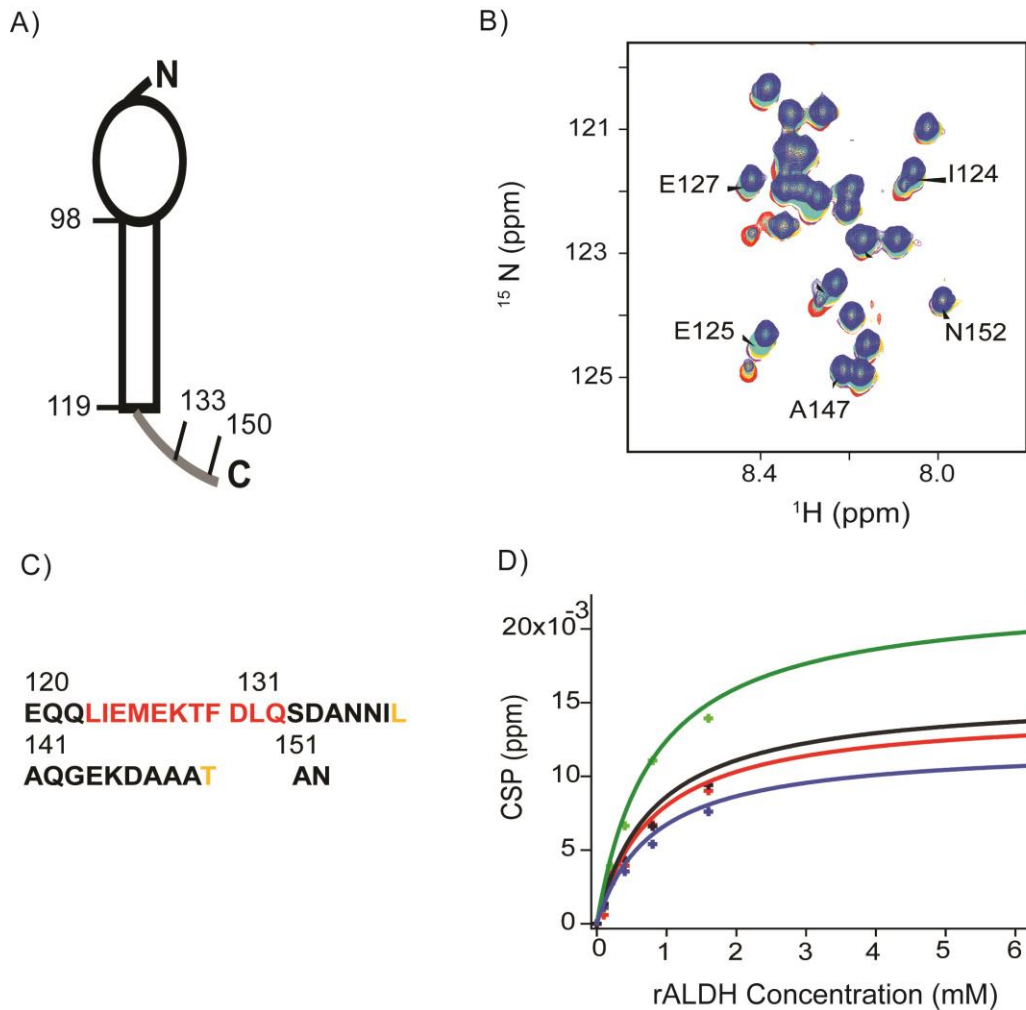


Figure 30: Binding of the presequence (rALDH) to Tom22ims. (A) The schematic representation of Tom22. Tom22 consists of an N-terminal cytosolic folded presequence receptor domain (oval) followed by a membrane spanning region (rectangle) and the C-terminal intermembrane space domain (extended curve in grey). The numbers correspond to the residues in the protein sequence of Tom22 in yeast. (B) The amino acid sequence of Tom22ims. Residues that show significant backbone chemical shift perturbations are marked in red. Two residues L140 and T150 that show chemical shift deviations outside the main binding site (residues 123-133) are marked in orange. (C) Superposition of a section of 2D ^1H - ^{15}N HSQC spectra of Tom22ims without rALDH (red) and at 100-fold molar excess of rALDH (Blue). (D) Chemical shift perturbation (CSP) as a function of rALDH concentration for residues 124 (black), 128 (red), 129 (green), and 133 (blue). Solid lines show the binding curves obtained by fitting to a one site model as in case of Tim23ims.

3.2.5 Tim21ims lacks presequence binding site

The intermembrane space domain of Tim21 has been proposed to interact with Tom22 and serves as the connecting link between two translocases (Mokranjac et al. 2005; Albrecht et al. 2006). After establishing the presequence binding region in Tom22ims, it was of a great interest for us to test the interaction between Tim21ims and

the presequence rALDH. The resonance changes in terms of chemical shift position and intensity in ^{15}N - ^1H HSQC of Tim21ims were absent even upon addition of 100 foldmolar excess of most soluble presequence rALDH (up to 3.2 mM) (Figure 31). These results suggest the absence of direct interaction between presequence and Tim21ims during preprotein import.

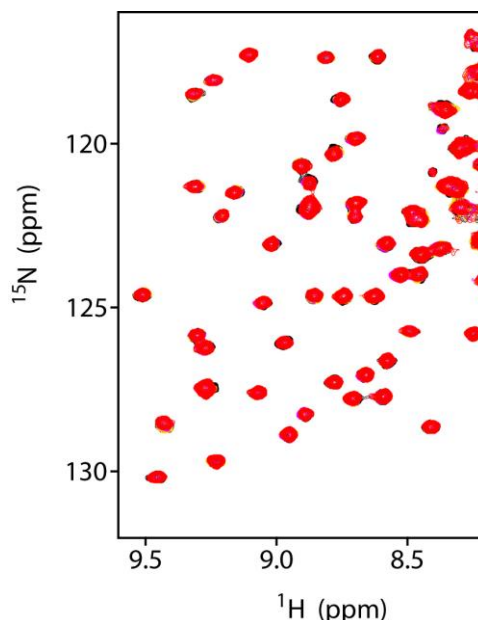


Figure 31: **Tim21ims lacks direct binding to presequence (rALDH):** Overlay of a region of ^1H - ^{15}N HSQC of Tim21ims without rALDH (black) and with increasing concentration of rALDH (4 fold excess :yellow, 16 fold excess: magenta, 64 fold excess: red) showing no change in the signal position and intensity.

3.2.6 Mapping presequence binding site on Tim50ims164-361

Tim50 acts as the receptor for presequence. The primary presequence binding site was recently assigned to the C-terminal intermembrane space domain residues 395-476 using photo crosslinking rALDH to Tim50 (Schulz et al. 2011). Additionally, a presequence binding groove was proposed in X-ray structure of Tim50¹⁶⁴⁻³⁶¹ (Qian, Gebert et al. 2011) (Figure 32). This motivated us to test the interaction of presequence (rALDH) with Tim50ims involving residues 164-364 using ^1H - ^{15}N -HSQC based titration. On comparing the ^1H - ^{15}N TROSY-HSQC spectra of Tim50¹⁶⁴⁻³⁶¹ with 16 fold molar excess of the presequence with free Tim50¹⁶⁴⁻³⁶¹ many changes were observed. This includes some residues exhibiting small changes in chemical shift as well as intensity to others shows a decrease in intensity as highlighted in Figure 32, B.

However, the spectra of bound form of Tim50¹⁶⁴⁻³⁶¹-presequence even at 16 fold excess of presequence, does not have additional resonances. These changes are characteristic of binding events occurring at fast to intermediate exchange regime at NMR time scale.

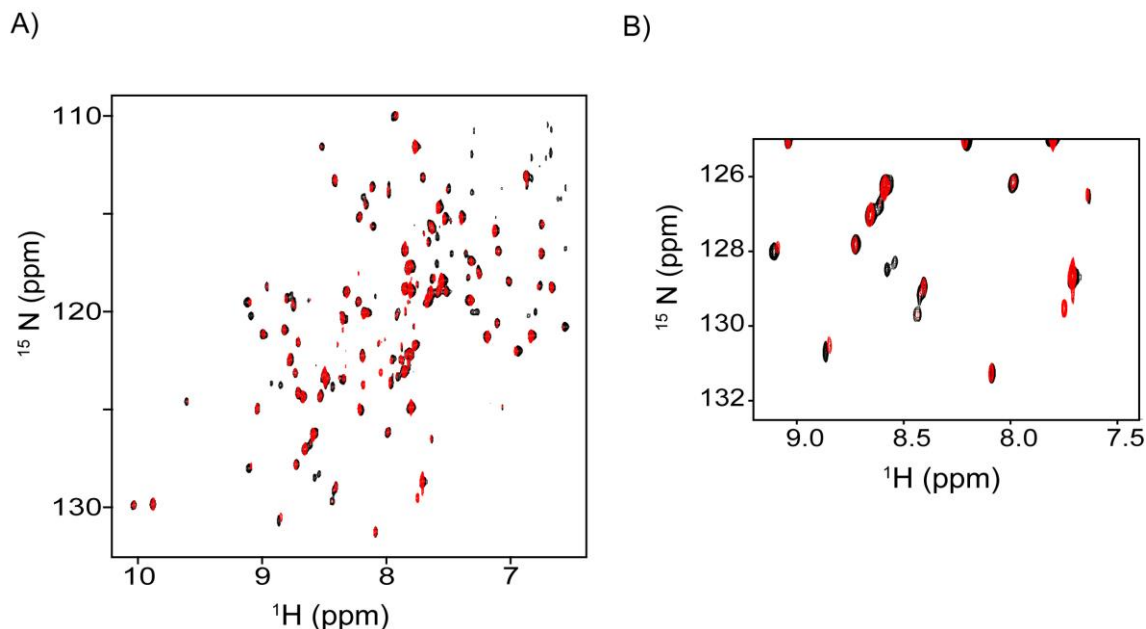


Figure 32: Tim50ims¹⁶⁴⁻³⁶¹ binds to presequence rALDH: Overlay of ¹H-¹⁵N TROSY-HSQC of Tim50¹⁶⁴⁻³⁶¹ims without rALDH (black) and with 16 fold excess of rALDH (red). (A) The complete ^H^N-N fingerprint region of Tim50ims. (B) Amide-proton amide-nitrogen regions highlighting changes of presequence bound resonances of Tim50ims in chemical shift and intensity.

The resonances of Tim50ims that gets perturbed in terms of chemical shift and intensity belong to ¹H downfield region at 9 ppm (β -strand region) and up field at 6.5 ppm (α -helix region). This suggests that the folded part of Tim50ims is most likely involved in binding to the presequence. More than 30 resonances in ¹H-¹⁵N TROSY-HSQC spectra of Tim50ims gets perturbed upon addition of a small 22 residues presequence (rALDH). This implies that either Tim50ims undergo a conformational change to accommodate the presequence or has a presequence binding groove as indicated in crystal structure to accommodate it. However, the assignment of Tim50¹⁶⁴⁻³⁶¹ is not available and the detailed binding studies of the presequence at higher ratios and in more steps will help in the precise mapping of the residues involved in binding to the presequence.

3.3 Interactions at the translocation contact site

The translocation contact site is generated by TOM-TIM super complex during preprotein import into mitochondria as an intermediate step to guide the preprotein to the inner mitochondrial translocase. The intermembrane space domains of Tom22, Tom40, Tim21, Tim23 and Tim50 have been proposed to be the part of complex and have been cross-linked *in vivo* (Tamura et al. 2009). Deletion of intermembrane space domain of Tom22 drastically affected the super complex assembly *in vivo* suggesting the importance of Tom22ims in maintaining the stability of super complex (Chacinska et al. 2003). Tim50 can be cross-linked to Tom22 *in vivo* (Shiota, et al. 2011). However, there is no evidence for the direct interaction between Tim50ims and Tom22ims. Interestingly, Tim21ims has been proposed to be a connecting link between two translocases (Mokranjac et al. 2005; Albrecht et al. 2006). Moreover, in Tim23, first 50 residues were suggested to insert into the outer mitochondrial membrane to facilitate the import by the formation of the translocation contact site. Contrasting results have been shown *in vivo* with Tim23 lacking N-terminal 50 residues where either import was affected severely or up to 10% than in wild type Tim23 (Donzeau et al. 2000; Chacinska et al. 2003; Tamura et al. 2009). However, Tim23 can be cross-linked to Tom22 *in vivo* (Tamura et al. 2009).

Evidence for direct interaction among the various intermembrane space domains at the translocation contact site is not well understood. To this end, unlabeled Tom22ims was titrated to ^{15}N labeled Tim21ims and Tim23ims. Furthermore, Tom40ims was also tested for its potential interaction with Tim23ims.

3.3.1 Tom22ims does not bind Tim21ims

Tim21 has been earlier shown to bind electrostatically to TOM complex, by affinity tagging Tim21 and incubating it with solubilized mitochondria. *In vitro* crosslink of purified Tim21ims and Tom22ims have also been observed (Albrecht et al. 2006). Tom22ims was soluble only up to a concentration of 0.8mM in the NMR buffer. The increasing amounts of Tom22ims were added to Tim21ims that resulted in no obvious change in peak position and intensity in ^1H - ^{15}N -HSQC. To elevate the

solubility problem, another mutant peptide of Tom22ims was titrated to ^{15}N -Tim21ims that resulted in small changes in the ^1H - ^{15}N -HSQC spectrum; however they are less significant than the changes in Tim21ims because of increased salt concentrations as described in Appendix 1.2

To provide further evidence for the observed changes in ^1H - ^{15}N -HSQC spectrum of Tim21ims upon the addition of Tom22ims and to probe the potential transient interaction between Tim21ims and Tom22ims, a paramagnetic tag (MTSL) was attached to Tom22 peptide and intermolecular PREs were measured on ^{15}N labeled Tim21ims. As shown in Figure 33 D, the intensity ratio for ^{15}N -Tim21-C30A with MTSL labeled Tom22 at E1C (oxidized, paramagnetic, I_p) and reduced with ascorbic acid (diamagnetic, I_d), except residues F200, R213, K217 and K221, which are highlighted in green, all residues have I_p/I_d above 0.8 indicating no interaction among Tim21ims and Tom22ims. These exceptional residues did not show any chemical shift perturbations and the intensity changes observed in the intermolecular PRE measurements could be attributed to the nonspecific interactions MTSL with aromatic phenylalanine and the surrounding positive charge patch.

Furthermore, reverse titration of unlabeled Tim21ims to ^{13}C , ^{15}N -Tom22ims doesn't lead to any significant changes in both ^1H - ^{15}N -HSQC and ^1H - ^{13}C -HSQC spectra confirming the absence of any direct binding between Tim21ims and Tom22ims (Figure 33, B).

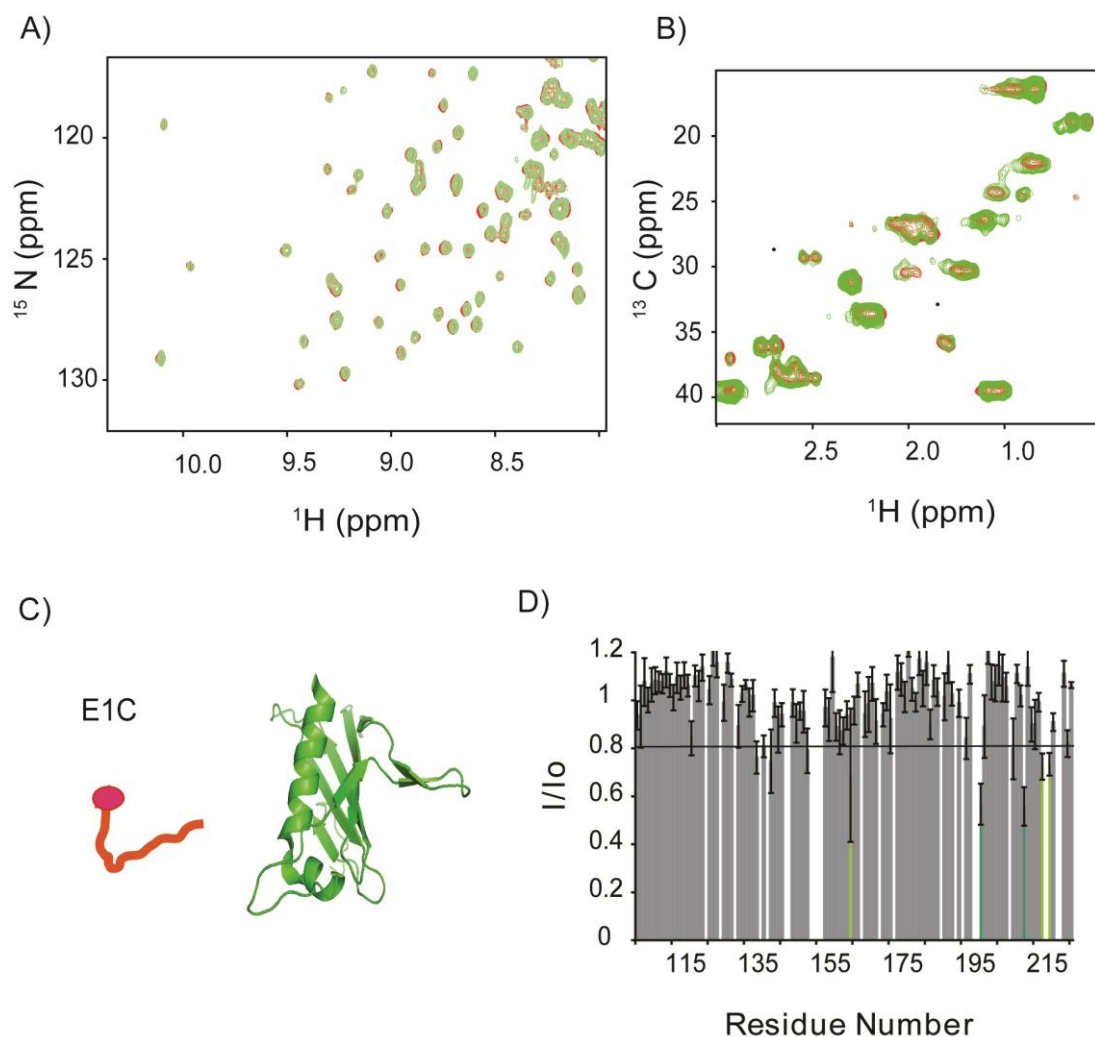


Figure 33: Intermembrane space domain of Tim21ims does not interact with Tom22ims (A) Overlay of a section of ^1H - ^{15}N -HSQC of Tim21ims with and without addition of Tom22ims mutant (red: reference, green: 16 fold molar excess of Tom22). (B) Overlay of a section of ^1H - ^{13}C -HSQC of ^{15}N - ^{13}C Tom22ims corresponding to C^β side chain region of apo Tom22 (red) and 16 fold molar excess of Tim21ims (green) showing no significant chemical shift changes. (C) The cartoon representation of PRE experiment with 1:1 ratio of Tom22ims and Tim21ims. Tom22ims E1C peptide carrying a MTS tag (oval) is colored in red and Tim21ims structure in green. (D) Intermolecular PRE broadening profile of ^{15}N -Tim21ims C128A up on addition of E1C-Tom22ims. Gaps in intensity ratio plot for non proline residues are either due to overlapping resonances or extremely broadened resonances in apo form. I_p/I_a ratio of 0.8 is used as the cutoff for selecting the non-interacting residues and is marked with a black line. The error bars are based on S/N ratio.

3.3.2 Tim23ims weakly interacts with Tom22ims

Tim23 has been shown to crosslink with Tom22 *in vivo* (Tamura et al. 2009). To identify the residues of Tim23ims involved in binding Tom22ims, the changes in the $^1\text{H}^{15}\text{N}$ -HSQC of Tim23ims with increasing amount of Tom22ims were followed. The addition of Tom22ims to ^{15}N -Tim23ims resulted in large chemical shift changes for the resonance belonging to residues V53 and L58-A62 as observed from spectra and are highlighted in the chemical shift perturbation plot of Tim23ims upon addition of Tom22ims (Figure 34, A and B). Due to the limited solubility of Tom22ims peptide, saturation could not be achieved and the K_d value of the above said fast exchanging peaks could not be estimated.

To ascertain the binding of Tim23ims with Tom22ims, reverse titration with the labeled Tom22ims in the presence of the increasing amount of Tim23ims was done. ^{13}C - ^1H -HSQC of Tom22ims showed the changes in resonances belonging to the aliphatic side chain region of Tom22ims confirming the interaction among the Tim23ims and Tom22ims (Figure 34, C).

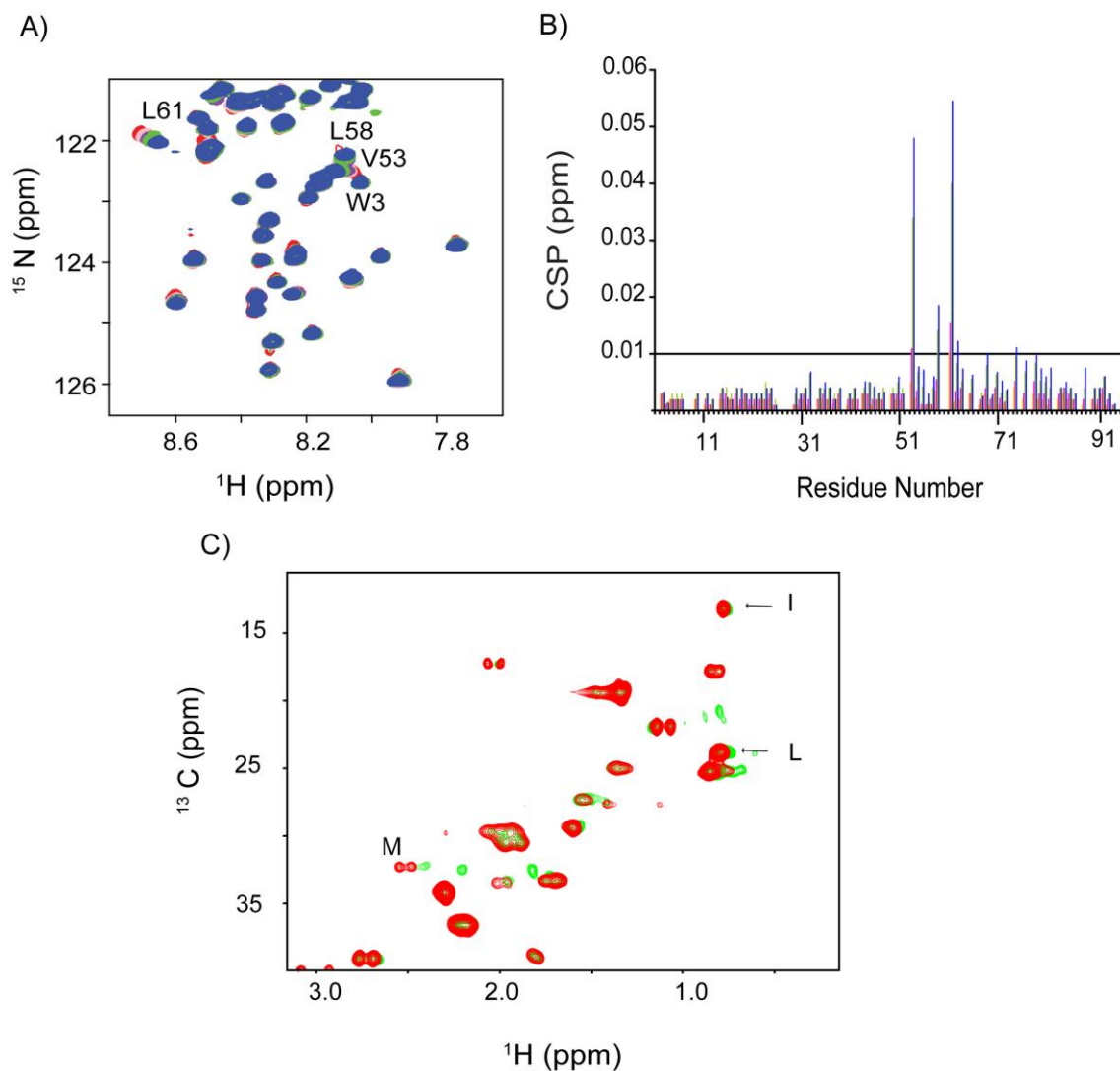


Figure 34: *Tom22ims* interacts with *Tim23ims*. (A) Overlay of a region of ^1H - ^{15}N HSQC of apo *Tim23ims* (red) and 32 fold molar excess of *Tom22ims* (blue). Labels inside the figure correspond to the most perturbed resonances with their corresponding labels as L58, V53, L61 and a non-interacting residue W3. (B) Chemical shift perturbation (CSP) plot as a function of primary sequence of *Tim23ims* upon addition of *Tom22ims* -1 fold molar excess (yellow), 2 fold molar excess (pink), 8 fold molar excess (green), 16 fold molar excess (blue). The black line represents twice the average CSP of non-shifting residues 1-50. (C) Overlay of a section of ^1H - ^{13}C HSQC of ^{15}N - ^{13}C *Tom22ims* corresponding to aliphatic carbon side chain region of apo *Tom22* (red) and 16 fold molar excess of *Tim23ims* (green) showing the chemical shift changes in *Tom22ims*. The correlation peaks corresponding to $\text{C}^\delta\text{H}^\beta$ of Leucine (L), $\text{C}^\delta\text{H}^\beta$ of Isoleucine and $\text{C}^\beta\text{H}^\beta$ of methionine (M) are marked.

3.3.3 Tim23ims binds to Tom40ims

NMR based titration of unlabeled Tom40ims with ^{15}N -Tim23ims resulted in selective chemical shift perturbation of residues V53, L58-L61 and L64 in Tim23ims beyond addition of 8 fold molar excess of Tom40ims (Figure 35,A). Based on the digital resolution of the acquired spectra, the accuracy of change in chemical shift perturbation is significant beyond 0.005ppm. Hence, the chemical shift perturbation changes in Tim23ims upon addition of Tom40ims beyond 8 fold excess are significant (Figure 35, B). The same directional chemical shift change (trajectory) was observed in above said residues of Tim23ims upon titration with 16- fold and 32-fold excess Tom40ims, which further confirms them as a part of Tom40ims binding site in Tim23ims. In line with the above results of small selective yet progressive chemical shift changes, suggests that the affinity of Tom40ims with Tim23ims is low, i.e. binding is too weak. However, determination of exact K_d value was hampered by the poor solubility of Tom40ims.

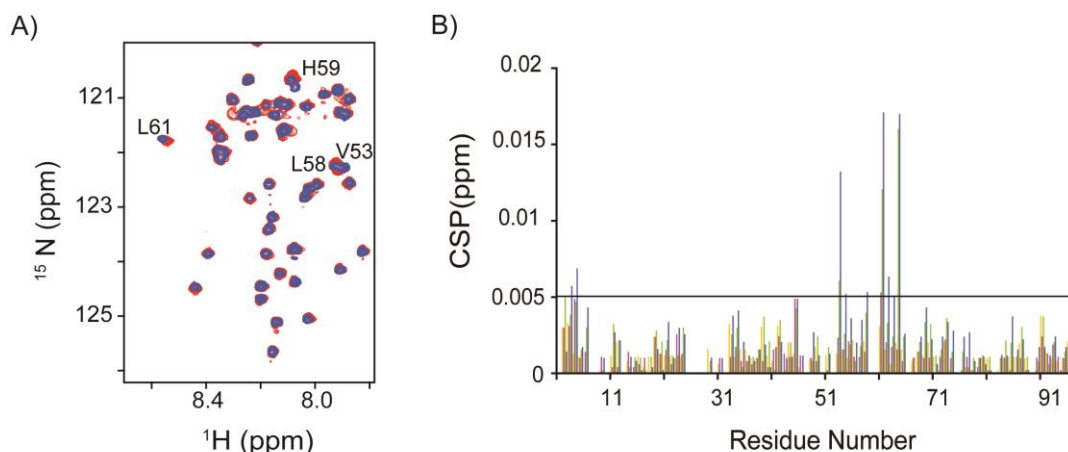


Figure 35: *Tom40ims weakly interacts with Tim23ims.* (A) Superposition of a region of $^1\text{H}^{15}\text{N}$ -HSQC of apo Tim23ims (red) and 16 fold molar excess of Tom40ims (blue). The most perturbed resonances are highlighted with their corresponding labels. (B) Chemical shift perturbations (CSP) as a function of residue number of Tim23ims upon addition of 4 fold molar excess (yellow), 8 fold molar excess (magenta), 16 fold molar excess (green) and 32 fold excess (blue) of Tom40ims.

3.4 Interaction studies of Tim23 with mitochondrial membrane mimics

The first twenty residues at the N-terminus of Tim23 were found to be sensitive to protease cleavage in mitochondria. These residues have been suggested to transverse the outer membrane (Donzeau et al. 2000). In addition, deletion of first 24 residues causes the growth defect (Davis et al. 2000) and contributes to the morphological defect with less number of cristae in mitochondria in yeast (Donzeau et al. 2000). The dynamic association and dissociation of N-terminal Tim23 with the outer membrane and its role in presequence carrying preprotein import has been suggested to be modulated by Tim50 (Tamura et al. 2009). The molecular and structural basis of such membrane protein interaction is perplexing. We have used two membrane mimicking environments i.e. liposomes imitating the mitochondrial membrane composition and micelles (DHPC) to identify and structurally characterize the residues of Tim23ims that interacts with the mitochondrial membrane in yeast.

3.4.1 Tim23 interacts with mitochondrial membrane mimicking liposomes

3.4.1.1 Outer mitochondrial membrane mimicking liposomes

The liposomes, an excellent mimic for the outer mitochondrial membrane (OMM), consisting of the OMM composition were titrated with Tim23ims and series of ^1H - ^{15}N -HSQC spectra of Tim23ims were obtained at 0, 0.5, 1.5 and 3 fold molar excess of total lipid to protein ratio.

On increasing the total lipid concentration of liposomes, the subsequent attenuation in the signal intensity without any change in the chemical shift position for the resonances corresponding to residues 1 to 12 (N-terminal region), 29-46 (intermediate region) and 53-58 (least perturbed region) of Tim23ims was observed. These changes are indicative of slow exchange on NMR timescale between the membrane bound and free form of Tim23ims. The signal intensity for the residues 17-26 and C-terminal residues 60-96 neither show changes in chemical shift nor the intensity

indicating that these residues are unaffected and retain the disorder in the liposome bound form of Tim23ims (Figure 36 A and B).

Upon liposome addition, the N-terminal region of Tim23ims involving residues 1-12, shows the signal intensity attenuation to 0.6 at 0.5 molar excess of total lipid and are almost completely broadened on addition liposomes (at total lipid to protein ratio 1 and 3) suggesting that the N-terminal liposome binding region involving residues 1-12 acts as a prime membrane interacting region in Tim23ims.

3.4.1.2 Effect of cardiolipin on binding of Tim23ims to mitochondrial membrane

Tim23ims can potentially interact with the inner mitochondrial membrane. Cardiolipin is a unique constituent of the inner mitochondrial membrane (IMM). It is a negatively charged lipid and is required for the functional reconstitution of protein complexes in the inner mitochondrial membrane (Claypool 2009; Marom et al. 2009). To address whether Tim23ims exclusively interacts with the outer mitochondrial membrane or can it also interact with the inner mitochondrial membrane.

Additionally, to the understand role of cardiolipin in membrane binding of Tim23, the following Tim23ims-liposome titrations were performed.

- A) Tim23ims with liposomes mimicking lipid composition (PE, PC and PS) of the inner mitochondrial membrane lacking cardiolipin.
- B) Tim23ims with liposomes mimicking lipid composition (PE, PC and PS) of the inner mitochondrial membrane containing cardiolipin.

On keeping the same total lipid concentration, but varying its composition to mimic either OMM or IMM containing phospholipids (without cardiolipin) liposomes, resulted in similar changes in the signal intensity in the Tim23ims (Figure 36, C). This shows that Tim23ims can bind to liposomes containing phospholipids PE, PC and PS irrespective of their constituent ratios.

However, addition of the cardiolipin containing IMM liposomes to Tim23ims resulted in two perturbation patterns in the peaks for the affected resonances.

- 1) The residues only with a reduction in the signal intensity (i.e. residues 1-67) indicative of slow exchange.

- 2) The residues showing small directional change in the chemical shift as well as the reduction in signal intensity (68-71) indicative of fast to intermediate exchange. The increase in the region of membrane interaction for Tim23ims from residue 46 onwards upon addition of the cardiolipin containing IMM liposomes could be attributed to the electrostatic interaction between the negatively charged cardiolipin on surface of liposomes with the positively charged residues (H52, R57, H59, and K66) of Tim23ims.

The interaction of Tim23ims with the mitochondrial membrane mimicking liposomes shows that Tim23ims possesses two membrane binding sites in 50 N-terminal residues of Tim23ims as N terminal region (1-12) and intermediate region (29-46).

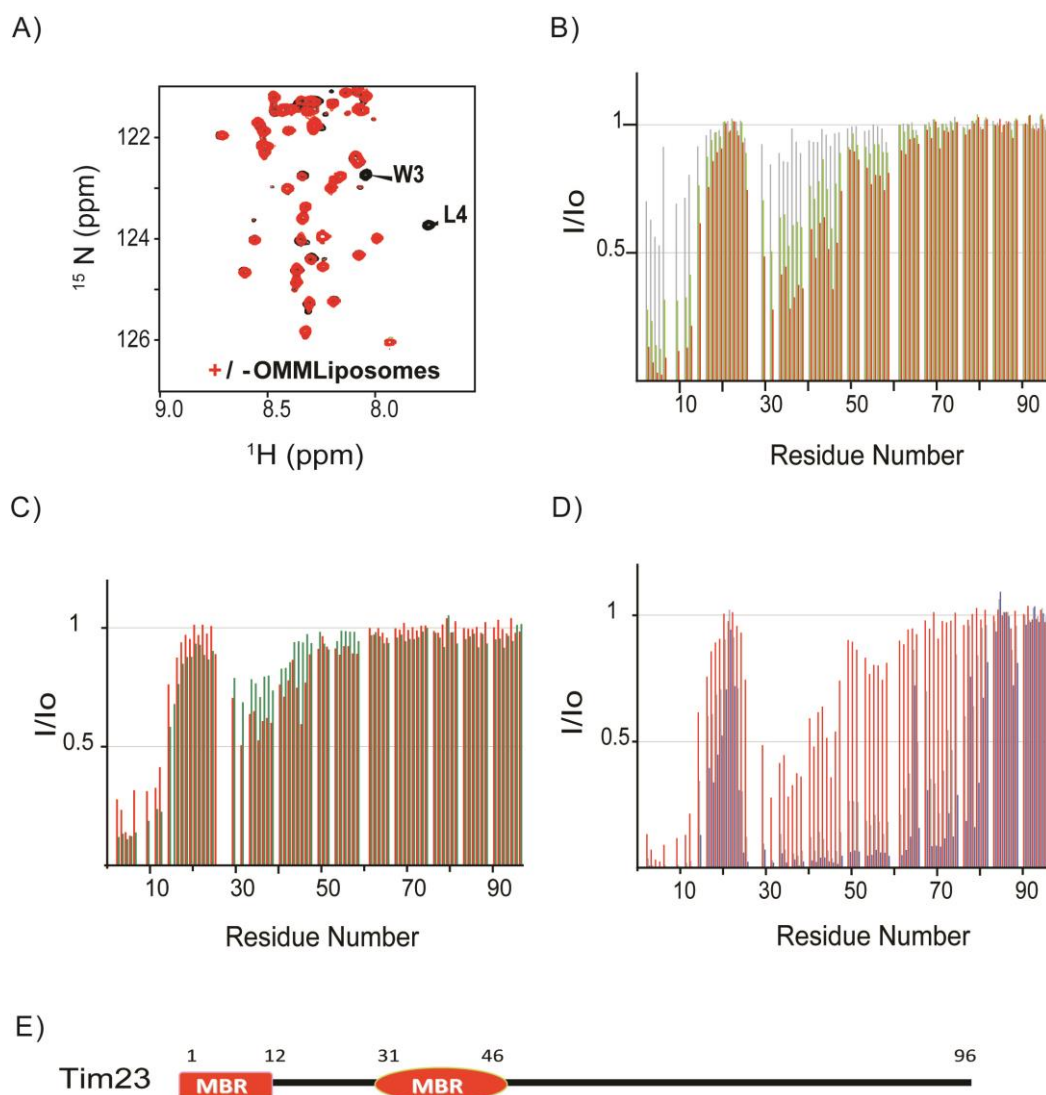


Figure 36: Membrane recognition by Tim23ims. (A) Superposition of 2D ^1H - ^{15}N HSQC spectra of Tim23ims without OMM (black) and with OMM (red). Resonances corresponding to W3 and L4 are labeled. In (B), (C) and (D) the intensity ratios (I/I_o) was calculated from the ratio of I (data height of a resonance from ^1H - ^{15}N HSQC spectra of Tim23ims at defined fold molar excess of liposome) and I_o (data height of the same resonance from ^1H - ^{15}N HSQC spectra of free Tim23ims). (B) The intensity changes upon addition of OMM as a function of residue number of Tim23ims with grey as 0.5, green as 1.5 and red as 3 fold molar excess of total lipid. (C) Intensity attenuation profile of Tim23ims with liposome mimicking composition of OMM (red) and IMM (green) without cardiolipin at 1 fold molar excess of total lipid to protein ratio. (D) Comparison of intensity attenuation profile of Tim23ims upon addition of OMM (3 fold molar excess of total lipid to protein ratio as red) and IMM with cardiolipin (grey as 0.8 and blue as 3 fold molar excess of total lipid to protein). (E) A cartoon representation of two membrane binding regions (MBR) of Tim23ims. The numbers represents the residues in Tim23ims sequence.

3.4.2 Interaction of Tim23 to micelles (DHPC)

3.4.2.1 Chemical Shift perturbation of Tim23ims with the detergent DHPC and its micelles

The presence of the slow exchange regime of Tim23-membrane interaction at NMR timescale where resonances from the bound form were not visible in NMR spectra hampered the structural characterization of the membrane bound form of Tim23ims. Therefore, we have used another membrane mimic (the DHPC detergent and its micelles) to structurally characterize the Tim23ims-membrane interaction.

Upon titration of Tim23ims with increasing amount of DHPC, changes in the chemical shift of backbone resonances of Tim23ims on binding to the DHPC molecules along with the changes in the resonance intensities using ^{15}N - ^1H -HSQC spectra were monitored. The addition of DHPC below its critical micellar concentration (CMC) (i.e. up to 16 mM) did not induce any changes in the Tim23ims spectra. However, the further stepwise addition of DHPC above its CMC, where micelles are formed (i.e. at 24mM and 32mM) caused progressive changes in the backbone amide resonances of Tim23ims for the residues involved in binding to DHPC micelles (Figure 37, A). The observance of chemical shift changes along with small, changes in resonance intensities (at 24mM and 32mM DHPC concentration) is a characteristic of presence of fast to intermediate exchange at NMR time scale. The observed changes in intensity can be attributed to the structural flexibility of the complex or to the chemical exchange (between the DHPC micellar bound and free form of Tim23ims)

Based on the degree of the chemical shift changes in Tim23ims upon micellar (DHPC) binding, the residues can be grouped into four categories: (1) the most perturbed residues 1-7, (2) the intermediately perturbed residues 33-46 and (3) the less perturbed 53-62 and (4) the least perturbed residues or non-interacting residue (7-33, 61-93) (Figure 37 B). The similar pattern of reduction in the resonance intensities was found in the aforementioned regions. The micellar binding regions of Tim23 are similar to the liposome binding region identified in section 3.4.1.

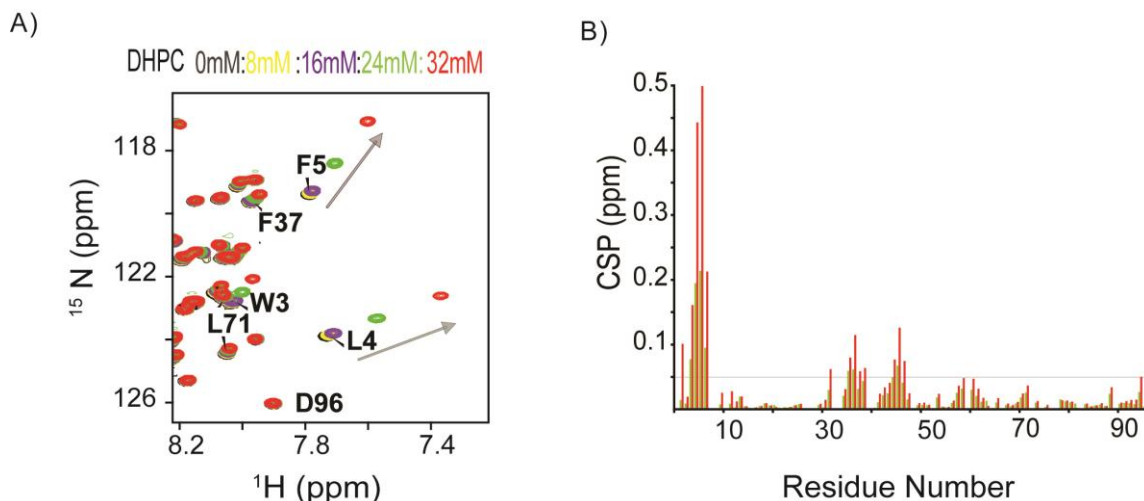


Figure 37: Tim23 interacts with DHPC micelles. (A) Superposition of 2D ^1H - ^{15}N HSQC spectra of Tim23 without DHPC (black), with 8mM DHPC (yellow), 16mM DHPC (purple), 24mM DHPC and 32mM DHPC (red). The progressive changes in the resonance positions of L4 and F5 are indicated by arrows. (B) Average chemical shift perturbation (CSP) of Tim23 resonances upon addition of 24mM DHPC (green) and 32mM DHPC (red). The grey line represents maximum CSP value of the C-terminal 50 residues.

3.4.2.2 ^1H - ^{15}N HetNOE of micelle-bound Tim23ims

In order to probe the backbone dynamics of micellar bound Tim23ims in detail, we measured the steady state ^1H - ^{15}N HetNOE of Tim23ims in presence of micelles (saturating amounts of 50mM DHPC). On average, positive ^1H - ^{15}N HetNOE value of 0.42 was obtained for an N-terminal membrane binding site (1-7) of Tim23ims in comparison of negative ^1H - ^{15}N HetNOE value observed in the free form (Figure 38). This observed change in the ^1H - ^{15}N HetNOE values is a direct evidence for the increase in rigidity of HN bond vector for these residues upon binding to DHPC micelles. However, the magnitude of ^1H - ^{15}N HetNOE values of the micellar bound Tim23ims is still less than as would be expected in for the presence of a rigid secondary structure in proteins (0.6-0.8). The residues 91-94 also show the increase in ^1H - ^{15}N HetNOE value in comparison to free Tim23ims but negative sign is indicative for flexible C-terminus of Tim23 even in micellar bound form.

For the intermediately perturbed residues, the less significant, small positive changes in ^1H - ^{15}N HetNOE values were observed suggesting them to be disordered in the micellar bound form. Moreover, the lack of overall change in ^1H - ^{15}N HetNOE

values of micellar bound Tim23ims, suggests that non-interacting residues retains a similar disordered state in the micellar bound form.

The chemical shift perturbations and ^1H - ^{15}N HetNOE data together suggests two regions i.e. N terminal (1-7) and intermediate (31-46) of Tim23ims are involved in binding to the membrane mimicking micelles (DHPC) but remain dynamic and exchange between the in the micelle bound and free form of Tim23ims.

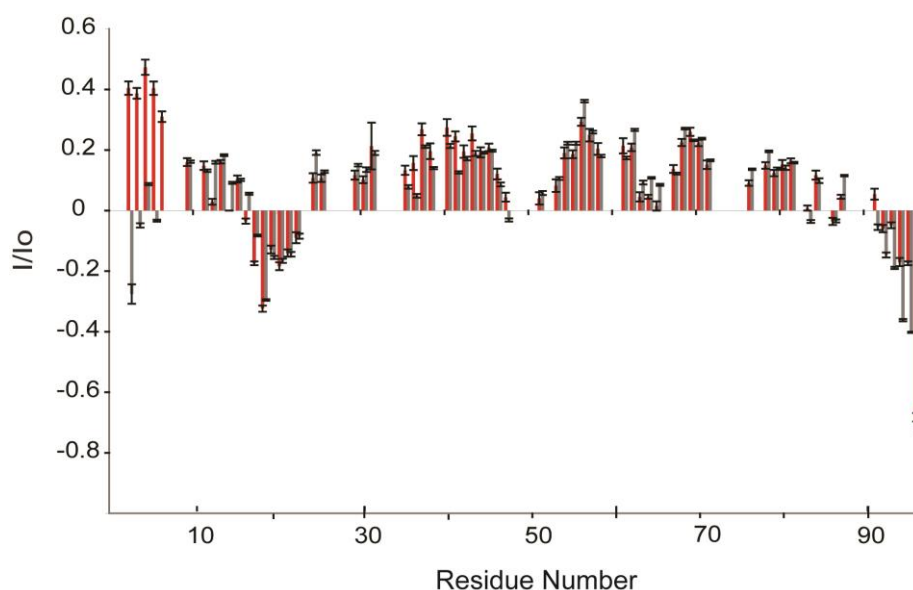


Figure 38: ^1H - ^{15}N -HetNOE intensity ratios ($I_{\text{sat}}/I_{\text{ref}}=I/I_o$) for Tim23ims in the presence (50mM) and the absence (grey) of DHPC micelles. The error bars are calculated based on the observed S/N ratio.

3.4.2.3 Structural analysis of micelle bound Tim23

In order to understand the change in the conformation of Tim23ims upon micelle binding (50mM DHPC), we characterized the changes in its backbone dihedral angles by evaluating the deviations of the measured C^α chemical shift from the sequence based random coil values i.e. chemical shift index (CSI) of C^α in micelle-bound Tim23 form (50mM DHPC). The average CSI value for all residues remained below $\pm 0.5\text{ppm}$

(except for Ser 2) indicating lack of rigid secondary structural elements present in Tim23ims even upon binding to DHPC micelles. In the main N-terminal micelle binding site (involving residues 1-7) only three residues 2, 3, 4 contain continuous stretch of positive values. Additionally in second micellar binding region of Tim23ims (involving residues 31-46), positive values were seen at stretch for residues 40-44. This precludes the existence of secondary structural elements (α -helix or β -sheet) in micellar bound Tim23ims.

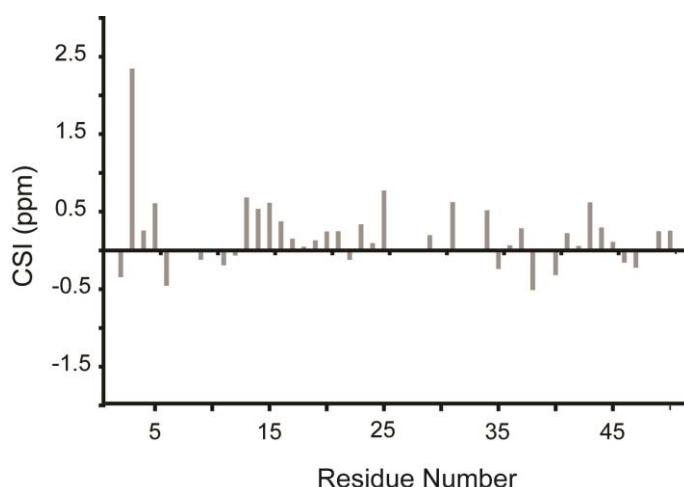


Figure 39: **Micelle bound Tim23ims lacks rigid secondary structure:** Deviations of the experimentally measured C^α chemical shift from that of random coil values is plotted as a function of Tim23ims sequence. For clarity only micellar binding regions present in first 50 residues of Tim23ims are shown.

3.4.2.4 Intermolecular NOE based structure of micelle bound Tim23¹⁻¹³

Further structural aspects of micellar bound Tim23 were investigated using the ^1H - ^1H NOESY of most perturbed N-terminal membrane binding region of Tim23 (1-13) in DHPC micelles (200mM). The intermolecular NOEs were observed between the hydrophobic side chains of DHPC detergent and side chain ^1H for the W3 and F5. The presence of sequential H^{N} - H^{N} NOE's among residues Trp³, Leu⁴, Phe⁵ and Gly⁶ indicates that these residues are presumably involved in the formation of a turn. The absence of medium and long range NOE's for the residues 7-13 resulted in these residues being unstructured. To gain the insights into the binding of these residues in micellar bound form of Tim23 (1-13), the structure of the micellar bound Tim23 (1-13) peptide was calculated using 70 short range, 4 medium range and 2 long range NOEs.

Furthermore, the ensemble of 15 low energy structures calculated showed poor convergence for the overall molecule. However, when the residues 2-6 are superimposed, the backbone and side-chain r.m.s.d were found to be 0.33 ± 0.13 Å and 0.59 ± 0.21 Å, respectively, indicating that residues 2-6 adopts a well ordered structure (Figure 40, B and C). On the other hand, when the residues 7-13 are superposed, the backbone and side-chain r.m.s.d was found to be 2.20 ± 0.52 Å and 3.36 ± 0.68 Å respectively. Most importantly, the side-chains of residues Trp³, Leu⁴ and Phe⁵ are oriented in same direction (Figure 40 C, D, and E). These side chains of Trp³, Leu⁴ and Phe⁵ forms a hydrophobic cluster and are involved in the interaction with the hydrophobic core of micelles DHPC.

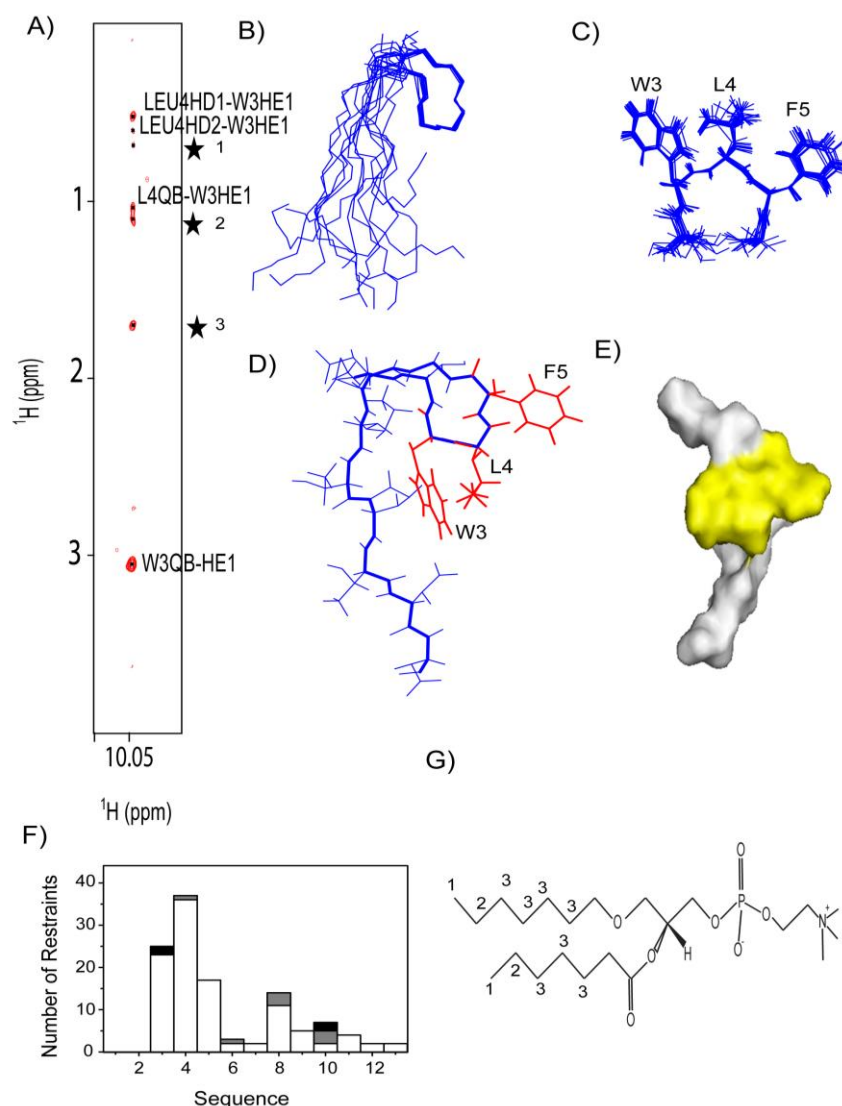


Figure 40: Structural analysis of micellar bound N-terminal (1-13) peptide of Tim23. (A) Intermolecular NOEs between side chain HE1 of Trp³ and the hydrophobic tails of protonated DHPC are highlighted with asteriks 1, 2, 3 as in (G) and interresidual side chain ¹H NOEs between side chain HE1 of Trp³ with side chain of L4 HB and HD1HD2 are labeled. (B) The ¹H-¹H NOE based structure of micellar Tim23(1-13) showing the good convergence of 15 conformers low energy for the N terminal residues 2-6. (C) Enlarged view of ensemble of 15 low energy structures highlighting hydrophobic cluster involving side chains of residues W3, L4 and F5 that pointing outward. (D) The hydrophobic cluster for lowest energy conformer was shown for clarity with residues of hydrophobic cluster marked in red. (E) Surface representation of Tim23 peptide 1-13 with hydrophobic cluster marked in yellow. (F) Histogram for number of NOE restraints incorporated for structure calculation: white, grey and black bars represents short, medium and long range restraints. (G) Chemical structure of DHPC. CH_n groups of the hydrophobic tails. The signals arising from the hydrophobic end (–CH₃ marked as 1 and antepenultimate –CH₂ as 2 and others –CH₂ as 3 except two unmarked. These signals are marked in (A) as labels+ asteriks.

3.4.3 Presequence binding site is accessible in liposome bound Tim23ims

When individual ligands i.e. either presequence or liposome (membrane imitating environment) are added to Tim23ims, the binding sites for both the ligands are well segregated in an intrinsically disordered Tim23ims. Tim23ims binds to the presequence through its C-terminus with residues 71-84 (section 3.2.1) whereas liposome binding sites are confined to the N-terminal involving residues 1-12 and 29-46 (section 3.4.1). It is intriguing to know what happens *in vitro* when one of the ligand is bound to Tim23ims and the binding site for other ligand is free. This might help us in understanding the effect of the simultaneous binding of Tim23ims with membrane and incoming presequence containing preprotein with respect to its biological function of preprotein import.

To this end, we have titrated the IMM mimicking liposome bound Tim23ims with increasing concentration of the presequence rALDH. As previously seen in section 3.4.1, upon addition of the IMM membrane mimicking liposomes the signal intensity of the resonances in the membrane-binding region of Tim23ims is reduced (Figure 41 A, black bars). This confirms that the Tim23ims is bound to (IMM) liposomes

The stepwise addition of presequence to the liposome bound Tim23ims resulted in a similar progressive chemical shift perturbation in the presequence binding region (residues 71-84) as observed for the titration of Tim23ims in the absence of liposomes (Figure 28). The observed exchange kinetics at NMR time scale and the direction of chemical shift change for the presequence interacting residues are also found to be similar. Furthermore, the magnitude of the chemical shift changes for the titrations of Tim23ims at 16 fold molar excess of presequence were identical both in the presence and absence of liposome (Figure 41 and Figure 42). Taken together, these results suggest that the formation of presequence-Tim23ims complex is not influenced by the presence of liposomes. The residues involved in liposomes binding (1-12 and 29-46) showed increased signal intensity upon progressive addition of presequence, presumably due to the dissociation of Tim23ims from the liposomes at higher concentrations of the presequence. Hence, two events are occurring simultaneously during this titration viz., binding of the presequence to Tim23ims (at residues 71-84) and the dissociation of Tim23ims from the liposomes. However, using NMR

spectroscopy, it is difficult to predict the order in which these events are occurring in equilibrium.

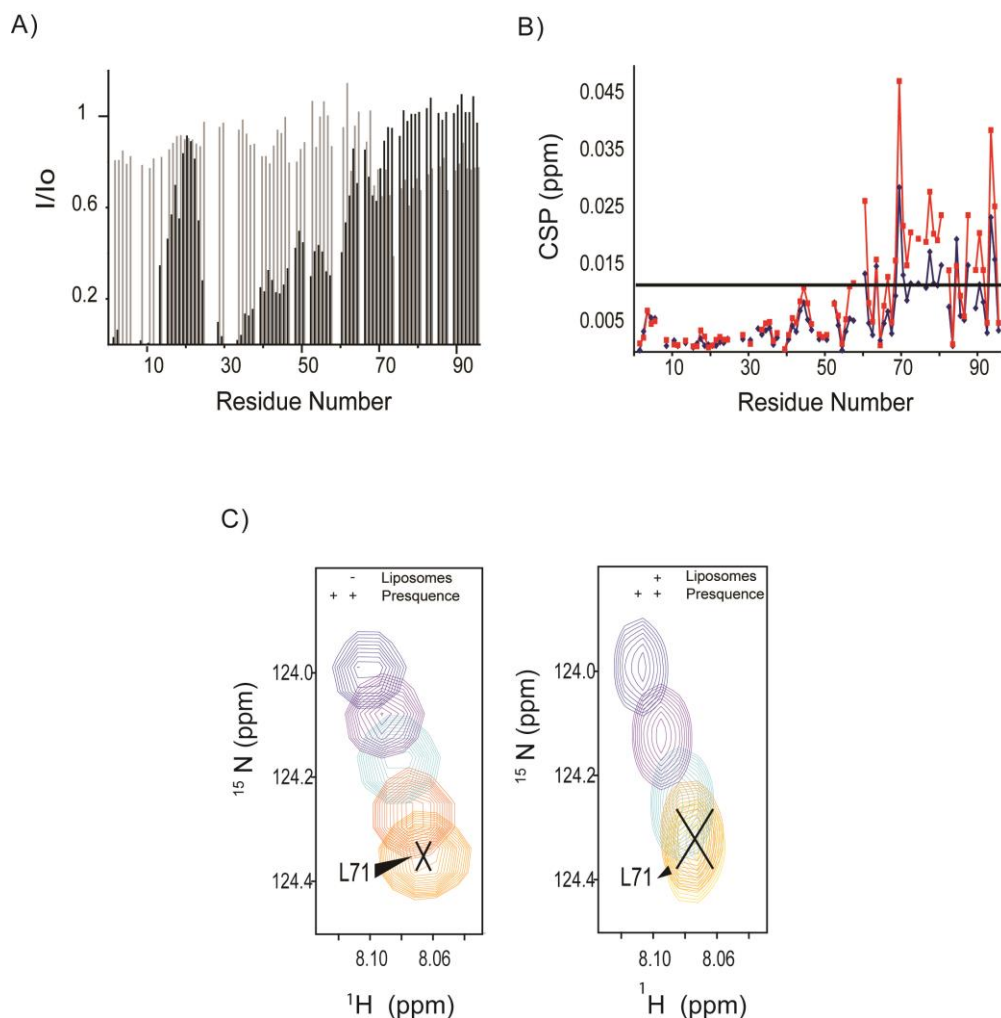


Figure 41: Binding of liposome bound Tim23 to presequence: In this figure, liposome bound Tim23ims represents the Tim23ims in 0.5 fold molar excess of total IMM lipids to protein ratio. For I/I_0 (Intensity ratio), I is obtained from data height of a resonance from ^1H - ^{15}N HSQC spectra of Tim23ims at defined fold molar excess of presequence and I_0 from the data height of the same resonance from ^1H - ^{15}N HSQC spectra of free Tim23ims. (A) Intensity changes upon addition of 16 fold molar excess of presequence added to liposome bound Tim23ims (grey) and liposome bound Tim23ims (black) i.e. zero equivalents of presequence. (B) Average chemical shift perturbation (CSP) upon addition of presequence to liposome bound Tim23ims, 8 fold excess (blue) and 16 fold excess (red) indicating that the presequence binding site is localized to C- terminal residues of Tim23ims. The black line represents the maximum value for the non-interacting 50 residue in N-terminus of Tim23ims in presence of 16 fold molar excess of presequence. (C) Comparison of presequence addition to free Tim23ims (left) and liposome bound

Tim23ims (right) at no excess (orange), one fold excess (red) two fold excess (cyan), four fold excess (purple) and 8 fold excess (blue) showing similar trajectory of resonance changes upon addition of presequence corresponding to L71 (Figure28) indicating the formation of complex *Tim23ims*-presequence is formed in *Tim23liposome* bound form and free*Tim23*.

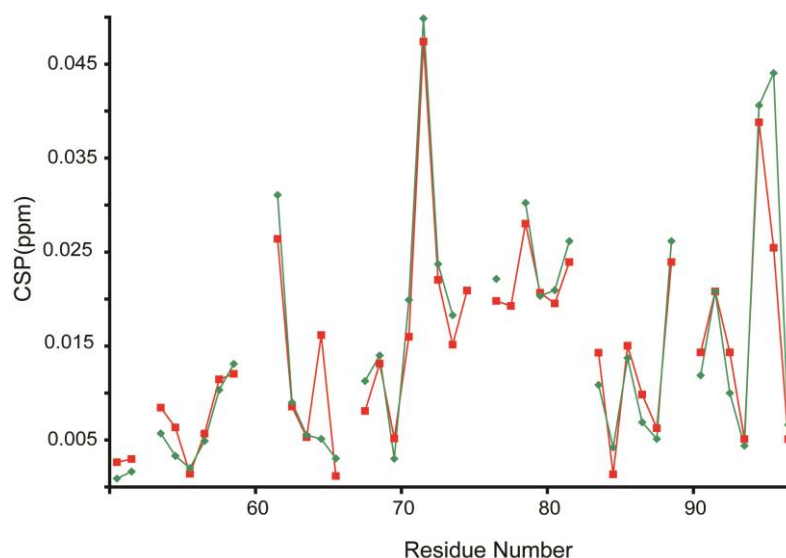


Figure 42: Average chemical shift perturbations (CSP) upon addition of 16 fold molar excess of presequence to liposome bound *Tim23ims* (red) and free *Tim23ims* (green) showing similar changes in magnitude of chemical shift perturbation (CSP) for C-terminal residues(50-96) of *Tim23ims*. This indicates that residues in presequence binding site of *Tim23ims* are affected in similar way and amount of presequence-*Tim23* complex is same in both cases. For clarity, only the C-terminal residues 51-96 of *Tim23ims* are shown.

3.5 Structural properties of intermembrane space domains of Tim21-Tim23

3.5.1 Localization of Binding sites of Tim21ims with Tim23ims

The subunits of TIM23 complex, Tim21 and Tim23 can be cross-linked *in vivo* (Tamura *et al.* 2009). It is possible that these subunits could interact with each other directly with their intermembrane space domains. However, the evidence for such a direct interaction between the intermembrane space domains of Tim21 and Tim23 is not known. To test the possibility of such a direct interaction between these two intermembrane space domains, ^1H - ^{15}N -HSQC based titration of ^{15}N labeled Tim23ims with unlabeled Tim21ims was performed. Upon addition of Tim21ims, progressive chemical shift changes and decrease in intensity were observed for the amide resonances of Tim23ims (Figure 43, A).

Upon addition of 16-fold molar excess of the Tim21ims (maximum concentration at which Tim21ims exist as monomer) resulted in chemical shift changes of above 0.02 ppm for the residues 4-7, 42-46, 69-79 and 88-95. In these regions, the largest chemical shift changes were observed for residues L4-G6, N40, I45, Y70-E74, S90-G92 and T94. Moreover, chemical shift changes were also observed for isolated residues Q33, G36, V53, L61 and A62. The residue 10-30 does not show any significant chemical shift changes. (Figure 43, B).

The overall decrease in intensity by ~30% was observed for resonances of Tim23ims upon addition of Tim21ims. This is evident from the decrease in intensity for the residues 10-30. Additionally, the residues 2-7 were broadened to an intensity ratio of ~0.4 whereas residues 69-94 were broadened to intensity ratio below 0.2 except for residues Q85 and I88 (Figure 43, C).

These changes in the backbone amide resonances in chemical shift position and signal intensity are indicative of fast to intermediate exchange at NMR time scale between Tim23ims-Tim21ims bound and free form of Tim23ims.

Cumulative analysis of the changes in intensity and the chemical shift perturbation in Tim23ims upon addition of Tim21ims resulted in three main perturbed regions as 2-7 and 69-79 and 90-95 (Figure 43, B). Specifically bulky hydrophobic resonances (W, L, F, I, Y) and neighboring resonances (G, S, D, E, and R) were affected all along the sequence of Tim23ims.

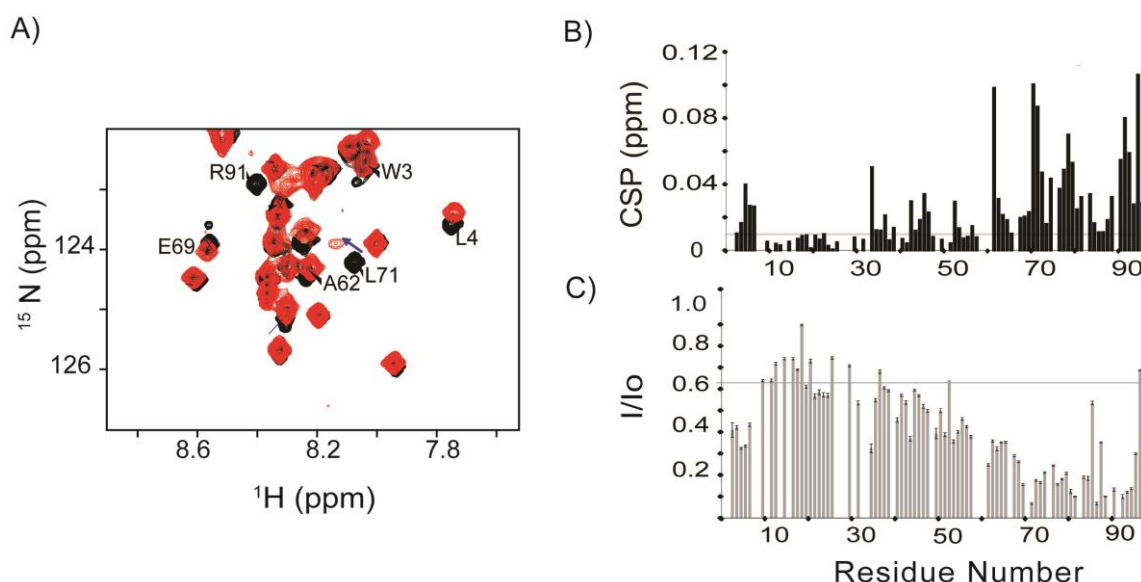


Figure 43: Interaction of Tim21ims with Tim23 ims. (A) Superposition of a section of 2D ^1H - ^{15}N HSQC of Tim23ims¹⁻⁹⁶ in the absence (black) and the presence of 16 fold molar excess of Tim21ims (red). The highlighted residues L4, A62, E69 and L71 exhibits both chemical shift changes and intensity attenuation. Average ^1H , ^{15}N chemical shift deviation (B) and intensity change (C) observed for backbone resonances of Tim23(1-96) in 2D ^1H - ^{15}N HSQC NMR spectra upon addition of 16 fold molar excess of unlabeled Tim21ims. The grey line in (B) and (C) represents the maximum chemical shift deviation and average intensity ratios for residues 10-30.

The direct interaction between the Tim23ims-Tim21ims is evident from the changes in the chemical shift and intensity in the backbone resonances of Tim23ims upon addition of Tim21ims. To further identify the binding interface of Tim21ims-Tim23ims complex and to understand the role of individual regions of disordered Tim23ims in the Tim23ims-Tim21ims complex, we did the reverse titration of ^{15}N labeled-Tim21ims with unlabeled Tim23ims.

The three constructs of Tim23ims, one involving the full length Tim23 (1-96) and the two smaller constructs (that showed maximum perturbation in ^{15}N -Tim23ims) corresponding to residues 1-13 (Tim23 1-13) and 61-96 (Tim23 61-96) were tested for their binding to ^{15}N Tim21ims. Upon stepwise addition of each of the construct of

Tim23ims to Tim21ims, the chemical shift changes were monitored in ^1H - ^{15}N -HSQC spectra.

The trajectory of chemical shift changes for all the three constructs of Tim23ims in ^1H - ^{15}N HSQC spectra of Tim21ims, is highlighted for residue L138 in Figure 46 A, B and C, which is similar for all the three constructs of Tim23ims tested. This indicates that the similar complex of Tim21ims-Tim23ims is formed irrespective of the construct length of Tim23ims titrated to Tim21ims. However, the magnitude of chemical shift changes in Tim21ims upon titration of shorter constructs Tim23 (1-13) and Tim23 (61-96) were different in comparison to full length Tim23 (1-96). Tim23ims (1-13) showed least magnitude of changes in chemical shift in Tim21ims. The maximum chemical shift changes in Tim21ims were induced upon binding to full length Tim23 (1-96) in comparison to shorter constructs Tim23 (61-96) and Tim23ims (1-13) (Figure 46 and Appendix Figure 6).

The changes induced upon addition of Tim23 (1-96) in Tim21ims were analyzed in detail. The binding curves for each of the perturbed residue of Tim21ims was plotted as the function of increasing concentration of Tim23 (1-96) to identify the residues present at binding interface. The hyperbolic binding curves were observed for the residues 138-144 (β -strand 1) and residues F109, V114, S114 and E117 (they lay on the same side of α -helix 1) whereas binding curves for nonspecific residues (M172, H169) lacks such hyperbolic pattern (Figure 46 and Appendix Figure 7).

To characterize the binding quantitatively, K_d values were obtained by fitting the curves. For Tim23 (1-96), the average K_d value of 152.57 μM for residues 138-142 was determined using one site binding model.

In case of Tim23ims (61-96), the binding curve for the most perturbed residues, L138 and A140 in Tim21ims showed K_d value of $360 \pm 65 \mu\text{M}$ and $432 \pm 143 \mu\text{M}$ respectively, that is approximately twofold higher than that of Tim23 (1-96). However for Tim23 (1-13) up to 8 fold molar excess addition to Tim21ims the chemical shift changes were significantly absent in ^1H - ^{15}N HSQC spectra of Tim21ims and thus, the K_d value could not be accurately determined.

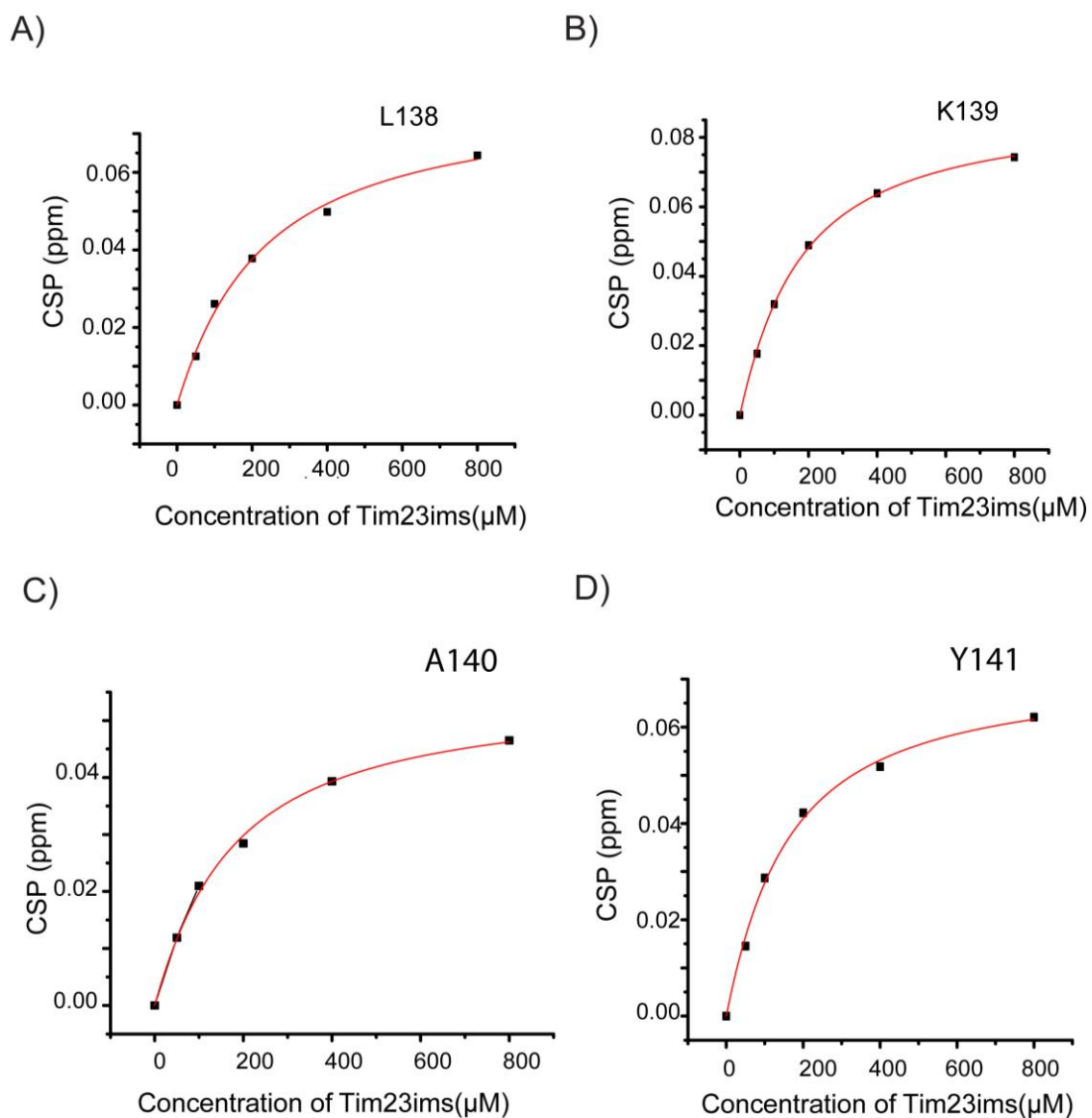


Figure 44: *The binding curves of residues of Tim21ims involved in binding to Tim23ims: Chemical shift perturbation (CSP) (in ppm) for residues 138-141 of Tim21ims as a function of increasing amount of Tim23 (1-96). Solid lines show the binding curves obtained by fitting to a one site model.*

Table 14: List of the K_d values obtained for the residues of Tim21ims upon binding to Tim23 (1-96). S.D means standard deviation obtained from fitting experimental values to the one site binding model.

| Residue | K_d (μ M) | S.D. (μ M) |
|---------|------------------|-----------------|
| L138 | 197 | 25 |
| K139 | 146 | 7 |
| A140 | 147 | 14 |
| Y141 | 130 | 14 |
| G142 | 141 | 7 |
| F109 | 114 | 27 |
| S114 | 233 | 43 |
| V113 | 238 | 8 |

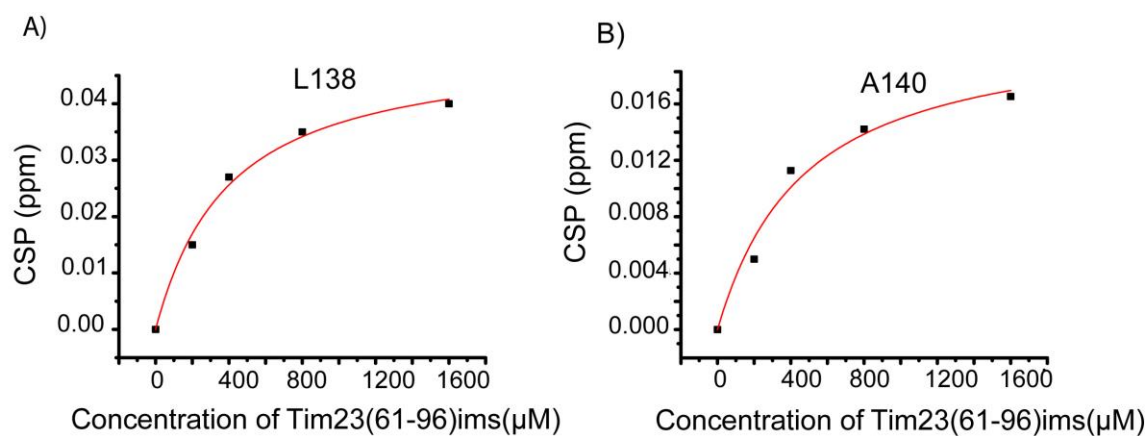


Figure 45: The binding curves for residues L138 and A140 of Tim21ims upon binding to shorter construct of Tim23 (61-96).

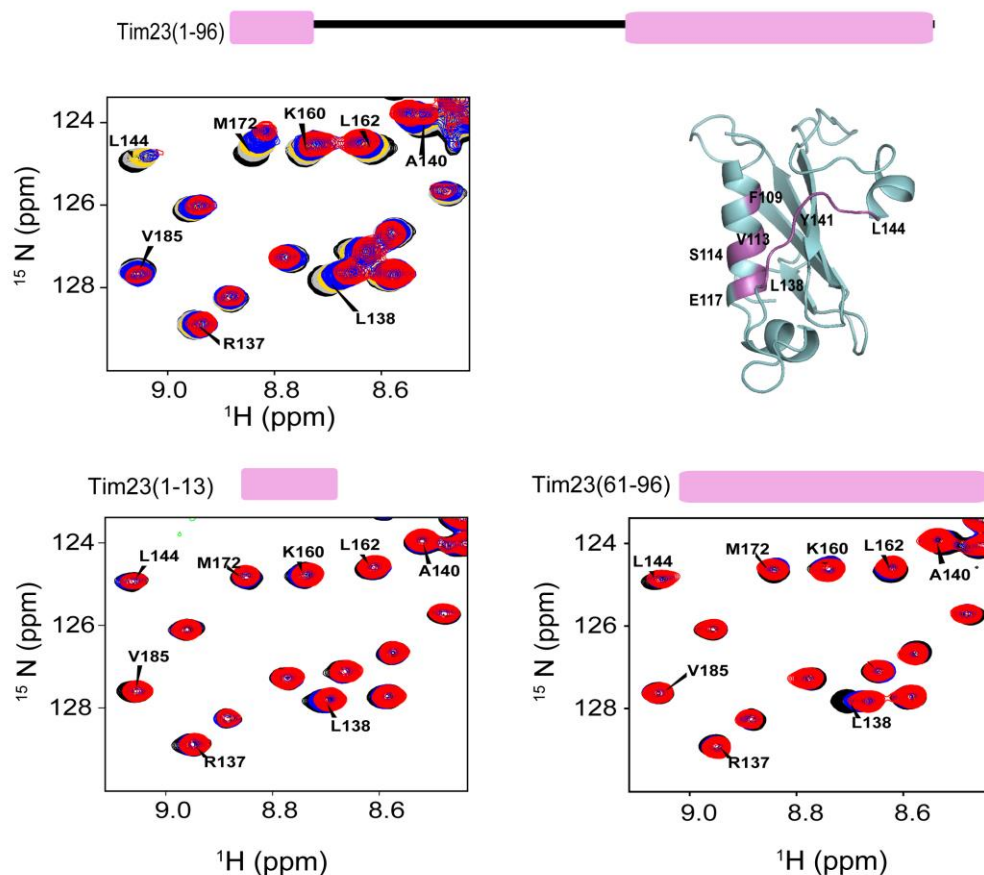


Figure 46: Interaction of various constructs of Tim23ims with ^{15}N -Tim21 ims. Each titrant construct of Tim23ims added to ^{15}N -Tim21ims is represented schematically above the panel with (A) Superposition of section of 2D ^1H - ^{15}N -HSQC of Tim21ims with increasing amount of Tim23 1-96 (0 fold excess : black, 1fold excess: grey, 2 fold excess: yellow, 8 fold excess: blue and 16 fold excess: red). The residues showing chemical shift changes and specific binding curves are highlighted on Tim21ims NMR model 1 with magenta color.(B) and (C) Superposition of section of 2D ^1H - ^{15}N HSQC of Tim21ims with increasing amount of Tim23 (1-13) and Tim23 (61-96) respectively (with reference: black, 8 fold excess: blue and 32 fold excess: red).

Additionally, the intermolecular paramagnetic relaxation enhancement effect (PRE) was employed to ascertain the binding site of Tim23ims in Tim21ims. The PRE tag (MTSL) was attached in Tim23 (1-96) at position close to the end of the interacting peptides of Tim23, at position 11 [from Tim23 (1-13)] and at position 67 [from Tim23 (61-96)].

Tim23 (1-96) tagged with MTSL at either position 11 (T11C) and 67 (G67C) was titrated to ^{15}N Tim21ims at 1:1 molar ratio. Both Tim23 (T11C) and Tim23 (G67C) exhibited similar broadening profile (i.e. ratio of NMR signal intensities in paramagnetic I_p and diamagnetic state I_d) in Tim21ims (Figure 47) with maximum broadening involved in region 138-144.

Interestingly, rest of residues of Tim21ims has an average intensity ratio (I_p/I_d) above 0.75 except for F200 and R114; presumably due to nonspecific interaction with MTSL.

Conjointly, chemical shift perturbation of three constructs of Tim23ims and two sets of intermolecular PRE data indicated following:

- (A) Each of peptide Tim23 (1-13) and Tim23 (61-96) can bind individually to same binding site in Tim21ims as Tim23 (1-96).
- (B) The binding elements of Tim23ims are present around the position 11 and 67 in individual peptides Tim23 (1-13) and Tim23 (61-96).
- (C) Tim23ims (61-96) binds with ~2 fold weaker affinity in comparison to Tim23 (1-96) into a single site on Tim21ims (involving residues 138-146 and residues 109, 114, 115, 117 of helix 1).

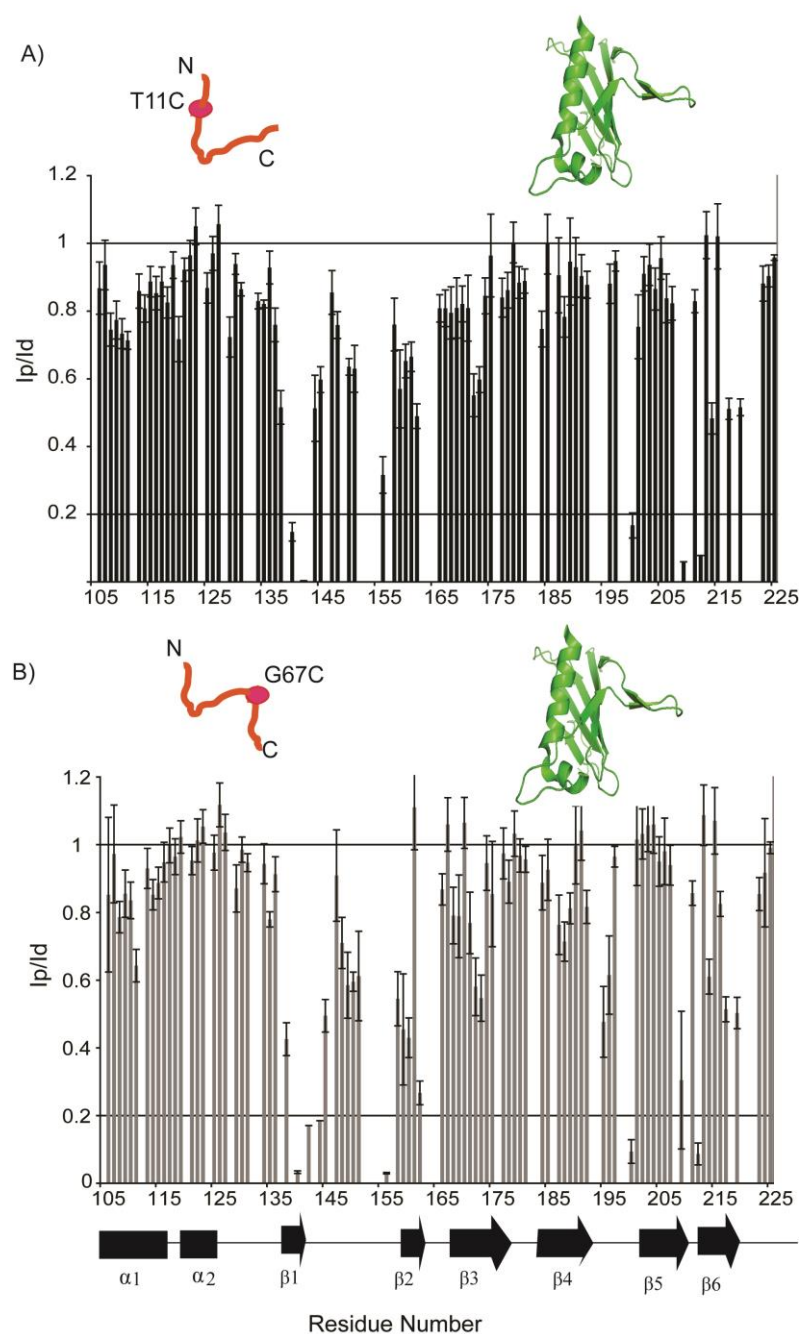


Figure 47: Intermolecular paramagnetic profile of Tim21ims upon addition of MTSL tagged Tim23ims at position (A) T11C as (black) and (B) G67C (grey). I_p/I_d is obtained from the intensity ratios from ^{15}N - ^1H HSQC Tim21ims with MTSL conjugated to Tim23ims at mentioned site (I_p) and addition of ascorbic acid to paramagnetic sample (I_d). The residues showing I_p/I_d (intensity attenuation) ratio values lower than 0.2 are within 5-10Å to the paramagnetic tag carrying Tim23ims. Spaces in the intensity attenuation profile correspond to prolines or residues that were overlapped in free Tim21ims. The schematic representation of the MTSL tag (as an oval) Tim23ims (as an extended chain) and Tim21ims (cartoon representation of structure) is shown in upper panel for both A and B.

3.5.2 Linear motifs of Tim23ims involved in binding to Tim21ims

In order to gain insights into the Tim21 binding region of Tim23ims, the intermolecular paramagnetic relaxation enhancement (PRE) effect was used. The intensity attenuation in paramagnetic state is observed if the two proteins are interacting. For the same, we used ^{15}N -Tim23ims and paramagnetic tagged Tim21ims (at below mentioned site of Tim21ims) at one fold molar excess. This intensity attenuation or signal broadening was observed due to the intermolecular interaction between Tim21ims and Tim23ims that causes paramagnetic relaxation enhancement of H^{N} and thus signal broadening for the residues of Tim23ims close to the paramagnetic tagged Tim21ims.

MTSL tagged S16C/C30A-Tim21ims was tested to identify the region of Tim23ims involved in binding to Tim21ims. The MTSL tagged S114C/C128A Tim21ims were added to one fold molar excess of Tim23ims. This resulted in paramagnetic broadening (an intensity ratio ($I_{\text{p}}/I_{\text{d}}$) at N-terminal residue 1-7 and C-terminal residues beyond residue 58 of Tim23ims to ~ 0.35 and below 0.2, respectively. The residues 35-55 also gets broadened to ~ 0.5 whereas the residues 9-34 remains unaffected ($I_{\text{p}}/I_{\text{d}} \sim 0.8$). This suggested that either C-terminal residues of Tim23ims lies closer to Ser114C/C128A Tim21ims than N-terminal 1-7 or C-terminal residues binds tighter than N-terminal 1-7 residues of Tim23ims.

In order to understand the binding of these two segments of Tim23ims, another position of Tim21ims for monitoring intermolecular PRE was selected i.e. native C128 that lies within 15 Å away from the S114.

The different paramagnetic tag i.e. nitrosyl radical based MTSL tag and paramagnetic lanthanide based Cys-Ph-TAHA tag (Peters et al. 2011) was attached to the native cysteine (residue128) of Tim21ims (Tim21ims C128A).

The broadening profile of Tim23ims due to the intermolecular PRE effect upon addition of Tim21ims C128 with the two different paramagnetic tags resulted in similar regions i.e. N terminal 1-11 and C-terminal 65-96 of Tim23ims. Within the C-terminal region 65-96 the residues 74-88 are less affected for both the tested tags. However, Cys-Ph-TAHA tagged Tim21-C128 caused severe broadening for residues involving 9-12 and C-terminal residues 90-95 in comparison to Tim21-C128-MTSL.

The two PRE tags differ in the length of molecule, the paramagnetic effect range and possess the differential mobility of the tagged-Tim21ims. This could explain the observed differences in the broadening profile for Tim23ims upon addition of tagged Tim21-C128.

Thus, three different data sets of intermolecular PREs of Tim23ims with Tim21ims were measured that cumulatively showed

A) The C-terminal residues of Tim23ims involving residues 61-96 were differentially perturbed with near complete line broadening in case of S114C *i.e.* S114C is much closer to binding site of Tim23 than C138A.

B) The residues 15-26 are unperturbed.

C) The N terminal residues 1-7 of Tim23ims are equally perturbed with MTSL loaded at either position of C128 or S114 of Tim21ims to an intensity ratio of 0.4

In summary, most of the polypeptide of Tim23ims is differentially perturbed with increasing effect with 1-13, 67-74 and 90-96 respectively and collectively, above discussed three intermolecular PRE data suggests three motifs 1-13, 67-76 and 90-96 of Tim23ims are involved in binding to Tim21ims. Three linear motif of Tim23ims also complement binding of individual peptides of Tim23 (1-13) and Tim23 (61-96) to Tim21ims.

Upon primary sequence analysis for Tim21 binding elements in Tim23ims, the presence of bulky hydrophobic and negatively charged residues ²SWLFGD⁷, ⁶⁷GVEYLDL⁷³, and ⁹⁰SRGWTDD⁹⁶ were observed that could be critical for interacting to Tim21ims.

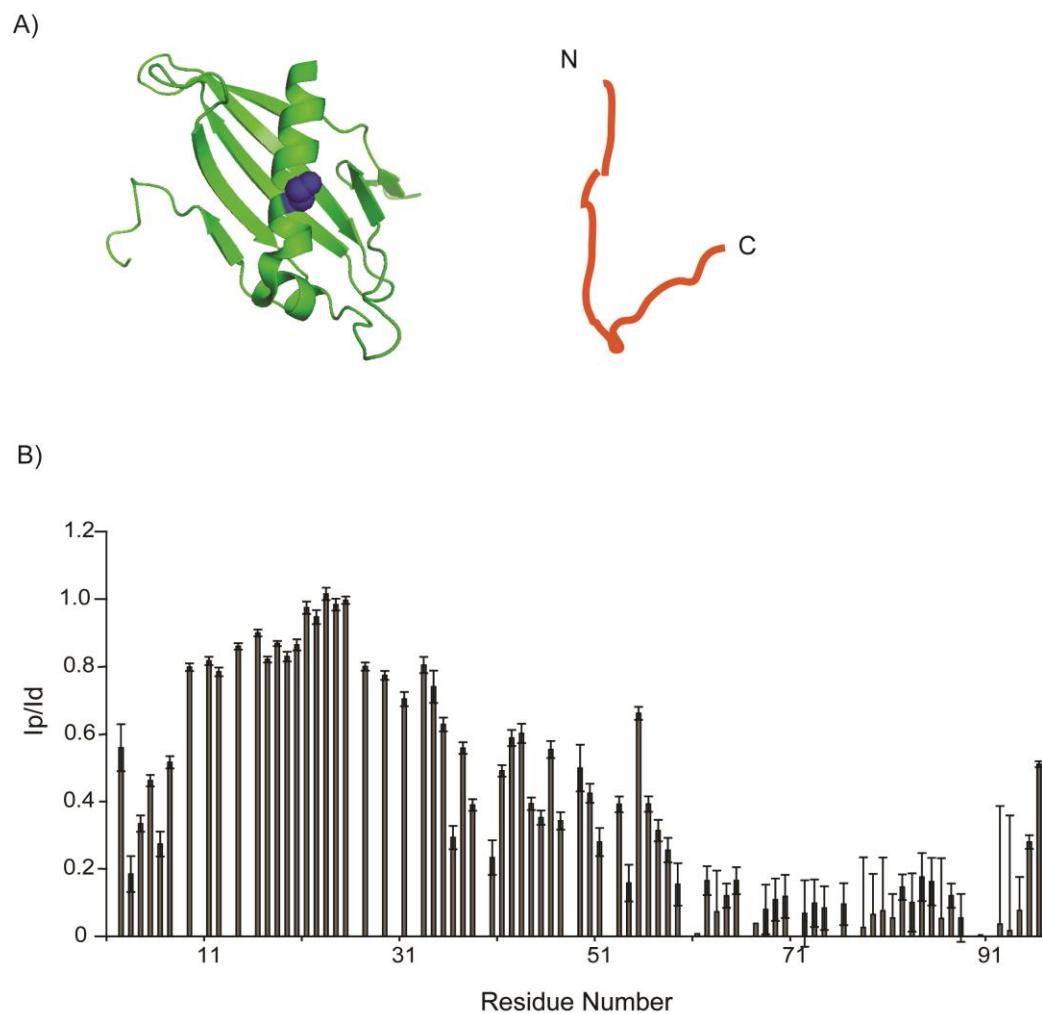


Figure 48: (A) The cartoon representation of Tim21ims (green) and Tim23ims (red) sample at 1:1 ratio used for measuring the intermolecular PREs. The paramagnetic tag (nitoxide spin label carrying MTSL) in Tim21ims S114C is conjugated with Tim21-S114C, side chain of S114 is represented in blue spheres (B) PRE (Intensity attenuation) profile of Tim23ims as a function of its sequence due to intermolecular paramagnetic relaxation enhancement of Tim23ims upon addition of 1 molar excess of Tim21ims at position S16C with MTSL (black). Intensity attenuation is a proof for transient binding of Tim21ims with Tim23ims.

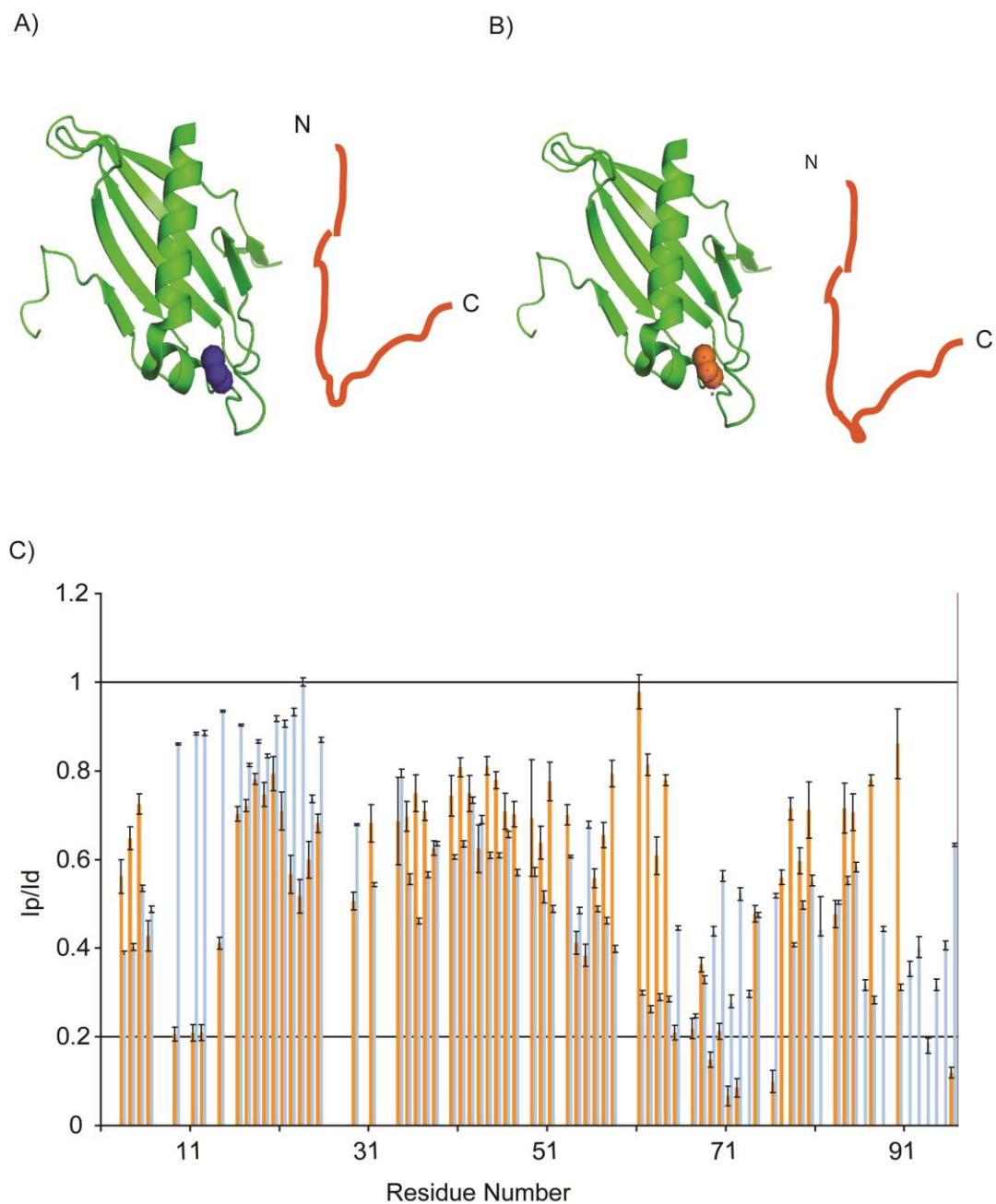


Figure 49: Two samples used for determination of intermolecular PRE based interaction between Tim21 and Tim23 (1:1ratio). The cartoon representation of Tim21ims (green) conjugated with paramagnetic tag and Tim23ims (red). The paramagnetic tag is conjugated at position C128 in Tim21ims is marked with spheres in blue for MTSL (A) and in orange for Cys-Ph TAHA (B). (C) PRE (Intensity attenuation) profile of Tim23ims as a function of its sequence due to intermolecular paramagnetic relaxation enhancement of Tim23ims upon addition of a molar excess of Tim21ims conjugated at position C-128 either with paramagnetic tag Cys-Ph TAHA (orange) and MTSL (blue). Intensity attenuation is proof for transient binding of Tim21ims with Tim23ims.

3.5.3 Model for Tim21-Tim23 complex

Intermolecular NOEs are most commonly used restraints for determination of protein complexes using NMR. The limited solubility of Tim21ims-Tim23ims complex and high K_d (150 μ M) prevented the use of intermolecular NOEs as restraints. The CBCACONH of Tim23ims in excess of Tim21ims (Tim23: Tim21:: 0.15mM: 1.2mM) was measured. The correlation peaks for residues of Tim23ims involved in binding to Tim21ims, which undergoes intermediate exchange on the NMR timescale between Tim23 free and bound state (that were broadened in ^1H - ^{15}N HSQC), could not encode the transfer of the magnetization from its H^{N} to C^{α} and C^{β} of the previous amino acid and hence were absent in the spectra. However, observable encoded residue showed no significant changes in C^{α} chemical shift indicating that non-interacting regions of Tim23ims remains disordered in Tim21ims bound state.

Further, ^1H - ^{15}N HetNOE of Tim23ims in complex with 16 fold excess of unlabeled Tim21ims was measured that shows an average value for all residues below ± 0.4 indicating that Tim23ims remain dynamic at nanosecond to picosecond time scale in its bound state. On comparing the ^1H - ^{15}N Het NOE values of Tim23ims in bound state (in complex with Tim21ims) with Tim23ims in free state, it was found that C-terminal residues (72-96) showed small yet significant changes. The above results suggest that the Tim23 remains dynamic in the Tim21 bound state with local ordering in the Tim21 binding region.

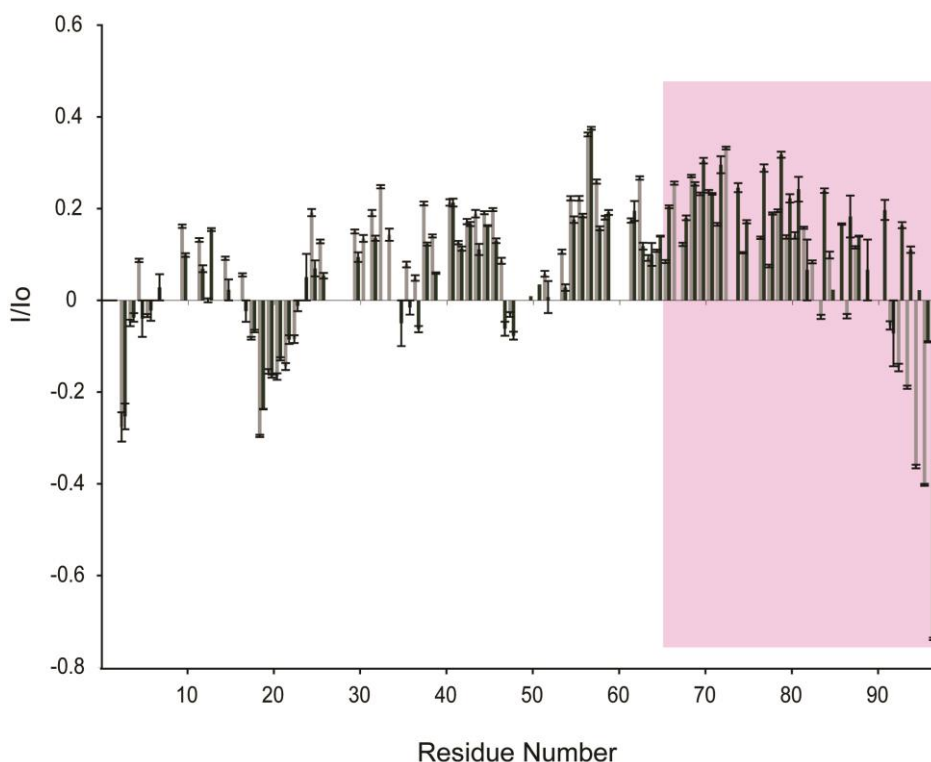


Figure 50: **Tim21 bound Tim23ims is highly dynamic:** ^1H - ^{15}N -HetNOE values of Tim23ims in 16 fold excess of Tim21 (grey) and no excess (black) as a function of residue number of Tim23ims. The region at C-terminus, which shows deviations, is highlighted with pink with box.

In order to determine the ensemble view of binding of the above determined three linear motifs namely peptide 1 (residue 2-7, SWLFGD), peptide 2 (residue 67-74 GVEYLDLE) and peptide 3 (residue 90-96, SRWTDD) to a single binding site in Tim21, Rosetta FlexPepDock was used to dock them. Rosetta FlexPepDock can provide the complete flexibility to peptide and the receptor at the binding interface and can be used to refine the interface residues if the approximate binding site is known (Raveh et al. 2011). Indeed, this was the scenario with Tim21-Tim23 where individual binding peptides of Tim23 were known along with binding site in Tim21ims but high resolution binding interface of Tim21-Tim23 complex needs to be defined.

We performed the *ab-initio* docking and results for the Tim21-peptide 1 complex is based on the top scoring rosetta models whereas for peptide 2 and 3 Tim21 peptide complex, 25000 conformers were obtained and based on rosetta energy score the top 500 models were clustered with a backbone RMSD of 2 Å and binding interface of each cluster was analyzed. The first five clusters for peptide 2 have 17, 11, 8, 5, and 6 conformers whereas peptide 3 has 58, 58, 37, 41 and 29 conformers.

The results of docking (Figure 51) revealed that the consensus interactions are formed involving the key residues K139 and Y141 in Tim21 with a negatively charged residue of Tim23 (D or E) and the hydrophobic residue (L) or aromatic residue (W) respectively. The Y 141 acts as a main hydrophobic anchor that interacts either to the leucine residue or to the aromatic residue of Tim23ims. Additionally, same peptide can also have the switch of hydrophobic anchor L to Y and W to Y interactions (peptide 1, Figure 51-lower panel).

Furthermore, these peptides are also accommodated at the binding site of Tim21ims with the help of π - π interactions with side chain of F109 if three bulky hydrophobic residues are present for example in peptide1 and 2 that has the WLF and EYLDL respectively.

All the three peptides interact to a single binding site and suggests that the peptides of Tim23ims can dynamically bind in the small hydrophobic cleft of Tim21ims with K139 and Y114 mediating the key interactions. Different clusters of peptide 2 and 3 converge well indicating that the conformer represented in Figure 51B and C are most probable near native solution for these binding motifs in Tim21-Tim23 complex. However for peptide 1, the switch of L to F to interact with Y141 of Tim21ims was observed. Thus, interaction of peptide 1 with Tim21 cannot be described by single conformation rather can be viewed as the presence of consensus interactions in multiple conformers with each of them representing a member of cluster (Figure 51 A, lower panel). In peptide 3, K139 and Y141 of Tim21 interact with W conferring cation- π and π - π interactions at the interface of Tim21-Tim23 complex.

In conclusion, three linear motifs of Tim23 can be structurally adapted in a shallow single binding site of Tim21ims and residues K139 and Y 141 of Tim21 ims plays an important role at binding interface.

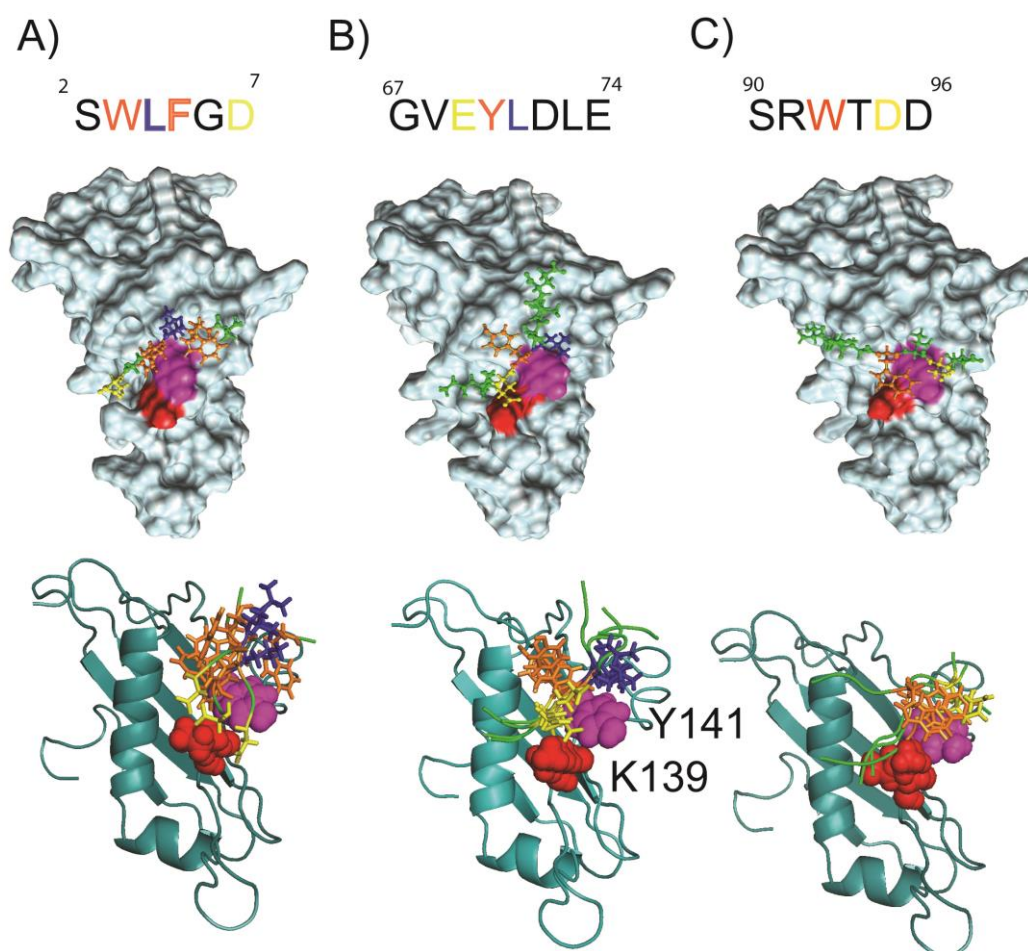


Figure 51: Model of Tim23ims-Tim21ims complex derived from docking Tim23 peptides to Tim21ims using Rosetta FlexPepDock. Three interacting peptides of Tim23ims based on chemical shift and intermolecular PRE data of Tim21ims and Tim23ims were placed near the binding site in Tim21 (10Å away). Column (A), (B) and (C) represents the results of three different peptides of Tim23 used for docking to monomeric lowest energy NMR structure of Tim21ims. Upper panel represents the peptide sequence and residue number of Tim23ims and the important residues at the docking site is color coded with leucine as blue, aromatic residues as orange and negatively charged residue as yellow. Middle panel shows the surface representation of binding interface of Tim21-Tim23 complex with Tim21ims (light blue) and docked Tim23 peptides. Lower panel shows ribbon representation of binding interface of Tim23 peptides in Tim21ims as elucidated from the lowest energy member of top 5 clusters derived from Rosetta FlexPepDock. The side chains of the key residues of Tim21, K139 and Y141 are shown in red and magenta spheres respectively and important side chains of Tim23ims are shown as sticks and color coded described previously.

3.6 Mapping Tim50ims interaction sites on Tim23ims

Direct interaction of Tim23 with another subunit of presequence translocase, Tim50ims has been studied and was suggested to involve residues 70-71 of Tim23ims (Gevorkyan-Airapetov et al. 2009; Tamura et al. 2009; Marom et al. 2011). Similar to Tim21ims, the effect of stepwise addition of unlabeled Tim50ims¹⁶⁴⁻³⁶¹ to ¹⁵N-Tim23ims were monitored using ¹H ¹⁵N-HSQC spectra. During the titration of Tim23ims with Tim50-¹⁶⁴⁻³⁶¹, differential signal broadening was monitored along the sequence of Tim23ims at each excess molar ratio of Tim50ims tested.

At 12 fold molar excess of Tim50ims, the maximum signal broadening was seen in the region 56-78, which got almost completely broadened to near zero value. Furthermore, N-terminal residues 1-7, were broadened to the average intensity ratio of 0.38 and the additional regions residues 36-53 and 83-94 were broadened to an average intensity ratio 0.6. The residues 11-34 were least perturbed with average intensity of 0.85.

Thus, increasing the concentration of Tim50ims¹⁶⁴⁻³⁶¹ caused the resonance line broadening of residues in Tim23ims without involving any change in the chemical shift. This is indicative of slow to the intermediate exchange regime for the residues of Tim23 involved in binding Tim50ims.

The signal broadening monitored upon addition of Tim50ims is a clear indication of binding of Tim23ims. Generally, resonance line broadening originates from two possible mechanisms, firstly the overall broadening due to an increase in the correlation time of a complex and secondly, selective broadening due to chemical exchange of residues belonging to interaction site. The overall broadening pattern is evident in the changes in the intensity of residues of least perturbed region 11-34 and selective broadening is seen in regions with residues 1-7, 56-73, and 83-94.

On comparing the intensity attenuation profile for Tim23ims with Tim50ims, it is evident that residues 56-73 are most affected and constitute the prime high affinity site of Tim23ims for binding to Tim50ims. However, residues 1-7 and 83-94 of Tim23ims can additionally contribute to the affinity of Tim23-Tim50ims complex.

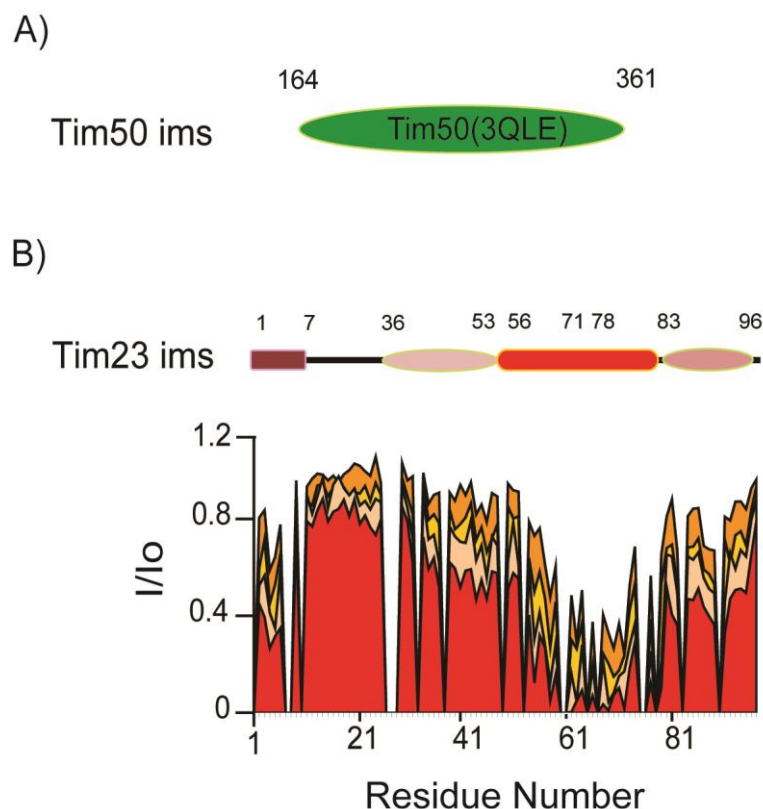


Figure 52: *Tim23ims* interacts with *Tim50ims*. (A) The schematic representation of *Tim50ims* construct involving residues 164-361(*Tim50ims*¹⁶⁴⁻³⁶¹) for which the crystal structure is available (PDB id: 3QLE) is shown. (B) The overlay of a region of ¹H-¹⁵N HSQC of *Tim23ims* without *Tim50ims* (black) and with 12 fold molar excess of *Tim50ims* (Red). (C) The intensity attenuation plot of *Tim23ims* upon addition of 2 fold, 4 fold, 8 fold and 12 fold molar excess for *Tim50ims*¹⁶⁴⁻³⁶¹ are shown in dark orange, light orange, skin color and red, respectively. The upper panel shows a schematic representation of the intensity attenuation profile along the sequence of *Tim23ims* that is color coded with gradient of red and red represents the most perturbed residues.

4 Discussion

4.1 Intermembrane space domains of mitochondrial translocases are highly dynamic with diverse structural properties

The translocation of the mitochondrial-matrix targeted preproteins involves transversing across three different compartments of mitochondria to reach their site of action in the mitochondrial matrix. The translocation of matrix-targeted preproteins is mediated by the hetero-oligomeric translocases in outer and inner mitochondrial membrane (TOM and TIM). The various components of these translocases (such as Tom22, Tom40, Tom5, Tom7, Tim23, Tim17, Tim50, Tim21, Pam17 etc.) are known to possess the soluble intermembrane space domains. Previous biochemical and genetic analyses in yeast have shown that the role of intermembrane space domains is three-fold: (a) as presequence receptor site, (b) as a part of the translocation contact site, (c) and in regulating the pore across the inner membrane. However, the molecular details pertaining to the oligomeric status, structural properties and the interaction sites of each of the intermembrane space domain is not well-characterized.

In vivo and *in vitro* crosslinking studies have also suggested the binding proximity of various intermembrane space domains such as Tim23-Tim50, Tim21-Tim23, Tom22-Tim50 and TOM40-Tim23, Tim17-Pam18, Pam17-Tim23 among others (Chacinska et al. 2005; Hutu et al. 2008; Tamura et al. 2009; Marom et al. 2011). The incoming preprotein can also be cross-linked to the ims domains of above said subunits. Of the various intermembrane space domains of the presequence translocases, only the ims domains of Tim21 and Tim50 have been structurally characterized (Albrecht et al. 2006, Qian et al 2011). Despite the availability of structures of these individual ims domains, the structural details of these ims domains in complex with

their interacting partner subunits still remains elusive. Furthermore, the exact composition of each of the component in the presequence translocase is not known. The lack of detailed structural studies of intermembrane space domains in bound form and presequence-receptor complexes has largely limited our understanding of how the various events during protein sorting and translocation into mitochondria occurs.

The functional form of the presequence translocase complex has been proposed to comprise a highly dynamic complex consisting of multiple subunits that adopt multiple conformations and involve in diverse interactions to facilitate preprotein import (Chacinska et al. 2005; Popov-Celeketic et al. 2008, Tamura et al, 2009, Chacinska et al. 2010; van der Laan et al. 2010; Marom et al. 2011). The tendency of the various subunits to adopt multiple conformations renders them intrinsically dynamic and poses a huge challenge to their structural characterization *in vitro*. The use of X-ray crystallography is limited by the production of crystals for proteins that are highly dynamic, largely unstructured and are involved in transiently weak interactions. Given these properties of the proteins involved in mitochondrial transport, NMR spectroscopy emerges as a viable and powerful alternative for the structural characterization. Furthermore, it can also be used to study the dynamic properties of proteins under physiological conditions. However, NMR studies require the large scale production of isotopically rich samples and optimization of the solution condition prior to their structural analyses.

With the aim of understanding the molecular details and to gain a deeper understanding of various ims domains both in the presence and absence of their targets, we have initiated structural studies on these domains. As a first step towards this goal, we have cloned, expressed, purified and isotopically labeled the Tom22ims, Tim21ims, Tim50ims and Tim23ims. The important functions of each intermembrane space domains are listed in Table 2.

4.1.1 Tom22ims is largely disordered and possess a transient helix in solution

A thirty three residues long peptide corresponding to the intermembrane space domain region of Tom22¹²⁰⁻¹⁵² has been reported to be essential for a stability of super complex involving subunits of the outer and inner membrane translocases during the preprotein import (Chacinska et al. 2003). It can also be cross-linked to presequence and

Tim50ims *in vivo* and has been proposed to be a part of *trans* site across OMM for incoming presequence carrying preprotein (Moczko et al. 1997; Shiota et al. 2011).

Additionally, peptide of Tom22ims corresponding to region 131-147 has been suggested crucial for its binding to Tim21ims (Albrecht et al. 2006). Hence, it is important to structurally characterize Tom22ims to gain deeper understanding for the basis for multifunctional role of Tom22ims. To this end; we have used a combination of NMR and CD spectroscopy in section 3.1.6, and have unambiguously established that Tom22ims is largely unstructured in solution. However, residues 120-127 have propensities to transiently populate α -helix (Figure 27). It is important to note that the lack of stable α -helix could be due to the absence of additional factors including its interacting partners, ligands such as membrane etc. required for the stabilization of the helix.

4.1.2 Tim21ims exhibits monomer-dimer equilibrium in solution and monomeric Tim21ims is highly dynamic in solution

Previous studies have characterized the crystal structure of monomeric Tim21ims (PDB id 2CIU). Our studies (section 3.1.4); have shown that Tim21ims undergoes monomer-dimer equilibrium under solution conditions using the size exclusion chromatography. Tim21¹⁰³⁻²²⁵ims exists predominantly in monomeric form up to a concentration of 1mM in solution, which is in accordance with less than 10% of dimeric species observed on the size exclusion chromatography. The resonances belonging to residues of β -strand 2 and 3 of Tim21ims also show chemical shift changes as a function of concentration (Figure 21). ¹⁵N-R₂ relaxation rates suggest that β -strand 2, 3 and 4 belonging to beta sheet region of Tim21ims undergo slow exchange at millisecond to microsecond time scale (Figure 23). Thus, it is plausible that the slow exchange could arise from the exchange between the monomeric and dimeric state of Tim21ims and the residues undergoing chemical exchange could also be important for the biomolecular interaction. Our attempt to further characterize the dimeric state of Tim21¹⁰³⁻²²⁵ was hampered by the low population of dimeric state in solution.

¹⁵N relaxation analysis along with our NMR structure of Tim21ims reveals another striking difference between solution and crystal structure of Tim21ims in region involving β 1* and β 2* strands (Figure24 and 25). The ¹H-¹⁵N Het NOE values and ¹⁵N-

R_2 relaxation rates revealed that residues in $\beta 1^*$ and $\beta 2^*$ are highly flexible (Figure 23). This observed flexibility in $\beta 1^*$ and $\beta 2^*$ strands is in accordance with the higher RMSD (up to 2.4Å) values among 20 lowest energy conformers determined by NMR spectroscopy. The residues in this region (146-154) also exhibit higher B factors in its crystal structure and were involved in extensive crystal contacts (Figure 25).

In conclusion, Tim21ims is highly dynamic and exhibit motion at wider time scales for different structured regions.

4.1.3 Tim50¹⁶⁴⁻³⁶¹ims is well folded with distinct dynamics

Tim50ims¹⁶⁴⁻³⁶¹ is well-folded and exhibits significant amide-proton chemical shift dispersion in ^1H - ^{15}N TROSY-HSQC spectra (section 3.1.5). This is in agreement with the X-ray structure of Tim50ims reported (Qian et al. 2011). It is interesting to note that Tim50ims exhibited variation in resonance intensities in the ^1H - ^{15}N TROSY-HSQC spectrum, much like Tim21ims. This is in line with the observed differences in B-factors at the Tim23 and presequence binding sites in the crystal structure of Tim50ims (Qian et al. 2011). Moreover, it has been observed that 50% of residues in Tim50ims (164-361) belong to flexible loops in the structure, although in the loop regions. Moreover 30 resonances were absent in the ^1H - ^{15}N TROSY-HSQC spectrum of Tim50im which could presumably arise from conformational exchange on the microsecond to millisecond timescale (Figure 26). At current stage, the absence of resonance assignments for Tim50ims hindered the detailed relaxation analysis as in case of Tim21ims. From the present studies, it can be concluded that Tim50ims undergoes conformational changes, which might play an important role in interaction with either the presequence or Tim23.

4.1.1 Tim23ims is intrinsically disordered in solution

Our studies in section 3.1.3, have shown that the first 96 residues at the N-terminus of Tim23 constituting the soluble intermembrane space domain globally lack secondary structural elements as evident from CD spectroscopy (Figure 18). Further evidence to support this comes from the poor amide proton chemical shift dispersion

and sharp resonances in the ^1H - ^{15}N HSQC of Tim23ims, characteristic for intrinsically disordered proteins.

To gain insights into the transient secondary structural elements in disordered Tim23ims, the residue specific secondary chemical shifts were obtained (Figure 19). The secondary chemical shift is highly skewed to the usage of the random coil database and thus provides semi qualitative relationship with the availability of transient secondary structural elements. However, secondary chemical shift analysis of Tim23ims revealed that only residues 73-87 have stretch of positive average values and hence could transiently populate the helix. This region also includes the predicted site for coiled-coiled interactions (residues 71-78), which has been suggested to be important for dimerization (Bauer et al. 1996). Additionally, ^1H - ^{15}N Het NOE for this region is low (less than ± 0.2) indicating that this region is flexible on the nanosecond to picosecond timescale. NMR parameters report on the ensemble average of protein molecules present in solution. Thus, with these results Tim23ims could be envisioned as an ensemble in which Tim23 conformers dynamically undergo rapid exchange and a small fraction of conformers in ensemble displaying helix. Furthermore, our data of paramagnetic labeling of Tim23 at N terminus (T11C) indicates the presence of transient long range interaction between N-terminal and C terminal residues in Tim23ims (Figure 20).

Moreover, ^1D (HN) in Pf1 phages surprisingly shows sign inversion for N-terminal residues in comparison to the rest of the polypeptide. The possible cause of change in sign for these residues is not known in free form. However, the same residues adopt turn upon binding to DHPC micelles (details in section 4.4.2).

In addition, A17-T23 was also seen to exhibit striking flexibility in ^1H - ^{15}N Het NOE similar to both N and C- terminus, indicating its role as a linker to connect N-terminal residues to C-terminal residues.

The aforementioned solution properties of intermembrane space domains of translocases indicate protein dynamics at various NMR timescales that could play important role in protein-protein interaction and precludes their study using X-ray crystallography.

4.2 Presequence recognition and translocation through intermembrane space for mitochondrial matrix-targeted preprotein

4.2.1 Presequence binding sites in intermembrane space domains of mitochondrial translocase

Receptor presequence interactions are required to recognize and sort the incoming preproteins to their functional site in mitochondria (Pfanner 2000; Mokranjac and Neupert 2005; Chacinska et al. 2009). Of the four available presequence receptor sites (i.e. *cis* site at the outer mitochondrial membrane translocase, *trans* site of outer mitochondrial translocase, *cis* site (ims facing) for the inner mitochondrial membrane translocase and the *trans* site (matrix facing) for the inner mitochondrial membrane translocase), the presequence receptors sites in the intermembrane space (*cis* and *trans* sites) are not well characterized. We therefore, focused on the presequence-receptor interactions in the intermembrane space of mitochondria of yeast using NMR spectroscopy (section 3.2). We have identified the direct presequence binding sites at a residue specific level in the ims domains of Tom22, Tim23 and Tim50. Moreover, we have also shown that Tim21ims lacks a presequence binding site.

In this study, (section 3.2.1) we identified the presequence receptor site in Tim23ims to be residues 71-84. The affinity of Tim23ims to the presequence is estimated to be in the millimolar range ($\sim 0.5\text{mM}$) (Figure 28). The presequence binding site of Tim23ims (residues 71-84) can adopt helical conformation to interact with amphipathic presequence. The presequence binding site in Tim23ims contains three charged residues (E74, E75 and E76) and three leucines (L71, L73 and L78). In order to gain insights into the Tim23ims-presequence complex, the helical wheel projections were determined and correlated with NMR titration data. The largest chemical shift changes were observed for residues L71, E74, E75, L78, and L81 that lie on the same side of the helical wheel projection of presequence binding site of Tim23ims suggesting that residues 71-84 of Tim23ims in a helical conformation with presequence (Figure 28 and 53). In the Tom20-presequence complex, hydrophobic interactions were suggested to be important for presequence receptor complex (Abe et al. 2001).

To identify the common features and differences in the presequence receptor complex at the TOM and TIM complex, we tested the interaction of mutant L71A/L78A Tim23ims with presequence and showed that indeed L71 and L78 are important for binding to presequence. Also, the residues in presequence binding site in Tim23 exhibit backbone dynamics at picosecond to nanosecond timescale (Figure 29). Although our data suggests that the hydrophobic leucines are important for binding to the presequences, one could not rule out the importance of the charged residues E74-76.

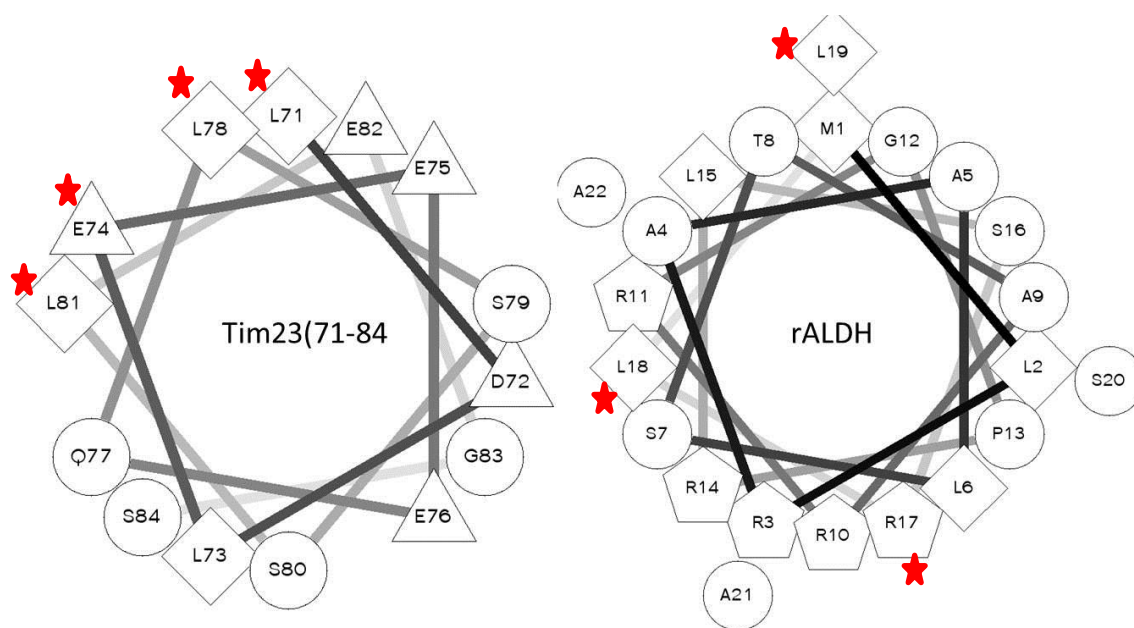


Figure 53: Helical wheel projections of the presequence binding site of Tim23ims and presequence rALDH. Amino acids are outlined according to their properties. Hydrophilic residues are represented by circles; hydrophobic residues are represented as diamonds negatively and positively charged residues are represented as triangles and pentagons respectively. Based on CSP, residues that shows largest chemical shift perturbation in Tim23ims upon addition of presequence and residues of rALDH that alter the binding affinity of Tim23ims-presequence complex are marked with red stars. Helical wheels were created using 'Wheel', created by D. Armstrong and R. Zidovetzki.

To further gain insights into the presequence receptor complex, we performed NMR titrations of truncated rALDH and its mutants with Tim23ims. Our titration results (Table 13) for N-terminal (1-11) and C-terminal (12-22) peptides of rALDH presequence show that the binding affinity to Tim23ims is reduced in comparison to the full length rALDH. This indicates that both N-terminal and C-terminal fragments of rALDH presequence have important role in the Tim23ims-presequence complex. This is in agreement with the *in vitro* import studies where the import of rALDH is abolished if either of the two segments of rALDH is removed (Wang and Weiner 1993).

Furthermore, our binding studies of the presequence mutants, rALDH R17Q and rALDH L18/L19Q showed that both mutants bind weaker to the Tim23ims in comparison to the wild type rALDH presequence. Hence, our data supports the binding chain hypothesis wherein both charged as well as hydrophobic interactions contribute to the affinity of presequence-Tim23 complex.

We also showed that in rALDH presequence mutant R17Q affects the binding to Tim23ims. R14Q or R17Q mutation in rALDH presequence was previously shown to be imported with the similar efficiency as that of wild type presequence *in vitro* whereas double mutant R14Q/R17Q showed reduced import to 50% than wild type presequence. Interestingly, it has also been shown that processing efficiency of preprotein carrying mutant presequence R17Q as well as R14Q/R17Q was affected in mitochondrial matrix (Hammen et al. 1996).

The impaired binding of R17Q rALDH to Tim23 could support the view that different check points in the mitochondrial preprotein import exists that additively provides quality control for the incoming preprotein. Hence, the binding of the double mutant R14Q/R17Q needs to be tested to understand the role of positive charge residues in Tim23-presequence complex formation.

It is known that -2 R or -3 R i.e. the presence of Arginine at position -2 or -3 from the C-terminus of a presequence is a prerequisite for processing preprotein to its functional form into the matrix by mitochondrial peptidase (Gavel and von Heijne 1990). Therefore, multiple presequences should be tested with mutation at -2 R or -3 R position to couple the binding efficiency of presequence-receptor complex with the processing activity of mitochondrial peptidase.

Previous studies have shown that Tom22ims is an important constituent of the presequence receptor at the *trans* site of TOM complex (Court et al. 1996; Komiya et al. 1998). Moreover, Tom22ims has been shown to act as a crucial presequence receptor site in absence of cytosolic receptors when the preproteins are blocked at the outer membrane in absence of membrane potential (Moczko et al. 1997). To this end, we have performed the NMR titrations of Tom22ims with presequence and identified the residues 123-133, which can form a transient helix in unbound state as the interaction site with incoming presequence (Figure27). The binding affinity of presequence-Tom22ims complex was found to be approximately 0.9mM. The presequence are

believed to bind their receptors by helix-helix interactions via either the hydrophobic face or electrostatic face of their amphipathic helix. Within the binding site of Tom22, the binding residues show the clustering of negatively charged residues on one side of helix (E120, E127 and D131 in Figure 54). It was believed that the electrostatic interaction mediate the binding of the presequence to the Tom22ims (Komiya et al. 1998). However, we observed large chemical shift changes in the binding site for L123, E125, M126, K128, T129 and F130 residues as well, suggesting that the combination of hydrophobic and electrostatic interactions play an important role in the presequence-Tom22ims receptor complex as observed for Tim23ims-presequence complex.

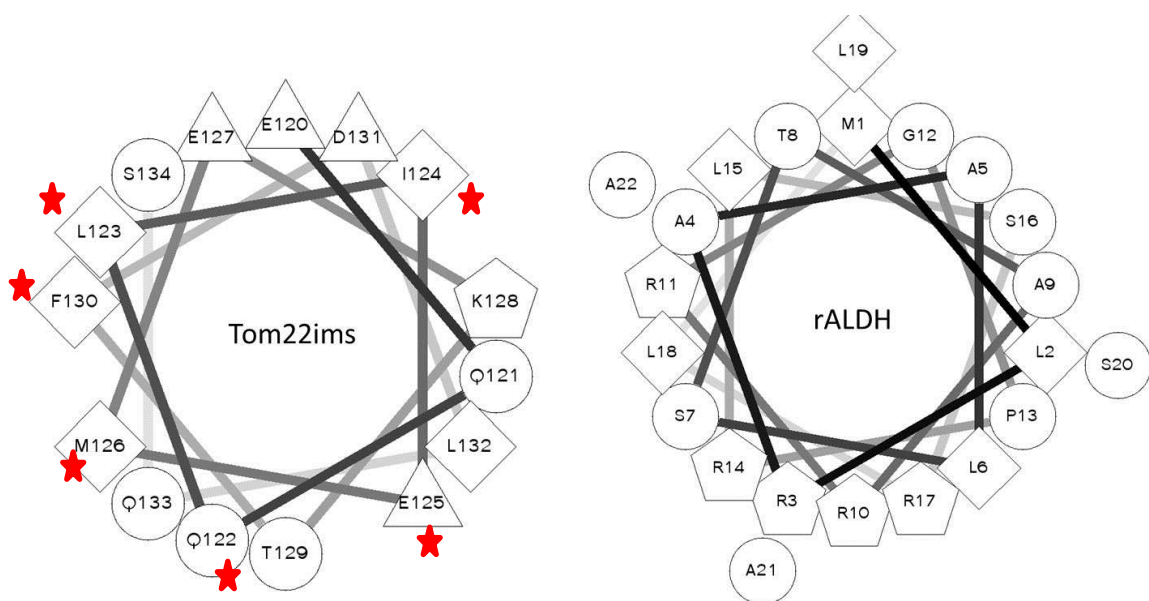


Figure 54: Helical wheel projections of the presequence binding site of Tom22ims and presequence rALDH. Amino acids are outlined according to their properties. Hydrophilic residues are represented by circles; hydrophobic residues are represented as diamonds negatively and positively charged residues are represented as triangles and pentagons respectively. The residues of Tom22ims that show highest chemical shift perturbation are highlighted with red filled stars. Helical wheels were created using 'Wheel', created by D. Armstrong and R. Zidovetzki.

Tim50 is another presequence receptor for the presequence translocase complex (Geissler et al. 2002; Yamamoto et al. 2002; Mokranjac et al. 2003; Mokranjac et al. 2009; Marom et al. 2011). However, the presequence binding site in Tim50 has been identified differently in literature (Qian et al. 2011; Schulz et al. 2011). The presequence was shown to bind Tim50ims encompassing residues 133-476 with dissociation constant of 45 μ M (Marom et al. 2011). Qian et al. (Qian et al. 2011) have mapped the interaction site to a presequence binding groove in Tim50ims structure

involving residues 164-361 whereas the Schultz et al (Schulz et al. 2011) have shown it to be the region involving residues 395-476. In addition, Schultz et al. (Schulz et al. 2011) also showed that a weak binding of presequence COX-IV to the 133-394 residues suggesting the possibility of additional presequence binding site in Tim50ims.

The NMR data in section 3.2.6 support that the presequence can indeed bind to Tim50¹⁶⁴⁻³⁶¹. Currently, the interacting residues could not be identified due to non-availability of backbone resonance assignment of Tim50ims. Moreover, comparative binding studies of presequence with Tim50ims¹⁶⁴⁻⁴⁷⁶ and Tim50ims³⁹⁶⁻⁴⁷⁶ would help to understand the molecular details of presequence binding in Tim50ims.

In contrast to Tom22ims, Tim23ims and Tim50ims, Tim21ims did not show any interaction with presequence up to the tested presequence concentration (3.2mM) (section 3.2.5). The absence of any NMR signal perturbation suggests that direct presequence binding site is either missing or is extremely weak in Tim21ims. This would imply that presequence-Tim21 complex does not play an important role *in vivo*.

4.2.2 Presequence binding and preprotein import

We have determined dissociation constant for the interaction of presequence as ~0.9mM and ~0.45mM for Tom22ims and Tim23ims, respectively. Hence, presequence carrying preprotein can bind Tom22ims and move to a higher presequence affinity site in Tim23ims in line with suggestion that difference in the affinities were suggested to facilitate the import of incoming preprotein. On the other hand, the presequence affinity for the Tim50ims has been reported to be than higher than Tim23ims (K_d ~45 μ M) (Marom et al. 2011) *in vivo* it has been proposed that to be first receptor from presequence translocase. Additionally, the binding affinity of presequence to Tom20 was found to be 20-30 μ M (Abe et al. 2000). Collectively, it suggests the affinity values determined for individual domains under different conditions does not agree well. In this study, the lower affinity values could be due to the free tumbling of the presequence and its receptor, which provide 3D space of search for the presequence and its receptor to form a complex. However, *in vivo* these intermembrane space domains are attached to transmembrane segments that would restrict the presequence receptor in a plane and might modulate the affinity so that presequence can optimally interact with its receptor.

Additionally, the presequence may have multiple binding sites *in vivo* with in same complex as could be envisioned in presequence translocase where both Tim23ims and Tim23ims can interact with presequence. Hence, it is possible that the presequence receptor site is made up of multiple subunits of translocases and determining the affinity of the presequence in complex as ternary complexes would be important to understand the chain of presequence binding sites in mitochondrial preprotein import. In line with this hypothesis it was suggested that the *trans* presequence binding site is formed by Tom22ims in combination with Tom40ims and Tom7ims. The deletion of Tom22ims affects the preprotein import when cytosolic presequence receptors were removed. This supports the importance of potential presequence binding sites in mitochondrial translocases(Moczko et al. 1997).

The presequence receptor complexes involving the intermembrane space of mitochondrial translocases (Tom22ims, Tim23ims and Tim50ims) under investigation shows fast exchange kinetics at NMR timescale and possesses weaker affinity for presequence. These results cumulatively suggest the transient formation of presequence receptor complexes in intermembrane space and correlate well with the view that a high number of preproteins need to be translocated in mitochondrial matrix.

4.3 Interactions at the translocation contact site

The formation of translocation intermediate in the mitochondrial import assays includes fusion of the presequence with carrier domain that can fold upon addition of a substrate. Following the initial translocation of presequence carrying carrier protein, the substrate is added that folds the carrier domain and this folded carrier domain cannot be translocated further. This results in an intermediate super-complex of TOM-TIM and preprotein, where N- terminal part of the preprotein has reached the mitochondrial matrix whereas the C-terminal part lies in the cytoplasm. Such a topology of the preprotein has been used to identify the subunits of the TOM and TIM by chemical crosslinking that cooperates in formation of translocation contact site subunits during the translocation of matrix targeting preprotein. Additionally, these intermediates are not stable and could be released in absence of ATP-driven pulling activity of mtHSP70 in mitochondrial matrix (Schwaiger et al. 1987; Rassow et al. 1990; Dekker et al. 1997).

Crosslinks for Tom22-Tim50 and Tim23-TOM40 have been obtained using this strategy (Chacinska et al. 2005; Tamura et al. 2009). In addition the direct interaction between intermembrane space domains of Tom22 and Tim21 has been shown (Albrecht et al. 2006).

The studies reported in section 3.3 aimed at identifying the direct interaction sites between the various intermembrane space domains at the translocation contact site. We have shown the direct interaction between Tim23ims-Tom22ims and Tim23ims-Tom40ims involving identical region V53-L64 of Tim23ims (Figure 34 and 35). This region has tendency to form coil-coil domain. However, in the free Tim23ims, this region is mostly disordered. The specificity of Tom22ims-Tim23ims interaction has been supported by the reverse titration of labeled Tom22ims with Tim23ims (Figure 34, B). The dissociation constant for Tom22-Tim23ims and Tom22ims-Tom40ims lies in mill molar range. Owing to the transient nature of the translocation intermediate, the weak interaction between the subunits is expected at the translocation contact site. The detailed information of the binding site for the above mentioned complexes from Tom40ims and Tom22ims side are absent. However, the primary sequence of both Tom22ims as well as Tom40ims is predicted to form the transient helices (Appendix Figure 9 and 10). To gain insights into the binding region of Tim23ims with Tom22ims and Tom40ims the helical wheel projections of the binding regions of the Tim23ims were obtained that shows the clustering of hydrophobic residues at one face of the helix in Tom40 as well as Tom22ims. The residues V53, L58, L61 and L64 in Tim23ims also lie on the same side of helix in Tim23ims.

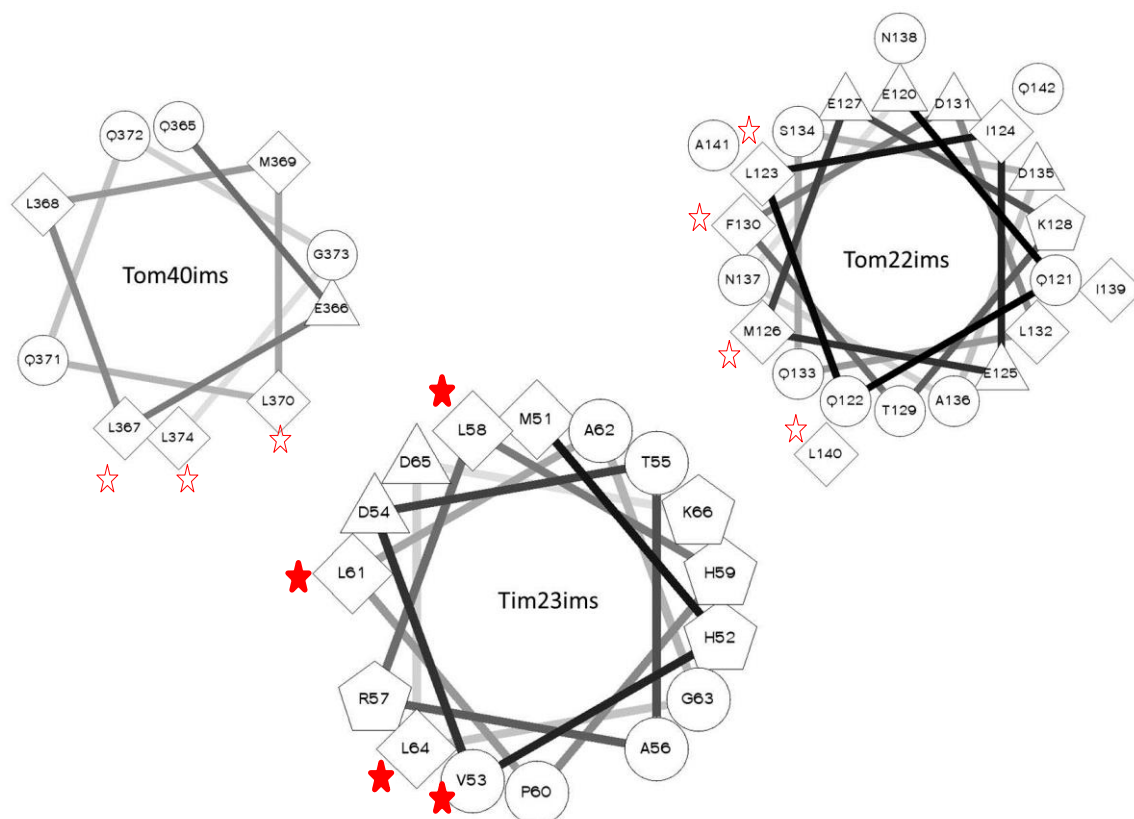


Figure 55: Helical wheel projections of the Tim23ims, Tom40ims and Tom22ims showing the hydrophobic amino acids clustered on one face of helix. Amino acids are outlined according to their properties. Hydrophilic residues are represented by circles; hydrophobic residues as diamonds, negatively and positively charged residues as triangles and pentagons respectively. The residues that show highest chemical shift perturbation in Tim23ims upon addition of Tom22ims and Tom40ims are marked with stars red color (filled). Potential interacting residues of Tom22ims and Tom40ims are marked with red stars (open). Helical wheels were created using 'Wheel', created by D. Armstrong and R. Zidovetzki.

Moreover, the membrane binding residues (1-7 and 31-44) of Tim23ims could also interact with the outer membrane and this Tim23-membrane anchor could facilitate the higher affinity *in vivo* in contrast to the solution conditions (detail in section 4.4.2). Notably, an interaction between Tim21ims and Tom22ims could not be detected (up to 0.8mM concentration of Tom22ims). On one side this would question their direct binding and on other hand, the reason behind this was poorly understood as same under same sample conditions Tim21ims successfully interacted with Tim23ims and Tom22ims. The only likely explanations are:

- 1) Interaction is too weak.
- 2) It may require the preformed helix in Tom22 (transient in our case) or
- 3) additional interacting agents such as membrane mimetic environment or unknown subunit that mediate the interaction.

To address the presence of a potentially very weak interaction paramagnetic tag labeled Tom22ims was titrated in ^{15}N Tim21ims that also showed no significant effect (Figure 33). This supports the absence of a direct interaction among Tim21ims and Tom22ims under tested solution conditions. Requirement of additional agent/subunits is most likely, as previously the direct interaction between Tim21ims and Tom22ims has been determined using the immobilized Tim21ims and solubilized mitochondria (Albrecht et al. 2006). The same study also showed a direct interaction of His tagged - Tim21ims and GST-tagged Tom22ims by affinity tagging whereas in our studies we have successfully removed the fusion tag, which prevents any interference from the tag.

4.4 Tim23ims acts as a hub protein in intermembrane space

We have demonstrated by CD and NMR spectroscopy that Tim23 exists as intrinsically disordered protein in solution (Figure 18 and 19). Both *in vitro* and *in vivo* studies have shown that various interactions including subunit-subunit and subunit presequence in intermembrane space of yeast mitochondrion are modulated by mutations at the C-terminus of the intermembrane part of Tim23 (Gevorkyan-Airapetov et al. 2009; Tamura et al. 2009). Unraveling and understanding the functional mechanism of mitochondrial matrix targeting protein translocation was the primary objective of this study with special emphasis on multiple interactions of Tim23ims. To this end, we correlated NMR titration based interactions with available *in vivo* data.

4.4.1 Conserved linear motifs of Tim23ims in yeast

Protein-protein interactions in disordered proteins can be mediated by short stretches of residues embedded locally in disordered regions or connected by flexible regions (Fuxreiter et al. 2007). Observation of residue specific NMR structural properties of Tim23 in section 3.1.3, revealed three major observations. (a) Transient alpha helical conformational preference for residues 73-84. (b) Highly flexible region involving residues V17-D22. (c) N-terminal first 12 residues exhibiting differential orientation with respect to magnetic field. Residues 1-7 and 58-78 can undergo disorder to order transition upon interaction according to software ANCHOR (Dosztanyi et al. 2009). The sequence alignment of Tim23ims revealed the conservation of hydrophobic

like M, W, L, F, P and I and charged D, E or R (Figure 56). Moreover, the highly conserved residues 87-96 of Tim23ims were not predicted to undergo any disorder to order transition based on ANCHOR prediction.

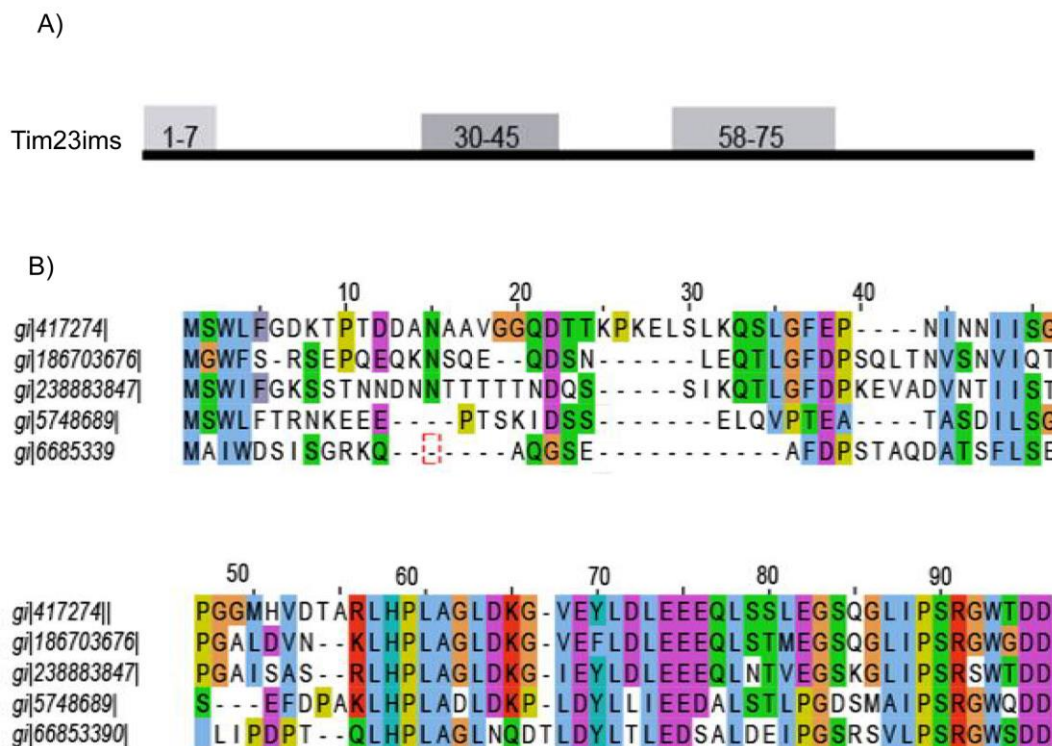


Figure 56: (A) ANCHOR based prediction of binding regions of disordered Tim23ims. (B) Sequence alignment for intermembrane space domain of Tim23 highlighting the conservation of residues among different kingdom Fungi with gi|4172774| as *Saccharomyces cerevisiae*, gi|186703676| as *Zygosaccharomyces rouxii*, gi|238883847| as *Candida albicans*, gi|5748689| as *Schizosaccharomyces pombe*, gi|6685339| as *Aspergillus fumigatus*. The sequence alignment was done using ClustalW and visualized with Jalview. Hydrophobic and aromatic residues are highlighted in blue, positive and negative charged as red and magenta respectively, neutral residues in green, glycine and proline in orange and yellow respectively.

Our NMR based binding studies for Tim23ims and its binding partners were able to provide the residue specific boundaries of their interaction motifs. The multiple interactions in disordered Tim23ims at residue level are described in section 3.2.1, 3.3.2, 3.3.3, 3.4, 3.5 and 3.6. We have identified the binding sites and affinities of Tim23ims with various ligands. Firstly, we addressed the role of Tim23ims in the translocation contact site by probing its binding with Tom22ims, Tom40ims (already explained in section 4.3.1) and mitochondrial membrane mimicking environments using micelles and liposomes. Secondly, we probed the binding of other subunits of TIM23 complex i.e. Tim50ims and Tim21ims to Tim23ims.

4.4.2 Tim23ims interact with the mitochondrial membrane at its N-terminus via two binding motifs.

The role of two membrane spanning topology of Tim23 and its relevance in mitochondrial preprotein import has been highly controversial (Chacinska et al. 2003; Popov-Celeketic et al. 2008; Tamura et al. 2009). The dynamic insertion and de-insertion of the N-terminal residues of Tim23 into the outer mitochondrial membrane and its clipping upon addition of protease is striking but structurally unclear. On the other hand, a morphological defect in yeast mitochondria lacking the first 50 residues and lethal phenotype of yeast upon removal of first 24 residues has also been reported (Davis et al. 2000; Donzeau et al. 2000).

In this study (section 3.4), NMR based titrations of Tim23ims with mitochondrial membrane mimicking liposomes and DHPC micelles clearly identify residues 1-7 and 31-46 as the two membrane binding sites in Tim23ims (Figure 36 and 37). Similar membrane mimics have been previously used to test lipid protein interaction using NMR (Fernandez et al. 2002; Billen et al. 2008; Bodner et al. 2010).

The N-terminal membrane binding region (residues 1-7) significantly show higher affinity than intermediate region (to all membrane mimetic tested) as NMR signal corresponding to these resonances got perturbed at lower protein to lipid ratios. The NMR based titrations of Tim23ims with the inner mitochondrial membrane mimicking liposomes (lacking cardiolipin) also resulted in the interaction with the similar membrane binding region (Figure 36 and 37).

We have also shown that cardiolipin; a negatively charged phospholipid of the inner mitochondrial membrane increases the binding affinity of Tim23 to liposomes and thus could modulate the interaction of Tim23ims with the inner membrane (Figure 36, D). Our data support the fact that the Tim23ims can bind either of the mitochondrial membrane. In parallel to the emerging view of cardiolipin micro domains in *E.coli*, where many selectively proteins exists in regions with high membrane curvature. It is interesting to speculate the interaction of Tim23ims with inner mitochondrial membrane and cardiolipin to regions of high membrane curvature (Sorice et al. 2009; Renner and Weibel 2011). Interaction of Tim23 with cardiolipin containing liposomes, thus could

provide interlink the defective morphology of the mitochondria lacking first fifty residues and lethal phenotype of yeast lacking first 24 residues in Tim23.

Our mitochondrial membrane mimicking liposomes have the similar composition as mitochondrial membrane for PE, PS, PC and cardiolipin but lacks PI. The OMM is rich in PI than IMM; however the amount of PI differs under different physiological conditions. Addition of PI would modulate the binding affinity of OMM and IMM. On the other hand, the translocation contact sites in mitochondria are rich in PE and cardiolipin (Ardail et al. 1990). These two phospholipids are present in liposomes used in this study. Despite the lack of PI in our mitochondrial membrane mimetic liposomes, we suggest that the binding of Tim23ims is not exclusive to the outer mitochondrial membrane but can also bind to the inner mitochondrial membrane.

The structural studies of Tim23ims bound to the mitochondrial membrane mimicking liposomes was not possible due to slow exchange regime of liposome bound Tim23ims and its free form, where signals corresponding to liposome bound Tim23ims could not be observed in NMR spectra. Hence, the structural characterization of the membrane binding of Tim23ims has been done in its micellar (DHPC) bound form. Micellar bound Tim23ims showed lack of significant changes in secondary carbon chemical shifts and excludes the formation of regular secondary structure (α -helix) in its first fifty residues. The first 12 residues of Tim23ims interacted strongly with liposomes as well as DHPC micelles. Hence, further structural studies were done on micellar bound 1-13 residues peptide of Tim23 which revealed that the hydrophobic cluster of W3, L4 and F5 involved in binding to the hydrophobic face of DHPC micellar core with its side chains. These results were similar to the interactions of hydrophobic clusters in unfolded Omp X to the DHPC micelles (Hiller et al. 2008).

In conclusion, this study has unambiguously identified the membrane binding regions in Tim23ims. Indeed the hydrophobic residues W, L, I and F are conserved in the membrane binding region of Tim23ims across many members of kingdom Fungi. But, our results could not structurally explain the phenomenon of membrane insertion and deinsertion as that would require a minimum of approximately 20 hydrophobic residues to transverse as a single membrane transversing helix. Besides this, how the clipping of Tim23ims in protease added assay can occur in intact mitochondria need to be answered which was linked to its two membrane spanning topology. Potentially this

could be due to the indirect effect of cleavage of the *cis* domains of outer mitochondrial translocase by protease that affects the intactness of outer membrane. However, direct evidence of the above mentioned speculation is missing.

Interestingly, the presequence binding region involving residues 71-84 were not affected by the addition of membrane mimics implying the segregation of presequence binding region from membrane binding.

Our results of NMR titration of presequence to liposome bound Tim23ims shows that presequence can bind to liposomes and the Tim23-presequence complex formed is similar irrespective of presence or absence of liposomes (Figure 41 and 42). Additionally, Tim23ims also get dissociated from liposomes upon increasing the concentration of presequence. The presequence are positively charged amphipathic helices and could interact themselves with the negatively charged liposomes due to their physiochemical properties. The magnitude of chemical shift in presequence interacting residues of Tim23 complex in absence of liposomes or in presence of liposomes at higher ratio excludes the direct interference of presequence by liposome interaction.

The dissociation of Tim23 ims from liposomes could be due to the disordering effect of lipids such as cardiolipin (CL) as seen previously in case of mitochondrial presequence cytochrome-*c* oxidase (Colotto et al. 1998). Moreover, it has observed that the presequence of the cytochrome *c* oxidase subunit IV (p25) can promote lipid mixing of large unilamellar liposomes containing CL and PE, without significant contents leakage ((Mandieu et al. 1995).

Hence, the Tim23ims has two membrane anchoring sites involving residues 1-12 and 31-46 and cannot form any regular secondary structure to transverse the outer mitochondrial membrane. Additionally, membrane binding sites act as linear motif, which is segregated from presequence binding motif in disordered Tim23ims.

4.4.3 Fuzzy complex involving Tim21ims-Tim23ims

Tim21-Tim23 crosslinks have been observed *in vivo* (Tamura et al. 2009). Two dynamic models have been proposed to explain the molecular basis for the dual sorting of an incoming preprotein either into the inner mitochondrial membrane or into the matrix by the presequence translocase. Tim21 and Pam17 are two important key players

in these models (section 1.3.4, Figure 10 and 11). However, the interaction studies of Tim21 with any of the subunits of presequence translocase are missing.

Using intermolecular PREs and NMR titration, this study (section 3.5) shows that Tim23ims directly binds to Tim21ims. The interaction of Tim21ims involves three separate binding motifs of Tim23ims involving residues 1-7, 67-73 and 90-96 (Figure 46 and 49). The binding motif of Tim23 corresponding to residues 67-73, is in accordance to the *in vivo* crosslinks of Tim21-Tim23 detected by incorporating the unnatural amino acid-BPA at position 71 and 78 of Tim23 (Tamura et al. 2009). All three binding motifs of Tim23ims has a single interaction site in Tim21ims (involving β -stand 1 and α -helix 1, Figure 46). The binding interface of Tim21ims-Tim23ims complex involves the shallow pocket with K139 and Y141 residues of Tim21ims mediating interactions with the aromatic or bulky hydrophobic side chain and negatively charged residues respectively in the interaction motif of Tim23ims. These two residues make the binding interface suitable for cation- π , π - π and hydrophobic interaction with Tim23ims (Figure 51). The residues in the binding interface for Tim21ims and interaction motifs of Tim23ims are conserved across the yeast ascertaining the importance of interface (Figure 56 and Appendix Figure 8).

The presence of a single binding site, with multiple binding motifs was previously seen in case of Cdc4-Sic1 complex. In Cdc4-Sic1 complex, the sub-optimal binding motifs of disordered Sic 1 are in rapid exchange with each other i.e. are in dynamic equilibrium between the bound and free form. Along the same line, our model (Figure 57) for Tim21ims-Tim23ims complex structurally envisions the binding of linearly placed hydrophobic motifs of Tim23ims into a single binding site in Tim21ims.

The individual motifs of Tim23ims can bind to the same site in Tim21ims with weaker affinity and exhibits fast exchange kinetics at NMR time scale. This indicates that interaction motifs in Tim23ims are in exchange to a single binding site of Tim21ims. The three binding motifs of Tim23ims contribute to the overall affinity of Tim21-Tim23 complex. The presence of three interaction motifs instead of one in Tim23ims increases the concentration of the binding motif and therefore the speed of binding (i.e. higher K_{on} rate). This is in accordance to the Fly casting mechanism, which could explain the kinetic advantage for the disordered proteins having the larger capture radius to interact weakly with its specific target. (Pontius 1993; Zhou 2012).The binding

motifs of Tim23ims seem to be representing the beads as binding motifs connected by flexible linkers that would enhance the association rate of Tim23 with Tim21.

The fast exchange kinetics of all three constructs of Tim23ims could allow the transposition of these binding motifs into the binding pocket of Tim21ims. The low binding affinity and higher K_{off} of each of binding motif of Tim23ims would facilitate the weak association of Tim21 with Tim23 at least with one of the available motif. As the motifs of Tim23ims doesn't attain a regular secondary structure in bound form and cannot be described with a single static conformation in bound state, the complex involving intermembrane space domains of Tim21-Tim23 is another example of fuzzy complex. The binding interface involves the transient interactions that are adapted depending upon the availability of motif rather than being static. The biological relevance of the multivalent interactions in Tim23ims is further explained in section 4.6.

Recently, a new component of presequence translocase known as Mgr2 has been proposed to mediate the interaction between Tim21 and Tim23^{core} subunits (i.e. Tim23, Tim50 and Tim17) (Gebert et al. 2012). Mgr2 is a ~10 kDa, inner mitochondrial membrane protein. The authors were able to detect a small amount of Tim21 bound to Tim23^{core} in absence of Mgr2 and the direct binding of Tim21ims-Tim23ims at higher temperature require efficient coupling of Tim21-Tim23^{core} via Mgr2. Thus in current scenario, it is most likely that Tim21ims-Tim23ims interaction established in this study could further get modulated either by the transmembrane part or intermembrane space domain of Mgr2.

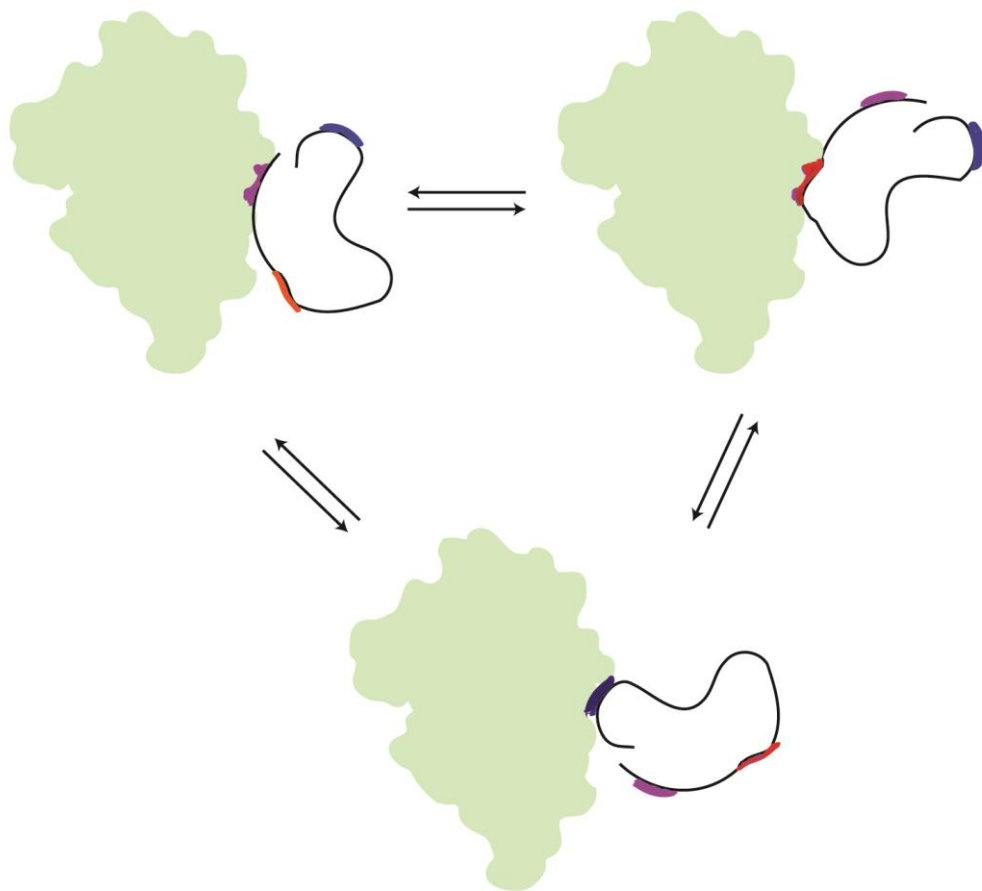


Figure 57: Model representing the interaction mode of linear motifs of disordered Tim23 in fuzzy complex of intermembrane space domains of Tim21-Tim23. Schematic representation of interaction of intrinsically disordered protein Tim23ims (black ribbon) at a single binding site of Tim21ims (filled green surface) via three linear binding motifs of Tim23ims (colored blue, red and purple). These binding motifs are in rapid exchange at a single binding site in Tim21ims. Three such representations are shown that exists in an ensemble of conformations showing either of three linear motifs interacting at single site in Tim21ims. Three linear motifs of Tim23ims confer overall affinity to Tim21ims-Tim23ims complex.

4.4.4 Tim50ims interacts with Tim23ims using multiple interaction motifs

Tim23ims and Tim50ims interact with each other and play an important role in the presequence carrying preprotein import. *In vitro* (by SPR, pull down assays, mutational analysis) and *in vivo* (indirectly using the chemical crosslinking) binding studies revealed that the residues Y70 and L71 were critical for this interaction (Gevorkyan-Airapetov et al. 2009; Tamura et al. 2009). The recent crystal structure of intermembrane space domain of Tim50 (164-361) proposed that β -hairpin with charged

residues play an important role in binding with Tim23ims (Qian et al. 2011). Recently, Schultz et al, showed the differential binding of full-length intermembrane space domain of Tim50 (133-476) and Tim50 (164-361) with Tim23ims. In order to understand the Tim23-Tim50 interaction in detail, we have identified the complete binding site of Tim50ims in Tim23ims. This include the four binding regions involving residues 1-7, 36-53, 56-78, and 80-93 with residues 56-78 forming the primary binding site for Tim50ims (Figure 52). These binding motifs are consistent with the previous results where mutations in this region at residues P60, Y70, L71 resulted in either no or weaker binding than wild type Tim23ims (Gevorkyan-Airapetov et al. 2009). We suggest that the terminal hydrophobic motifs 1-7 and 80-94 of Tim23 contributes towards the affinity of complex in line with the SPR data where replacement of residue 83 with small but hydrophobic alanine instead of glycine contributed to much tighter binding than the wild type Tim23ims (Gevorkyan-Airapetov et al. 2009). Hence, the interaction between the Tim23ims and Tim50ims is primarily hydrophobic and is mediated by four linear motifs comprised of residues 1-7, 36-53, 58-78 and 90-96 in Tim23ims. Much like Tim23-Tim21ims, our NMR titration results of Tim23ims-Tim50ims interaction also suggest the presence of multiple interaction motifs in Tim23ims involved in binding to Tim50ims. But how these motifs binds to Tim50ims and their importance and role in Tim23-Tim50 complex is still under investigation.

4.5 Conserved linear motifs of Tim23ims actively mediate its interaction with Tom22, Tom40ims, Tim21ims, Tim50ims and presequence.

We have identified the direct interaction between Tim23ims with Tom22ims, Tom40ims, Tim50ims, Tim21ims and mitochondrial membrane mimics that provided a residue specific interaction network of Tim23ims. This led us to determine the linear interaction motifs of Tim23ims for each of the ligand tested. The affinity of Tim23ims for various ligands is in the following order: Tim50ims (~50 μ M) > Tim21 (~150 μ M)> presequence (~500 μ M)> Tom22ims/Tom40ims (~1mM).

The five residue specific linear interaction motifs of Tim23ims identified in this study are shown in Figure 58. We observed significant overlap in the interaction residues especially in N-terminal residues 1-7 and residues 58-78, indicating that key residues of Tim23ims interaction network lies in these region. Moreover, the sequence alignment of these regions showed that they are conserved among most members of the fungi kingdom (Figure 56). Analogous findings for importance of these regions in Tim23 have been found *in vivo* where deletion of the first twenty residues and mutations in residues Y70 and L71 caused a lethal phenotype and growth defects in yeast (Davis et al. 2000; Gevorkyan-Airapetov et al. 2009; Mokranjac et al. 2009; Tamura et al. 2009). However, mutation of residues L64 and L78 did not cause growth defects individually but when combined with L71 mutation showed pronounced growth defects (Tamura et al. 2009). This is in accordance with our data where L71 residue is part of a common motif (58-71) for binding five important ligands i.e. Tom22, Tom40, Tim50, Tim21 and presequence. This suggests that the region 58-78 is a hub interaction site of Tim23ims.

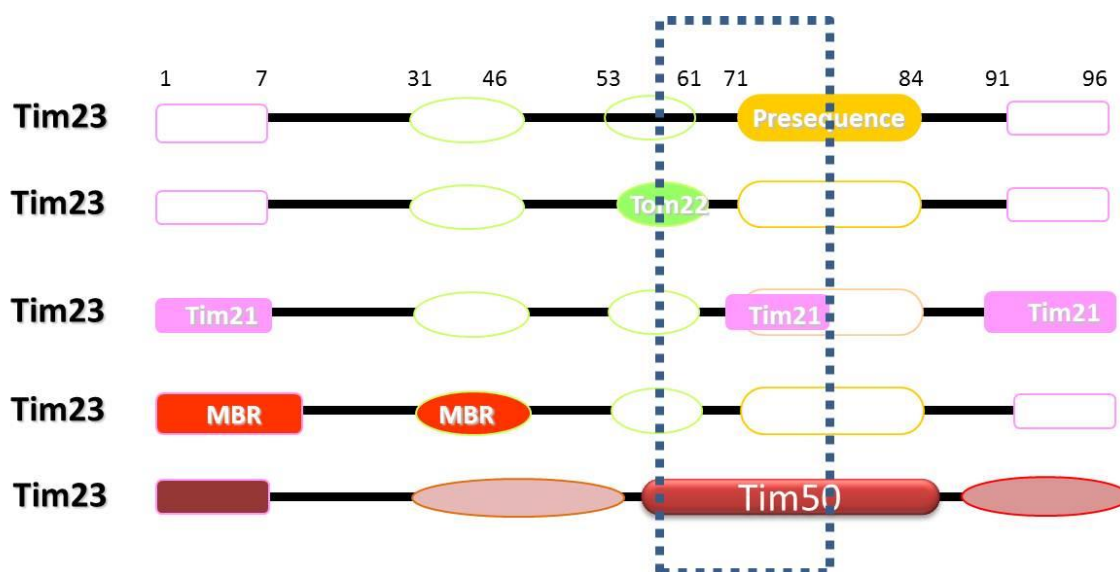


Figure 58: **Five linear interaction motifs in Tim23ims** as deduced from our NMR based interaction studies revealed multiple overlapping interacting residues such as 1-7 involved in binding of Tim21, Tim50 and membrane mimicking liposomes and residues 58-73 as main hub region in C-terminus of Tim23 mediating interactions with intermembrane space domains of all subunits of translocases i.e. Tom22, Tim50, Tim21 and presequence. MBR means the membrane binding region. Tim23 ims is schematically shown in black line and the interaction motifs are highlighted in different geometrical shapes. Various interaction motifs in Tim23ims can be envisioned as the beads on the disordered Tim23.

Simultaneous *in vivo* crosslinks obtained for Tim23 (with BPA) at position 71 with Tim21 and Tim50 (Tamura et al. 2009) can also be explained with our data as Ims of both proteins share common interaction motif involving residue 71. Tom22ims and Tom40ims weakly interacted with Tim23 and play a role in the translocation contact site. They both interact with similar residues (53, 58, 61 and 64) of Tim23ims, suggesting the common degenerate mechanism for direct interaction of Tom22ims or Tom40ims with Tim23ims (Figure 33 and 34).

This study has identified the complete binding site of Tim23ims involving the segments of residues of Tim23 interacting with Tim21, Tim50 and mitochondrial membrane. Thus, results of this study correlated well with available *in vivo* data.

4.6 Multivalency and tuned hydrophobicity governs the working of Tim23

Structurally, protein-protein interaction is one of the physical basis for providing dynamics to the mitochondrial protein import machinery. The TIM23 complex consisting of core components Tim17, Tim50 and Tim23 and additionally Tim21, is dynamic machinery that changes its subunit composition in response to incoming precursor as Tim23^{sort} and Tim23^{motor} form (or the conformational changes in subunits). At the intermembrane space side, functional working of TIM23 complex requires the association and dissociation (or a conformational change) of Tim21 subunit to Tim23^{core}. The insertion of intermembrane space domain of Tim23 into the outer mitochondrial membrane has been shown to increase as a function of incoming precursor and Tim23-Tim50 interaction (Kozany et al. 2004; Chacinska et al. 2005; Popov-Celeketić et al. 2008; Tamura et al. 2009).

This study provides insight into the dynamic interaction network of presequence translocase with central focus on Tim23ims and shows that the intrinsically disordered Tim23ims acts as a hub protein with residues 58-78 involved in binding the presequence, Tim21ims, Tim50ims, Tom22ims and Tom40ims (Figure 58). This hub region is highly conserved in yeast (Figure 56). Interestingly, another highly conserved membrane binding motif involving residues 1-7 is also used by Tim23 to interact with Tim50 and Tim21 and contributes to overall affinity in complex (Figure 56 and 58). The accessory need of membrane binding motif (residues 1-7) for binding to Tim50

provides an explanation of previously observed effect of Tim50 on clipping efficiency and outer membrane interaction of Tim23.(Kozany et al. 2004; Chacinska et al. 2005; Popov-Celeketi et al. 2008; Tamura et al. 2009).

We have attempted to understand the *in vivo* function of Tim23 ims and underline the role of interplay of linear motifs in Tim23ims identified in our studies to facilitate mitochondrial preprotein import by binding multiple partners. Tim23ims has been shown to use more than one linear interaction motif to bind Tim21ims and Tim50ims (Figure 58). This is consistent with the need of multiple mutations *in vivo* to see the effect of interaction among Tim23-Tim50 interaction (Tamura et al. 2009). Further structural studies, with Tim23-Tim50 and Tim21-Tim50 are required to determine molecular basis of interaction of Tim23 in detail to recapitulate its mechanism *in vivo*. In addition to degeneracy in the linear motifs used by Tim23ims to interact with Tim21ims and Tim50ims, for membrane interaction Tim23ims selectively uses its N-terminal motifs (1-12 and 31-46).

Interestingly, swapping several weak binding motifs of Tim23 at single site in Tim21 might be used to accommodate the available motif of Tim23 to facilitate its binding. This would explain the availability of binding motifs for the other interaction partners like Tim50 and presequence. Additionally, multivalent hydrophobic linear motifs might be needed for providing the affinity to effectuate the biological function. In conclusion, interactions of disordered Tim23ims agrees well with an emerging significance of intrinsically disordered proteins that involves a multitude of interactions with fine tuning between specificity and endurance to accomplish the biological function, which can be envisioned by precise usage of interaction motifs in binding to membrane, a common motif involved in regulating the interactions with Tim21, Tim50 and presequence, and multiple motifs for binding same ligands i.e. Tim21ims and Tim50ims.

5 Summary and Outlook

5.1 Intermembrane space domains of mitochondrial translocases

In this study, ensemble view of structural properties and dynamics in intermembrane space domains of mitochondrial translocases have been investigated. We showed that Tim23ims and Tom22ims are disordered in solution possessing transient secondary structural elements. On the other hand, Tim21ims has been shown to exhibit concentration dependent effects and exist in monomer-dimer equilibrium in solution. The monomeric Tim21ims display dynamics in β -strands 2, 3 and 4 and further studies using relaxation dispersion and nuclear spin relaxation ($R_{1\rho}$ and R_{ex}) experiments are required to understand the role of these dynamics for either dimerization or various other protein-protein interactions. The expression, purification and solution conditions for Tim50ims have been optimized and further characterization including sequence-specific resonance assignments are currently in progress.

5.2 Presequence-intermembrane space domain interactions

Presequence recognition guides the preprotein translocation in mitochondria. This study has identified the presequence binding domains of Tim23ims and Tom22ims at single residue level as residues 71-84 and 122-133 respectively. Mutational analysis of the presequence-Tim23 complex showed that both the N and C-terminus of presequence binds to the presequence receptor and R17 in C terminus of presequence is important for binding Tim23ims. Collectively, the Tom22ims-presequence and

Tim23ims-presequence interaction suggest that a combination of hydrophobic and electrostatic interactions is critical for formation of the presequence-receptor complex. The affinity Tim23ims-presequence and Tom22ims-presequence interactions are in mill-molar range indicative of a lower stability of the complex. Our data also suggests that Tim21ims does not act as presequence receptor.

This study has also established the presence of a presequence binding site in Tim50¹⁶⁴⁻³⁶¹. With the availability of assignment, structural insights into the binding site will be obtained and can be compared to full length Tim50ims.

5.3 Interactions between ims domains of the mitochondrial translocases

Biological macromolecular machineries such as TIM23 translocases are functionally complex in nature. The functions of such machineries are encoded by its interacting subunits. Along the same line, the interface of its interacting subunits could primarily be important to modulate the comprehensive function of this biological machinery.

The study of various protein-protein interactions between various subunits of TIM23 translocases is important for the understanding of underlying molecular mechanisms involved in their function. We have used NMR spectroscopy based titrations of various subunits to understand the molecular mechanisms for the functionality of TIM23 at residue specific level (cf. Table 15).

This study envisions the need of disorder in Tim23ims to be multifunctional and establishes it as a hub protein in the intermembrane space of yeast mitochondria.

Most importantly, interactions of disordered Tim23ims are central to this study and that showed active involvement of common hydrophobic linear motifs in mediating multiple interactions with various ligands such as mitochondrial membrane, Tom22ims, Tom40ims, Tim50ims, Tim21ims and presequences. However, we have deduced that the inter-subunit interactions differ in affinity ranging from a micro-molar range affinity in case of Tim23-Tim50 to a milli-molar in case of Tim23ims-Tom22ims. Interestingly, the various low affinity interactions could be collectively involved in the functioning of the TIM23complex.

Furthermore, this study has identified the interaction between Tim21-Tim23ims and modeled the structural basis for interaction of Tim21ims-Tim23 mediated by three hydrophobic linear motifs of Tim23 into a single binding site of Tim21ims.

K139 and Y141 of Tim21ims are found to be key residues in the binding interface of Tim21ims-Tim23ims have been mutated. NMR titrations of mutant Tim21ims with Tim23ims would be performed to validate the model of the complex.

Our studies on Tim21ims and Tom22ims showed the absence of any significant interaction between them at the translocation contact site whereas Tim23ims-Tom22ims and Tim23ims –Tom40ims interacts with a lower affinity in the milli-molar range. Future studies will focus on identification of Tom40ims residues involved in binding to Tim23ims. For this, the labeled Tom40ims peptide would be obtained using a similar strategy as that of Tom22ims using Z₂-fusion tag.

Table 15: Comprehensive view of the binding site and affinities of interactions between the intermembrane space domains in yeast mitochondrion deduced during this study. The tabular check board involves unlabeled subunit along the row and labeled subunit along the column. Each checkbox represents an interaction. Within each check box, the first line denotes presence, absence, not tested interaction by YES, NO, ND respectively, second line denotes the residues involved, followed by its affinity in third line.

| Unlabeled/ Labeled | rALDH | Tom40 | Tom22 | Tim21 | Tim23 | Tim50 |
|-----------------------|-------------------------|----------------------|-----------------------|--|---------------------------|--|
| ¹⁵ NTom22 | YES | ND | - | NO | YES 53,58,61 1mM | NO |
| ¹⁵ NTim21 | NO | ND | NO | YES | YES (138-146) 150μM | ND |
| ¹⁵ NTim23 | YES (71-84) 0.5mM | YES (53-61) ND | YES (53-61) 1mM | YES (1-7,67-73, 90-96) 0.15mM | NO | YES (2-7, 53- 73, 83-94) 0.05mM |

This study also involved preliminary structural studies on Tim50ims and its role in various interactions. Solution condition of Tim50ims has been optimized. Thus, backbone assignment and structural characterization of binary interactions of Tim50ims with Tim23ims, Tom22ims and Tim21ims will form the basis for future studies.

Additional interactions of Tim23ims with Pam17ims, Tim17ims and the potential role of Mgr2 in modulating the Tim21-Tim23 complex will also be of interest in future. In conclusion, we showed that disordered Tim23ims acts as the hub protein and demonstrated its role in various interactions involving the other subunits of the TIM23 complex at single residue level. This study provides a basis for the future studies to understand the molecular mechanisms for the function of the TIM23 complex in mitochondrial import.

6 References

- Abe, Y., Shodai, T., Muto, T., Mihara, K., Torii, H., Nishikawa, S., Endo, T., and Kohda, D. (2000). Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. *Cell* 100, 551-560.
- Ahting, U., Thun, C., Hegerl, R., Typke, D., Nargang, F.E., Neupert, W., and Nussberger, S. (1999). The TOM core complex: the general protein import pore of the outer membrane of mitochondria. *J Cell Biol* 147, 959-968.
- Albrecht, R., Rehling, P., Chacinska, A., Brix, J., Cadamuro, S.A., Volkmer, R., Guiard, B., Pfanner, N., and Zeth, K. (2006). The Tim21 binding domain connects the preprotein translocases of both mitochondrial membranes. *EMBO Rep* 7, 1233-1238.
- Alder, N.N., Jensen, R.E., and Johnson, A.E. (2008a). Fluorescence mapping of mitochondrial TIM23 complex reveals a water-facing, substrate-interacting helix surface. *Cell* 134, 439-450.
- Alder, N.N., Sutherland, J., Buhring, A.I., Jensen, R.E., and Johnson, A.E. (2008b). Quaternary structure of the mitochondrial TIM23 complex reveals dynamic association between Tim23p and other subunits. *Mol Biol Cell* 19, 159-170.
- Ardail, D., Privat, J.P., Egret-Charlier, M., Levrat, C., Lerme, F., and Louisot, P. (1990). Mitochondrial contact sites. Lipid composition and dynamics. *J Biol Chem* 265, 18797-18802.
- Attardi, G., and Schatz, G. (1988). Biogenesis of Mitochondria. *Annual Review of Cell Biology* 4, 289-331.
- Battiste, J.L., and Wagner, G. (2000). Utilization of Site-Directed Spin Labeling and High-Resolution Heteronuclear Nuclear Magnetic Resonance for Global Fold Determination of Large Proteins with Limited Nuclear Overhauser Effect Data†. *Biochemistry* 39, 5355-5365.
- Bauer, M.F., Sirrenberg, C., Neupert, W., and Brunner, M. (1996). Role of Tim23 as voltage sensor and presequence receptor in protein import into mitochondria. *Cell* 87, 33-41.

- Becker, T., Böttlinger, L., and Pfanner, N. (2012). Mitochondrial protein import: from transport pathways to an integrated network. *Trends in Biochemical Sciences* 37, 85-91.
- Berjanskii, M.V., and Wishart, D.S. (2005). A Simple Method To Predict Protein Flexibility Using Secondary Chemical Shifts. *Journal of the American Chemical Society* 127, 14970-14971.
- Bertoncini, C.W., Jung, Y.S., Fernandez, C.O., Hoyer, W., Griesinger, C., Jovin, T.M., and Zweckstetter, M. (2005). Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein. *Proc Natl Acad Sci U S A* 102, 1430-1435.
- Billen, L.P., Kokoski, C.L., Lovell, J.F., Leber, B., and Andrews, D.W. (2008). Bcl-XL inhibits membrane permeabilization by competing with Bax. *PLoS biology* 6, e147.
- Bodner, C.R., Maltsev, A.S., Dobson, C.M., and Bax, A. (2010). Differential phospholipid binding of alpha-synuclein variants implicated in Parkinson's disease revealed by solution NMR spectroscopy. *Biochemistry* 49, 862-871.
- Bogomolovas, J., Simon, B., Sattler, M., and Stier, G. (2009). Screening of fusion partners for high yield expression and purification of bioactive viscotoxins. *Protein Expr Purif* 64, 16-23.
- Bolender, N., Sickmann, A., Wagner, R., Meisinger, C., and Pfanner, N. (2008). Multiple pathways for sorting mitochondrial precursor proteins. *EMBO Rep* 9, 42-49.
- Bolliger, L., Junne, T., Schatz, G., and Lithgow, T. (1995). Acidic receptor domains on both sides of the outer membrane mediate translocation of precursor proteins into yeast mitochondria. *EMBO J* 14, 6318-6326.
- Bomer, U., Meijer, M., Maarse, A.C., Honlinger, A., Dekker, P.J., Pfanner, N., and Rassow, J. (1997). Multiple interactions of components mediating preprotein translocation across the inner mitochondrial membrane. *EMBO J* 16, 2205-2216.
- Brzovic, P.S., Heikaus, C.C., Kisselev, L., Vernon, R., Herbig, E., Pacheco, D., Warfield, L., Littlefield, P., Baker, D., Klevit, R.E., et al. (2011). The acidic transcription activator Gcn4 binds the mediator subunit Gal11/Med15 using a simple protein interface forming a fuzzy complex. *Mol Cell* 44, 942-953.
- Cavanagh, J., Fairbrother, W., Palmer III, A., Rance, M., and Skelton, N. (2007). *Protein NMR Spectroscopy (Second Edition)* (Elsevier Inc.), p. 757.
- Chacinska, A., Koehler, C.M., Milenkovic, D., Lithgow, T., and Pfanner, N. (2009). Importing mitochondrial proteins: machineries and mechanisms. *Cell* 138, 628-644.

- Chacinska, A., Lind, M., Frazier, A.E., Dudek, J., Meisinger, C., Geissler, A., Sickmann, A., Meyer, H.E., Truscott, K.N., Guiard, B., et al. (2005). Mitochondrial presequence translocase: switching between TOM tethering and motor recruitment involves Tim21 and Tim17. *Cell* 120, 817-829.
- Chacinska, A., Rehling, P., Guiard, B., Frazier, A.E., Schulze-Specking, A., Pfanner, N., Voos, W., and Meisinger, C. (2003). Mitochondrial translocation contact sites: separation of dynamic and stabilizing elements in formation of a TOM-TIM-preprotein supercomplex. *EMBO J* 22, 5370-5381.
- Chacinska, A., van der Laan, M., Mehnert, C.S., Guiard, B., Mick, D.U., Hutu, D.P., Truscott, K.N., Wiedemann, N., Meisinger, C., Pfanner, N., et al. (2010). Distinct forms of mitochondrial TOM-TIM supercomplexes define signal-dependent states of preprotein sorting. *Mol Cell Biol* 30, 307-318.
- Claypool, S.M. (2009). Cardiolipin, a critical determinant of mitochondrial carrier protein assembly and function. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1788, 2059-2068.
- Colotto, A., Martin, I., Ruysschaert, J.M., Sen, A., and Epand, R.M. (1998). Structural study of the interaction between the mitochondrial presequence of cytochrome c oxidase subunit IV and model membranes. *Bioscience reports* 18, 251-263.
- Court, D.A., Nargang, F.E., Steiner, H., Hodges, R.S., Neupert, W., and Lill, R. (1996). Role of the intermembrane-space domain of the preprotein receptor Tom22 in protein import into mitochondria. *Mol Cell Biol* 16, 4035-4042.
- Csermely, P., Sandhu, K.S., Hazai, E., Hoksza, Z., Kiss, H.J., Miozzo, F., Veres, D.V., Piazza, F., and Nussinov, R. (2012). Disordered proteins and network disorder in network descriptions of protein structure, dynamics and function: hypotheses and a comprehensive review. *Current protein & peptide science* 13, 19-33.
- D'Silva, P.D., Schilke, B., Walter, W., Andrew, A., and Craig, E.A. (2003). J protein cochaperone of the mitochondrial inner membrane required for protein import into the mitochondrial matrix. *Proc Natl Acad Sci U S A* 100, 13839-13844.
- D'Silva, P.R., Schilke, B., Hayashi, M., and Craig, E.A. (2008). Interaction of the J-protein heterodimer Pam18/Pam16 of the mitochondrial import motor with the translocon of the inner membrane. *Mol Biol Cell* 19, 424-432.
- Davis, A.J., Sepuri, N.B., Holder, J., Johnson, A.E., and Jensen, R.E. (2000a). Two intermembrane space TIM complexes interact with different domains of Tim23p during its import into mitochondria. *J Cell Biol* 150, 1271-1282.

- Davis, A.J., Sepuri, N.B., Holder, J., Johnson, A.E., and Jensen, R.E. (2000b). Two intermembrane space TIM complexes interact with different domains of Tim23p during its import into mitochondria. *J. Cell. Biol.* 150, 1271-1282.
- de la Cruz, L., Bajaj, R., Becker, S., and Zweckstetter, M. (2010). The intermembrane space domain of Tim23 is intrinsically disordered with a distinct binding region for presequences. *Protein Sci* 19, 2045-2054.
- Dekker, P.J., Martin, F., Maarse, A.C., Bomer, U., Muller, H., Guiard, B., Meijer, M., Rassow, J., and Pfanner, N. (1997). The Tim core complex defines the number of mitochondrial translocation contact sites and can hold arrested preproteins in the absence of matrix Hsp70-Tim44. *EMBO J* 16, 5408-5419.
- Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6, 277-293.
- Donzeau, M., Kaldi, K., Adam, A., Paschen, S., Wanner, G., Guiard, B., Bauer, M.F., Neupert, W., and Brunner, M. (2000). Tim23 links the inner and outer mitochondrial membranes. *Cell* 101, 401-412.
- Dosztanyi, Z., Chen, J., Dunker, A.K., Simon, I., and Tompa, P. (2006). Disorder and sequence repeats in hub proteins and their implications for network evolution. *Journal of proteome research* 5, 2985-2995.
- Dosztanyi, Z., Meszaros, B., and Simon, I. (2009). ANCHOR: web server for predicting protein binding regions in disordered proteins. *Bioinformatics* 25, 2745-2746.
- Dunn, C.D., Lee, M.S., Spencer, F.A., and Jensen, R.E. (2006). A genomewide screen for petite-negative yeast strains yields a new subunit of the i-AAA protease complex. *Mol Biol Cell* 17, 213-226.
- Fernandez, C., Hilty, C., Wider, G., and Wuthrich, K. (2002). 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.
- Fuxreiter, M. (2012). Fuzziness: linking regulation to protein dynamics. *Mol Biosyst* 8, 168-177.
- Fuxreiter, M., and Tompa, P. (2012). Fuzzy complexes: a more stochastic view of protein function. *Advances in experimental medicine and biology* 725, 1-14.
- Fuxreiter, M., Tompa, P., and Simon, I. (2007). Local structural disorder imparts plasticity on linear motifs. *Bioinformatics* 23, 950-956.
- Gavel, Y., and von Heijne, G. (1990). Cleavage-site motifs in mitochondrial targeting peptides. *Protein engineering* 4, 33-37.

- Gebert, M., Schrempp, S.G., Mehnert, C.S., Heißwolf, A.K., Oeljeklaus, S., Ieva, R., Bohnert, M., von der Malsburg, K., Wiese, S., Kleinschroth, T., et al. (2012). Mgr2 promotes coupling of the mitochondrial presequence translocase to partner complexes. *J Cell Biol* 197, 595-604.
- Geissler, A., Chacinska, A., Truscott, K.N., Wiedemann, N., Brandner, K., Sickmann, A., Meyer, H.E., Meisinger, C., Pfanner, N., and Rehling, P. (2002). The mitochondrial presequence translocase: an essential role of Tim50 in directing preproteins to the import channel. *Cell* 111, 507-518.
- Gevorkyan-Airapetov, L., Zohary, K., Popov-Celeketic, D., Mapa, K., Hell, K., Neupert, W., Azem, A., and Mokranjac, D. (2009). Interaction of Tim23 with Tim50 Is essential for protein translocation by the mitochondrial TIM23 complex. *J Biol Chem* 284, 4865-4872.
- Goddard, T.D., and Kneller, D.G. (2006). SPARKY 3. (San Francisco, University of California).
- Guntert, P. (2004). Automated NMR structure calculation with CYANA. *Methods Mol Biol* 278, 353-378.
- Hammen, P.K., Waltner, M., Hahneemann, B., Heard, T.S., and Weiner, H. (1996). The role of positive charges and structural segments in the presequence of rat liver aldehyde dehydrogenase in import into mitochondria. *J Biol Chem* 271, 21041-21048.
- Herbig, E., Warfield, L., Fish, L., Fishburn, J., Knutson, B.A., Moorefield, B., Pacheco, D., and Hahn, S. (2010). Mechanism of Mediator recruitment by tandem Gcn4 activation domains and three Gal11 activator-binding domains. *Mol Cell Biol* 30, 2376-2390.
- Hill, K., Model, K., Ryan, M.T., Dietmeier, K., Martin, F., Wagner, R., and Pfanner, N. (1998). Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins [see comment]. *Nature* 395, 516-521.
- Hiller, S., Wider, G., Imbach, L.L., and Wuthrich, K. (2008). Interactions with hydrophobic clusters in the urea-unfolded membrane protein OmpX. *Angew Chem Int Ed Engl* 47, 977-981.
- Hutu, D.P., Guiard, B., Chacinska, A., Becker, D., Pfanner, N., Rehling, P., and van der Laan, M. (2008). Mitochondrial protein import motor: differential role of Tim44 in the recruitment of Pam17 and J-complex to the presequence translocase. *Mol Biol Cell* 19, 2642-2649.
- Igura, M., Ose, T., Obita, T., Sato, C., Maenaka, K., Endo, T., and Kohda, D. (2005). Crystallization and preliminary X-ray analysis of mitochondrial presequence receptor Tom20 in complexes with a presequence from aldehyde dehydrogenase. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 61, 514-517.

- Jensen, R.E., and Dunn, C.D. (2002). Protein import into and across the mitochondrial inner membrane: role of the TIM23 and TIM22 translocons. *Biochim Biophys Acta* 1592, 25-34.
- Komiya, T., Rospert, S., Koehler, C., Looser, R., Schatz, G., and Mihara, K. (1998). Interaction of mitochondrial targeting signals with acidic receptor domains along the protein import pathway: evidence for the 'acid chain' hypothesis. *EMBO J* 17, 3886-3898.
- Kozany, C., Mokranjac, D., Sichting, M., Neupert, W., and Hell, K. (2004). The J domain-related cochaperone Tim16 is a constituent of the mitochondrial TIM23 preprotein translocase. *Nat Struct Mol Biol* 11, 234-241.
- Lee, D., Hilty, C., Wider, G., and Wuthrich, K. (2006). Effective rotational correlation times of proteins from NMR relaxation interference. *J Magn Reson* 178, 72-76.
- Lietzow, M.A., Jamin, M., Dyson, H.J., and Wright, P.E. (2002). Mapping long-range contacts in a highly unfolded protein. *J Mol Biol* 322, 655-662.
- Liu, Q., Krzewska, J., Liberek, K., and Craig, E.A. (2001). Mitochondrial Hsp70 Ssc1: role in protein folding. *J Biol Chem* 276, 6112-6118.
- Lohret, T.A., Jensen, R.E., and Kinnally, K.W. (1997). Tim23, a protein import component of the mitochondrial inner membrane, is required for normal activity of the multiple conductance channel, MCC. *J. Cell Biol.* 137, 377-386.
- Mandieau, V., Martin, I., and Ruyschaert, J.M. (1995). Interaction between cardiolipin and the mitochondrial presequence of cytochrome c oxidase subunit IV favours lipid mixing without destabilizing the bilayer structure. *FEBS Lett* 368, 15-18.
- Marom, M., Azem, A., and Mokranjac, D. (2011a). Understanding the molecular mechanism of protein translocation across the mitochondrial inner membrane: Still a long way to go. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1808, 990-1001.
- Marom, M., Dayan, D., Demishtein-Zohary, K., Mokranjac, D., Neupert, W., and Azem, A. (2011b). Direct interaction of mitochondrial targeting presequences with purified components of the TIM23 protein complex. *J Biol Chem* 286, 43809-43815.
- Marom, M., Safonov, R., Amram, S., Avneon, Y., Nachliel, E., Gutman, M., Zohary, K., Azem, A., and Tsfadia, Y. (2009). Interaction of the Tim44 C-terminal domain with negatively charged phospholipids. *Biochemistry* 48, 11185-11195.

- Martinez-Caballero, S., Grigoriev, S.M., Herrmann, J.M., Campo, M.L., and Kinnally, K.W. (2007). Tim17p regulates the twin pore structure and voltage gating of the mitochondrial protein import complex TIM23. *J Biol Chem* 282, 3584-3593.
- McBride, H.M., Neuspiel, M., and Wasiak, S. (2006). Mitochondria: More Than Just a Powerhouse. *Current Biology* 16, R551-R560.
- Meier, S., Neupert, W., and Herrmann, J.M. (2005). Conserved N-terminal negative charges in the Tim17 subunit of the TIM23 translocase play a critical role in the import of preproteins into mitochondria. *J Biol Chem* 280, 7777-7785.
- Meinecke, M., Wagner, R., Kovermann, P., Guiard, B., Mick, D.U., Hutu, D.P., Voos, W., Truscott, K.N., Chacinska, A., Pfanner, N., et al. (2006). Tim50 maintains the permeability barrier of the mitochondrial inner membrane. *Science* 312, 1523-1526.
- Miao, B., Davis, J.E., and Craig, E.A. (1997). Mge1 functions as a nucleotide release factor for Ssc1, a mitochondrial Hsp70 of *Saccharomyces cerevisiae*. *J Mol Biol* 265, 541-552.
- Mignotte, B., and Vayssiere, J.-L. (1998). Mitochondria and apoptosis. *European Journal of Biochemistry* 252, 1-15.
- Mittag, T., and Forman-Kay, J.D. (2007). Atomic-level characterization of disordered protein ensembles. *Curr Opin Struct Biol* 17, 3-14.
- Mittag, T., Kay, L.E., and Forman-Kay, J.D. (2010). Protein dynamics and conformational disorder in molecular recognition. *Journal of molecular recognition : JMR* 23, 105-116.
- Mittag, T., Orlicky, S., Choy, W.Y., Tang, X., Lin, H., Sicheri, F., Kay, L.E., Tyers, M., and Forman-Kay, J.D. (2008). Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor. *Proc Natl Acad Sci U S A* 105, 17772-17777.
- Moczko, M. (1997). The intermembrane space domain of mitochondrial Tom22 functions as a trans binding site for preproteins with N-terminal targeting sequences. *Mol. Cell. Biol.* 17, 6574-6584.
- Moczko, M., Bomer, U., Kubrich, M., Zufall, N., Honlinger, A., and Pfanner, N. (1997). The intermembrane space domain of mitochondrial Tom22 functions as a trans binding site for preproteins with N-terminal targeting sequences. *Mol Cell Biol* 17, 6574-6584.
- Mokranjac, D., Bourenkov, G., Hell, K., Neupert, W., and Groll, M. (2006). Structure and function of Tim14 and Tim16, the J and J-like components of the mitochondrial protein import motor. *EMBO J* 25, 4675-4685.

- Mokranjac, D., and Neupert, W. (2005). Protein import into mitochondria. *Biochem Soc Trans* 33, 1019-1023.
- Mokranjac, D., and Neupert, W. (2010). The many faces of the mitochondrial TIM23 complex. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1797, 1045-1054.
- Mokranjac, D., Paschen, S.A., Kozany, C., Prokisch, H., Hoppins, S.C., Nargang, F.E., Neupert, W., and Hell, K. (2003a). Tim50, a novel component of the TIM23 preprotein translocase of mitochondria. *EMBO J* 22, 816-825.
- Mokranjac, D., Popov-Celeketi, D., Hell, K., and Neupert, W. (2005a). Role of Tim21 in mitochondrial translocation contact sites. *J Biol Chem* 280, 23437-23440.
- Mokranjac, D., Sichting, M., Neupert, W., and Hell, K. (2003b). Tim14, a novel key component of the import motor of the TIM23 protein translocase of mitochondria. *EMBO J* 22, 4945-4956.
- Mokranjac, D., Sichting, M., Popov-Celeketi, D., Mapa, K., Gevorgyan-Airapetov, L., Zohary, K., Hell, K., Azem, A., and Neupert, W. (2009). Role of Tim50 in the transfer of precursor proteins from the outer to the inner membrane of mitochondria. *Mol Biol Cell* 20, 1400-1407.
- Mukrasch, M.D., Bibow, S., Korukottu, J., Jeganathan, S., Biernat, J., Griesinger, C., Mandelkow, E., and Zweckstetter, M. (2009). Structural polymorphism of 441-residue tau at single residue resolution. *PLoS biology* 7, e34.
- Narayanan, R.L., Durr, U.H., Bibow, S., Biernat, J., Mandelkow, E., and Zweckstetter, M. (2010). Automatic assignment of the intrinsically disordered protein Tau with 441-residues. *J Am Chem Soc* 132, 11906-11907.
- Nargang, F.E., Rapaport, D., Ritzel, R.G., Neupert, W., and Lill, R. (1998). Role of the negative charges in the cytosolic domain of TOM22 in the import of precursor proteins into mitochondria. *Mol Cell Biol* 18, 3173-3181.
- Oldfield, C.J., Meng, J., Yang, J.Y., Yang, M.Q., Uversky, V.N., and Dunker, A.K. (2008). Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. *BMC genomics* 9, 1471-2164.
- Ottiger, M., Delaglio, F., and Bax, A. (1998). Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J Magn Reson* 131, 373-378.
- Ozenne, V., Bauer, F., Salmon, L., Huang, J.-r., Jensen, M.R., Segard, S., Bernadó, P., Charavay, C., and Blackledge, M. (2012). Flexible-meccano: a tool for the generation of explicit ensemble descriptions of intrinsically disordered proteins and their associated experimental observables. *Bioinformatics* 28, 1463-1470.
- Pawson, T., and Nash, P. (2003). Assembly of cell regulatory systems through protein interaction domains. *Science* 300, 445-452.

- Peters, F., Maestre-Martinez, M., Leonov, A., Kovacic, L., Becker, S., Boelens, R., and Griesinger, C. (2011). Cys-Ph-TAHA: a lanthanide binding tag for RDC and PCS enhanced protein NMR. *Journal of biomolecular NMR* 51, 329-337.
- Pfanner, N. (2000). Protein sorting: recognizing mitochondrial presequences. *Curr Biol* 10, R412-415.
- Pontius, B.W. (1993). Close encounters: why unstructured, polymeric domains can increase rates of specific macromolecular association. *Trends in Biochemical Sciences* 18, 181-186.
- Popov-Celeketi, D., Mapa, K., Neupert, W., and Mokranjac, D. (2008). Active remodelling of the TIM23 complex during translocation of preproteins into mitochondria. *EMBO J* 27, 1469-1480.
- Qian, X., Gebert, M., Hopker, J., Yan, M., Li, J., Wiedemann, N., van der Laan, M., Pfanner, N., and Sha, B. (2011). Structural basis for the function of Tim50 in the mitochondrial presequence translocase. *J Mol Biol* 411, 513-519.
- Rassow, J., Hartl, F.U., Guiard, B., Pfanner, N., and Neupert, W. (1990). Polypeptides traverse the mitochondrial envelope in an extended state. *FEBS Lett* 275, 190-194.
- Raveh, B., London, N., Zimmerman, L., and Schueler-Furman, O. (2011). Rosetta FlexPepDock ab-initio: simultaneous folding, docking and refinement of peptides onto their receptors. *PLoS One* 6, 0018934.
- Renner, L.D., and Weibel, D.B. (2011). Cardiolipin microdomains localize to negatively curved regions of Escherichia coli membranes. *Proc Natl Acad Sci U S A* 108, 6264-6269.
- Rezaei-Ghaleh, N., Blackledge, M., and Zweckstetter, M. (2012). Intrinsically disordered proteins: from sequence and conformational properties toward drug discovery. *Chembiochem* 13, 930-950.
- Rimmer, K.A., Foo, J.H., Ng, A., Petrie, E.J., Shilling, P.J., Perry, A.J., Mertens, H.D., Lithgow, T., Mulhern, T.D., and Gooley, P.R. (2011). Recognition of mitochondrial targeting sequences by the import receptors Tom20 and Tom22. *J Mol Biol* 405, 804-818.
- Roise, D., Horvath, S.J., Tomich, J.M., Richards, J.H., and Schatz, G. (1986). A chemically synthesized pre-sequence of an imported mitochondrial protein can form an amphiphilic helix and perturb natural and artificial phospholipid-bilayers. *Embo Journal* 5, 1327-1334.
- Schneider, R., Huang, J.R., Yao, M., Communie, G., Ozenne, V., Mollica, L., Salmon, L., Jensen, M.R., and Blackledge, M. (2012). Towards a robust description of intrinsic protein disorder using nuclear magnetic resonance spectroscopy. *Mol Biosyst* 8, 58-68.

- Schulz, C., Lytovchenko, O., Melin, J., Chacinska, A., Guiard, B., Neumann, P., Ficner, R., Jahn, O., Schmidt, B., and Rehling, P. (2011). Tim50's presequence receptor domain is essential for signal driven transport across the TIM23 complex. *J Cell Biol* 195, 643-656.
- Schwaiger, M., Herzog, V., and Neupert, W. (1987). Characterization of translocation contact sites involved in the import of mitochondrial proteins. *J Cell Biol* 105, 235-246.
- Schwartz, M.P., and Matouschek, A. (1999). The dimensions of the protein import channels in the outer and inner mitochondrial membranes. *Proc. Natl Acad. Sci. USA* 96, 13086-13090.
- Schwarzinger, S., Kroon, G.J., Foss, T.R., Chung, J., Wright, P.E., and Dyson, H.J. (2001). Sequence-dependent correction of random coil NMR chemical shifts. *J Am Chem Soc* 123, 2970-2978.
- Shiota, T., Mabuchi, H., Tanaka-Yamano, S., Yamano, K., and Endo, T. (2011). In vivo protein-interaction mapping of a mitochondrial translocator protein Tom22 at work. *Proc Natl Acad Sci U S A* 108, 15179-15183.
- Sklenář, V., and Bax, A. (1987). Spin-echo water suppression for the generation of pure-phase two-dimensional NMR spectra. *Journal of Magnetic Resonance* (1969) 74, 469-479.
- Smith, D.B., and Johnson, K.S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* 67, 31-40.
- Sorice, M., Manganelli, V., Matarrese, P., Tinari, A., Misasi, R., Malorni, W., and Garofalo, T. (2009). Cardiolipin-enriched raft-like microdomains are essential activating platforms for apoptotic signals on mitochondria. *FEBS Lett* 583, 2447-2450.
- Tamura, Y., Harada, Y., Shiota, T., Yamano, K., Watanabe, K., Yokota, M., Yamamoto, H., Sesaki, H., and Endo, T. (2009). Tim23-Tim50 pair coordinates functions of translocators and motor proteins in mitochondrial protein import. *J Cell Biol* 184, 129-141.
- Truscott, K.N., Kovermann, P., Geissler, A., Merlin, A., Meijer, M., Driessen, A.J., Rassow, J., Pfanner, N., and Wagner, R. (2001). A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by Tim23. *Nat Struct Biol* 8, 1074-1082.
- Uversky, V.N. (2011a). Flexible nets of malleable guardians: intrinsically disordered chaperones in neurodegenerative diseases. *Chemical reviews* 111, 1134-1166.
- Uversky, V.N. (2011b). Intrinsically disordered proteins from A to Z. *Int J Biochem Cell Biol* 43, 1090-1103.

- Uversky, V.N. (2011c). Multitude of binding modes attainable by intrinsically disordered proteins: a portrait gallery of disorder-based complexes. *Chemical Society reviews* 40, 1623-1634.
- Uversky, V.N. (2012). Intrinsic disorder-based protein interactions and their modulators. *Current pharmaceutical design*.
- van der Laan, M., Chacinska, A., Lind, M., Perschil, I., Sickmann, A., Meyer, H.E., Guiard, B., Meisinger, C., Pfanner, N., and Rehling, P. (2005). Pam17 is required for architecture and translocation activity of the mitochondrial protein import motor. *Mol Cell Biol* 25, 7449-7458.
- van der Laan, M., Hutu, D.P., and Rehling, P. (2010a). On the mechanism of preprotein import by the mitochondrial presequence translocase. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1803, 732-739.
- Van Der Laan, M., Rissler, M., and Rehling, P. (2006a). Mitochondrial preprotein translocases as dynamic molecular machines. *FEMS Yeast Research* 6, 849-861.
- van der Laan, M., Wiedemann, N., Mick, D.U., Guiard, B., Rehling, P., and Pfanner, N. (2006b). A role for Tim21 in membrane-potential-dependent preprotein sorting in mitochondria. *Curr Biol* 16, 2271-2276.
- van Loon, A.P.G.M., Brändli, A., and Schatz, G. (1986). The presequences of two imported mitochondrial proteins contain information for intracellular and intramitochondrial sorting. *Cell* 44, 801-812.
- van Wilpe, S., Ryan, M.T., Hill, K., Maarse, A.C., Meisinger, C., Brix, J., Dekker, P.J., Moczko, M., Wagner, R., Meijer, M., et al. (1999). Tom22 is a multifunctional organizer of the mitochondrial preprotein translocase. *Nature* 401, 485-489.
- Vonheijne, G. (1986). Mitochondrial targeting sequences may form amphiphilic helices. *Embo Journal* 5, 1335-1342.
- Waksman, G. (2005). *Proteomics and Protein-Protein Interactions Biology, Chemistry*. Springer, Heidelberg
- Wang, Y., and Weiner, H. (1993). The presequence of rat liver aldehyde dehydrogenase requires the presence of an alpha-helix at its N-terminal region which is stabilized by the helix at its C termini. *J Biol Chem* 268, 4759-4765.
- Warde-Farley, D., Donaldson, S.L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C.T., et al. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* 38, W214-220.

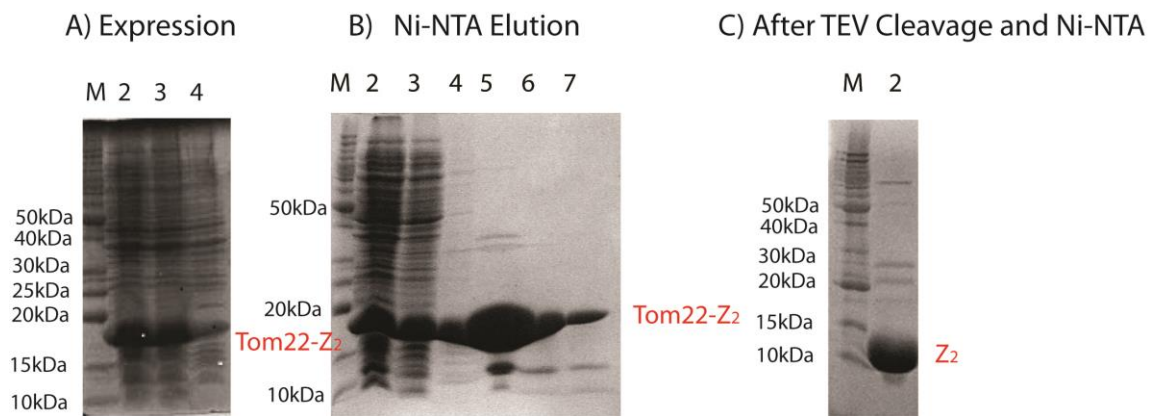
- Weiss, C., Oppliger, W., Vergeres, G., Demel, R., Jenö, P., Horst, M., de Kruijff, B., Schatz, G., and Azem, A. (1999). Domain structure and lipid interaction of recombinant yeast Tim44. *Proc Natl Acad Sci U S A* 96, 8890-8894.
- Wiedemann, N., van der Laan, M., Hutu, D.P., Rehling, P., and Pfanner, N. (2007). Sorting switch of mitochondrial presequence translocase involves coupling of motor module to respiratory chain. *J Cell Biol* 179, 1115-1122.
- Wright, P.E., and Dyson, H.J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol* 293, 321-331.
- Xue, B., Brown, C.J., Dunker, A.K., and Uversky, V.N. Intrinsically disordered regions of p53 family are highly diversified in evolution. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*.
- Yamamoto, H., Esaki, M., Kanamori, T., Tamura, Y., Nishikawa, S., and Endo, T. (2002). Tim50 is a subunit of the TIM23 complex that links protein translocation across the outer and inner mitochondrial membranes. *Cell* 111, 519-528.
- Yamano, K., Yatsukawa, Y., Esaki, M., Hobbs, A.E., Jensen, R.E., and Endo, T. (2008). Tom20 and Tom22 share the common signal recognition pathway in mitochondrial protein import. *J Biol Chem* 283, 3799-3807.
- Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC bioinformatics* 9, 40.
- Zhou, H.X. (2012). Intrinsic disorder: signaling via highly specific but short-lived association. *Trends Biochem Sci* 37, 43-48.

7 Appendix

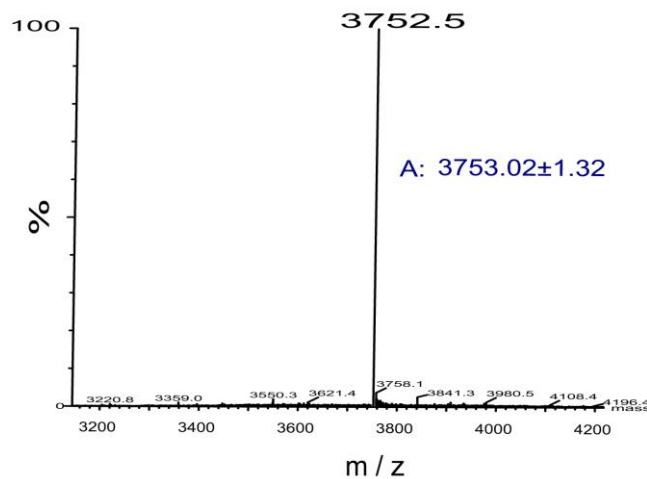
7.1 Purification of Intermembrane space domains

7.1.1 Tom22ims

The overexpression of Z₂-Tom22ims was confirmed using 15% SDS-PAGE electrophoresis (Appendix Figure 1 A, lane 3) and the majority of the expressed protein was present in the soluble fraction, as evident from the strong band around 18 kDa (as opposed to the expected mass of 12.6 kDa). The use of Z₂ domain has facilitated the solubility and overexpression of Tom22ims peptide (corresponding to 3.5kDa) in *E.coli*. The expression of such small peptides is generally difficult in *E.coli*. The identity and homogeneity of the Tom22ims was confirmed by mass spectrometry (Appendix Figure 2).



Appendix Figure 1: Expression and Purification of Z₂-Tom22ims. (A) 15% SDS-PAGE gel showing the overexpression of Z₂-Tom22ims: Lane M: marker, Lane 2: whole cell lysate, Lane 3: supernatant and Lane 4: cell pellet. (B) Ni²⁺-NTA purification of Z₂-Tom22ims: Lane 2: cell lysis supernatant, Lane 3: wash, Lane 4, 5, 6, and 7 are elution fractions of Ni²⁺-NTA. (C) TEV cleavage and removal of the Z₂tag: Lane M: Marker and Lane 2: Z₂ bound to Ni²⁺-NTA resin.



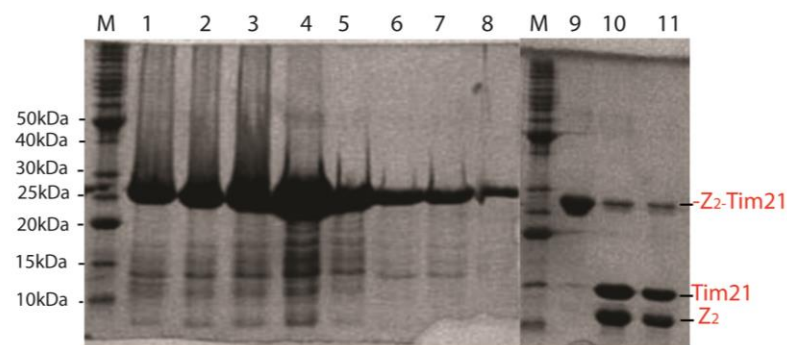
Appendix Figure 2: ESI-MS of Tom22ims confirming the identity of the Tom22ims peptide.

7.1.2 Tim21ims

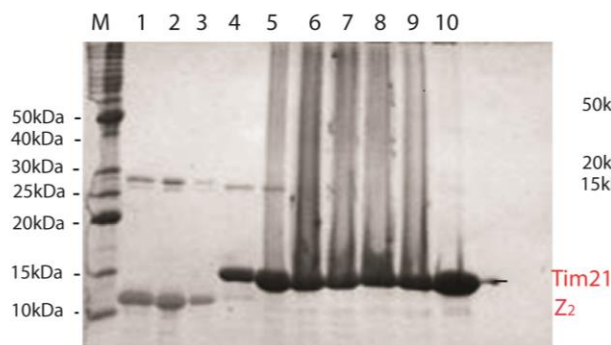
Tim21ims was expressed and purified as described in the Methods (section 2.2.4 and 2.2.5). Tim21ims has the tendency to aggregate and precipitate at higher concentrations (above 1mM). This was overcome by the use of size exclusion chromatography prior to NMR studies to remove efficiently the dimeric and/or aggregated species. In addition to wild type Tim21ims, several mutants were designed

in this study. All the mutants (Tim21 C128A, Tim21 W150S, Tim21 S114A/C128A and Tim21 C128A, Tim21 K139A and Tim21 Y141A) were generated using site directed mutagenesis. They all showed similar expression and were purified using the same procedure as described for *wt* Tim21ims. The results of the purification for *wt* Tim21ims are summarized in Appendix Figure 3.

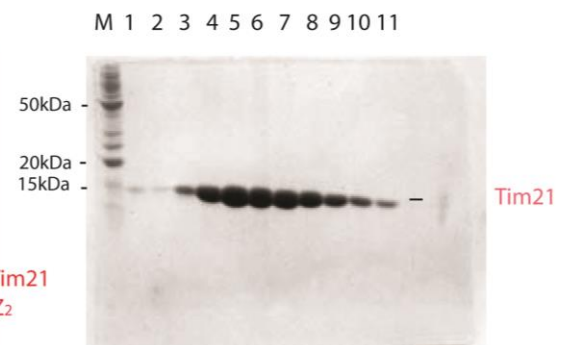
A) Ni NTA elution and TEV cleavage



B) SP Sepharose elution



C) Elution fractions post gel filtration

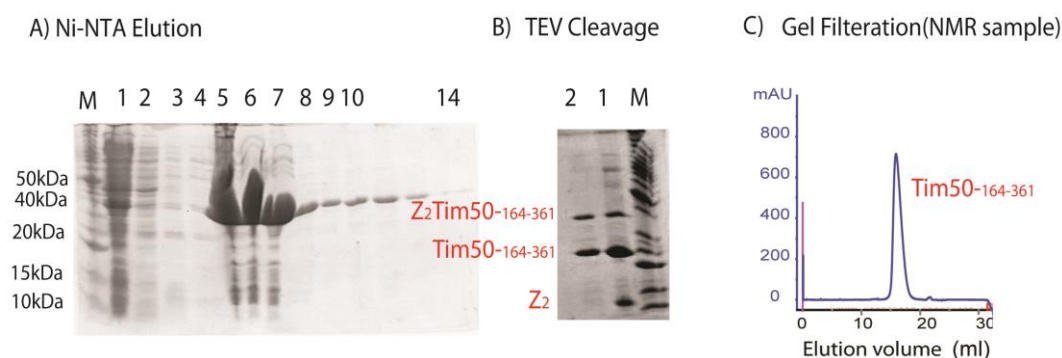


Appendix Figure3: **Purification of Tim21ims.** 15% SDS –PAGE gel showing (A) Ni²⁺-NTA purification of Z₂-Tim21 ims¹⁰³⁻²²⁵. Lane M: marker, Lane 1- 8: fractions eluted with 100mM imidazole and TEV cleavage, Lane 9: before TEV cleavage, Lane 10, 11: after TEV Cleavage. (B) Cation exchange chromatography (SP-sepharose): Lane M: marker, Lane 1: flow through, Lane 2, 3, 4: Wash, Lane5-10: elution fractions C) Elution fractions from gel filtration are shown in Lane 2-11.

7.1.3 Tim50ims (164-361):

The expression and purification of Tim50ims (residues 164-361) is summarized in methods (section 2.2.4 and 2.2.5) and Appendix Figure 4. Moreover, Tim50ims has the tendency to aggregate and precipitate at concentrations higher than 300 μ M. Hence, an additional purification step involving the analytical size exclusion chromatography was used prior to NMR studies to ensure its homogeneity.

Z₂Tim50¹⁶⁴⁻³⁶¹



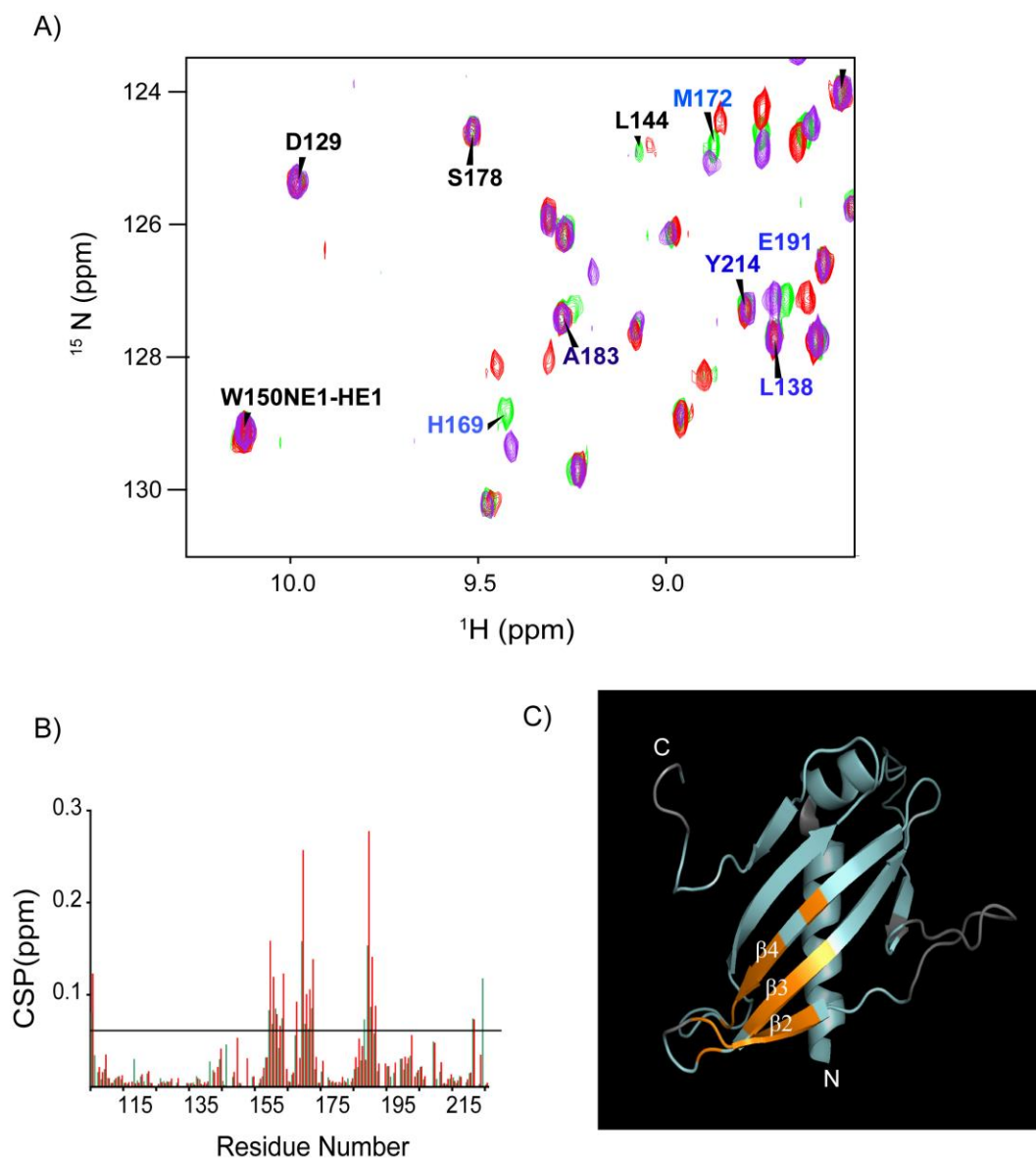
Appendix Figure4: Purification of Tim50ims¹⁶⁴⁻³⁶¹ .(A) 15% SDS gel showing Ni²⁺ NTA purification of Z₂-Tim50ims¹⁶⁴⁻³⁶¹. Lane M: marker, Lane 1: cell lysis supernatant, Lane 2,3,4: wash, Lane 5, 6, and 7: elution fractions with 100mM imidazole, Lane 8-14: elution fractions with 500mM Imidazole. (B) TEV cleavage of Z₂-Tim50ims domain. Lane M: Marker, Lane 1: after TEV cleavage, Lane 2: after removing Z₂ domain with Ni-NTA and before size exclusion chromatography. (C) Gel filtration chromatogram showing single peak to confirm the purity of Tim50ims¹⁶⁴⁻³⁶¹

7.2 Tim21

7.2.1 Effect of pH on backbone resonances of Tim21ims

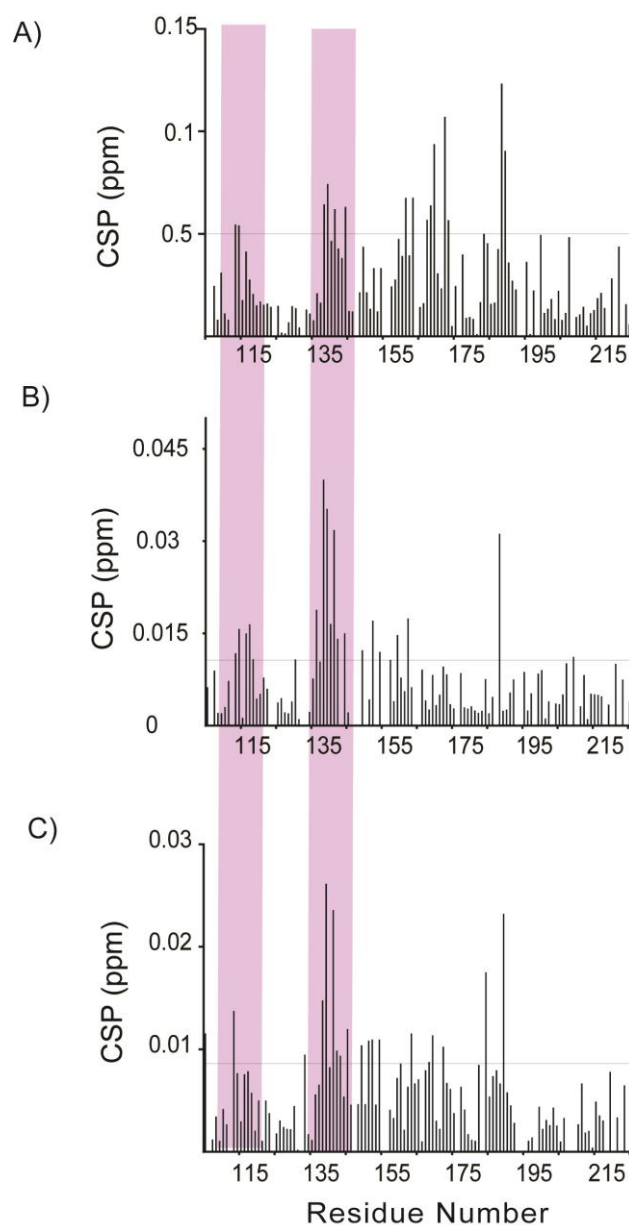
Tim21ims was selected to probe its interaction with various biological ligands has been suggested to play important role in formation of translocation contact site. Additionally, it has potential to interact with other subunits of presequence translocase due to its physical proximity (Albrecht et al. 2006; Tamura et al. 2009). Prior to NMR based mapping of binding interface, pH and salt induced changes in Tim21 were examined to identify sensitive residues and their neighbors involved in electrostatic interaction (protonation state of titrable groups such as His).

Series of ^1H - ^{15}N HSQC spectra of Tim21ims were acquired at different pH (i.e. 6.8, 7.2 and 7.6). On comparing them, it was found that most of the residues belonging to three beta strand (2, 3 and 4) of Tim21ims were significantly perturbed (cut off above 0.6ppm) (Appendix Figure 5). However, the chemical shift position for rest of resonances remained unperturbed. Additionally, the chemical shift dispersion range of the Tim21ims remains unaffected indicating that overall structure of Tim21ims does not change in the range of pH tested

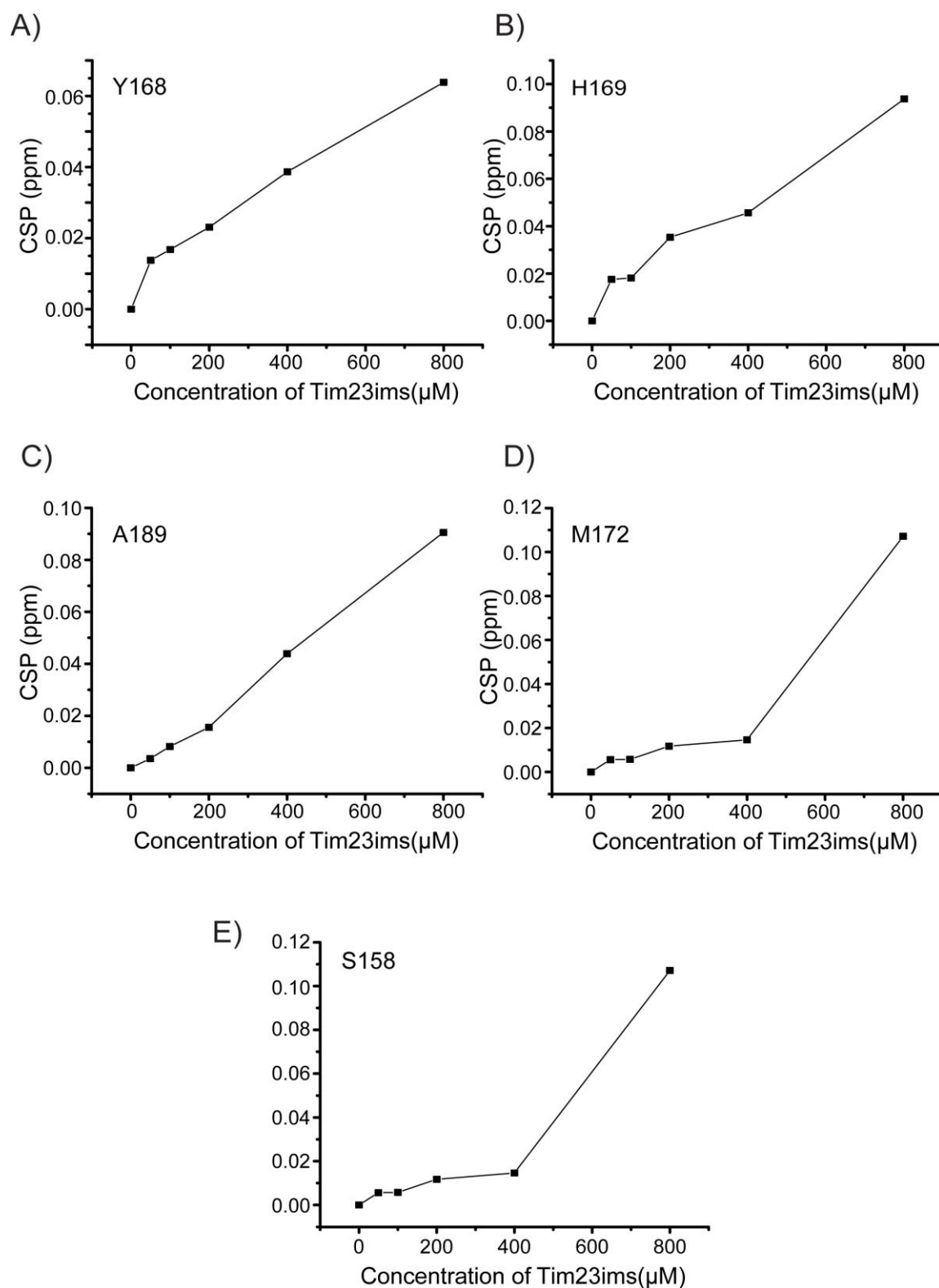


Appendix Figure 5: Influence of near physiological pH ranging from 6.8-7.6 on structural properties of Tim21ims (A) Overlay of ^1H ^{15}N -HSQC of selected region of Tim21 at three different pH (Red: 6.8, green: 7.2 and purple: 7.6). (B) Average Chemical shift perturbations (CSP) deduced from ^1H ^{15}N HSQC spectra's along the polypeptide sequence of Tim21 with pH 7.2 as reference, red (pH=6.8) and green (pH=7.6). (C) Tim21 ims structure (NMR model 1) highlighting the residues that show maximum spectral changes (orange as CSP above 0.06 ppm and intensity ratio below 0.2 as grey) as a function of pH.

7.2.2 Chemical shift changes in observed in backbone resonances of Tim21ims upon addition of different constructs of Tim23ims.

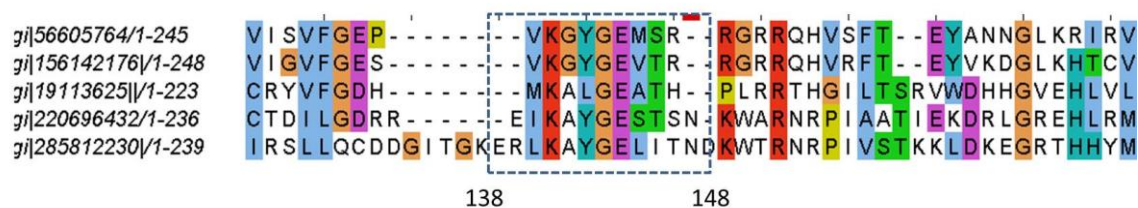


Appendix Figure6: Chemical shift perturbation plot (Chemical shift changes vs residue number) of Tim21ims upon addition of 16 fold excess of Tim23 (1-96) in (A), 32 fold excess of Tim23 (61-96) in (B) and Tim23 (1-13) in (C). The residues highlighted in pink constitute the interaction site of Tim23ims in Tim21ims. The line represents the highest chemical shift perturbation in non-interacting residues 195-215.



Appendix Figure 7: **Binding curves for the residues of Tim21ims outside the Tim23 binding site:** The changes in chemical shift of the non- interacting residues of Tim21ims as a function of increasing concentration of Tim23ims.

7.2.3 Sequence alignment of Tim23 binding site in Tim21ims



Appendix Figure 8: **Sequence alignment for a subset of residues in intermembrane space domain of Tim21** highlighting the conservation of residues in evolution for K139 and Y141 of *S. cerevisiae*. The residues 138-148 in *S. cerevisiae* are marked and highlighted in an open blue box. gi|56605764| represents the protein sequence of Tim21ims for *Rattus norvegicus*, gi|156142176| for *Homo sapiens*, gi|19113625| for *Schizosaccharomyces pombe*, gi|220696432| for *Aspergillus flavus*, gi|285812230| for *Saccharomyces cerevisiae*. The sequence alignment was done using ClustalW and visualized with Jalview. Hydrophobic and aromatic residues are highlighted in blue, positive and negative charged as red and magenta respectively, neutral residues in green, glycine and proline in orange and yellow respectively.

7.3 Primary sequence of presequence rALDH and its variants used in section 3.2.3

| | |
|---------------------------|------------------------|
| rALDH ^{wt} | MLRAALSTARRGPRLSRLLSAA |
| rALDH ¹⁻¹¹ | MLRAALSTAR |
| rALDH ¹²⁻²² | RGPRLSRLLSAA |
| rALDH ^{L18QL19Q} | MLRAALSTARRGPRLSRQQSAA |
| rALDH ^{R1Q} | MLRAALSTARRGPRLSQLLSAA |

7.4 Primary sequence and secondary structure prediction of Tom22ims

The prediction is done based on Jpred software
<http://www.compbio.dundee.ac.uk/www-jpred/>

Jnet represents the final secondary structure prediction for Tom22ims. The two predicted helices include residues 123-128 and 133-142 and is represented as H

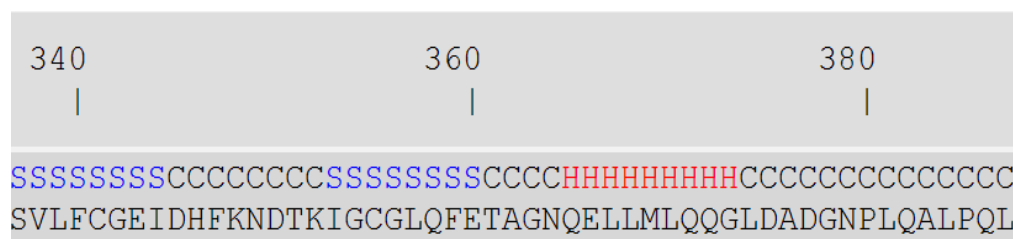
EQQLIEMEKTFDLQSDANNILAQGEKDAAATAN : Tom22ims

---HHHHHH---HHHHHHHHHH----- : Jnet

Appendix Figure 9: Jpred based secondary prediction of Tom22ims in yeast. H means helical propensity of a residue.

Primary sequence of presequence Tom40 ims

Tom40ims: Prediction was derived from software I-TASSER (Zhang 2008).



Appendix Figure 10: **I-TASSER** based secondary prediction of C-terminal residues of Tom40ims in yeast. 'S' denotes the strand, C means coil and H means helical propensity of a residue.

Following last β -strand (353-360), predicted Tom40ims residues 361-387 corresponding to sequence TAGNQELLMLQQGLDADGNPLQALPQL acetylated at the N-terminus and amidated at the C-terminus was chemically synthesized.

Appendix Table1: List of primers used in this study. RE means Restriction Endonuclease site. FP and RP denotes forward and reverse primer respectively, .aa denote amino acid.

| Construct Name | RE Site | Vector | FP/RP | Oligonucleotide Sequence |
|--------------------------------|----------|---|-------|---|
| Tim23 (aa 1-96) | Bam HI | pGEX-2T | FP | 5'GTCAAGCTTGGATCCATGTCGTGGCTTT TTGGAGATAAGAC 3' |
| Tim23 (aa 1-96) | Eco RI | pGEX-2T | RP | 5'GACCATATGGAATTCCTAGTCATCGGTC CACCCACGG 3' |
| Tim23 (aa 1-96), L71A | | pGEX-2T | FP | 5'GTCTAGACAAGGGTGTGGAGTATGCAG ATCTGGAAGAAGAAC 3' |
| Tim23 (aa 1-96), L71A | | pGEX-2T | RP | 5'GTTCTTCTTCCAGATCTGCATACTCCACA CCCTTGCTAGAC 3' |
| Tim23 (aa 1-96), L78A | | pGEX-2T | FP | 5'GATCTGGAAGAAGAACAAGCATCCTCGT TAGAAGGCTCAC 3' |
| Tim23 (aa 1-96), L78A | | pGEX-2T | RP | 5'GTGAGCCTTCTAACGAGGATGCTTGTTCT TTCTTCCAGATC 3' |
| Tim23(aa 1-96), G67C | | pGEX-2T | FP | 5'GGCTGGTCTAGACAAGTGTGTGGAGTAT TTAGATCTGGAAGAAG 3' |
| Tim23(aa 1-96) G67C | | pGEX-2T | RP | 5'CTTCTTCCAGATCTAAATACTCCACACA CTTGCTAGACCAGCC 3' |
| Tim21(aa 103-225) S16C | | pET28aHis ₇ Z ₂ TEV | FP | 5'CACAGCTTTTCAACAGAGCAGTTTGTAT GGTAGAGAAAAAC 3' |
| Tim21(aa 103-225) S16C | | pET28aHis ₇ Z ₂ TEV | RP | 5'GTTTTTCTCTACCATACAACTGCTCTGT TGAAAAGCTGTG 3' |
| Tim23(aa 1-96), T11C | | pGEX-2T | FP | 5'CGTGGCTTTTTGGAGATAAGACACCTTG CGATGATGCGAATGCTGC 3' |
| TIM21 aa. 103. | Bam HI | pET28aHis ₇ Z ₂ TEV | FP | 5'GCTTCTAGAGGATCCGGTGATACACAGC TTTTCAACAGAG 3' |
| TIM21, aa. 225. | Hind III | pET28aHis ₇ Z ₂ TEV | RP | 5'GTCGAATTCAAGCTTTTAATTAGAAACC GGATGCAATTTTGGCTTG 3' |
| TOM22, aa. 120. | Bam HI | pET28aHis ₇ Z ₂ TEV | FP | 5'GCTTCTAGAGGATCCGAACAACAGCTAA TCGAAATGG 3' |
| TOM22, aa. 152 C- term. | Hind III | pET28aHis ₇ Z ₂ TEV | RP | 5'GTCGAATTCAAGCTTTTAATTGGCTGTT GCTGCAGCATC 3' |
| Tim21ims, aa103- 225, C128A | | pET28aHis ₇ Z ₂ TEV | FP | 5'GATATAAGAAGTTTGTACAGGCCGACG ATGGCATTACGGGAAAAAG 3' |
| Tim21ims(aa103- 225), C128A | | pET28aHis ₇ Z ₂ TEV | RP | 5'CTTTTCCCCTAATGCCATCGTCGGCCTGT AACAACTTCTTATATC 3' |
| Tim50 ims, aa. 133 | Bam HI | pET28aHis ₇ Z ₂ TEV | FP | 5'GCTTCTAGAGGATCCAGGGATTGGGAGC CTCAAGAG-3' |
| Tim50 ims - C-term., aa 476 | Hind III | pET28aHis ₇ Z ₂ TEV | RP | 5'GTCGAATTCAAGCTTTTATTGGATTCA GCAATCTTCTTCTTTTC 3' |
| Tim50 ims - C-term., aa 374 | Hind III | pET28aHis ₇ Z ₂ TEV | RP | 5'GTCGAATTCAAGCTTTTAGCCACCAGAT TTATGATCTCCG 3' |
| Tim50 ims, aa. 164 | Bam HII | pET28aHis ₇ Z ₂ TEV | FP | 5'CTTCTAGAGGATCCTTCAACTCAATGTT CACCTACTCCAAGAGCC 3' |

Curriculum Vitae

January 2013, Göttingen

PERSONAL DATA

| | |
|---------------------|------------------------------|
| Name | Rakhi Bajaj |
| Birthdate and place | 06-08-1984, New Delhi, India |
| Citizenship | Indian |

EDUCATION

| | |
|----------------|--|
| 2/2009-Present | Max Planck Institute of Biophysical Chemistry, Göttingen, Germany. |
|----------------|--|

Doctoral Thesis

“Residue level characterization of molecular interactions of intermembrane space domains governing the preprotein import into the mitochondrial matrix ”

Advisor: Prof. Dr. Markus Zweckstetter

| | |
|----------------|---|
| 8/2007-12/2008 | Junior Research Fellow , Indian Institute of Science, Bangalore, India Project Title: “ <i>Cloning and Interaction study of various domains of AHAS enzyme</i> ” |
|----------------|---|

| | |
|-----------------|---|
| 06/2005-04/2007 | M.Sc., Life Sciences , Jawaharlal Nehru University, New Delhi, India Project Title: “ <i>Functional significance of conserved residues in NBD-1 signature motif of Cdr1p, an ABC transporter of Candida albicans</i> ” |
|-----------------|---|

| | |
|-----------------|--|
| 06/2002-04/2005 | B.Sc. (H), Zoology , Hansraj College, Delhi University, India Botany and Chemistry Ancillaries |
|-----------------|--|

HONORS

| | |
|---------|---|
| 2012 | ISOTEC Travel Award, ICMRBS Conference, France |
| 2009 | Junior Group Stipend, GGNB, Gottingen, Germany |
| 07/2007 | Research Fellowship, Indian Council of Medical Research, India |
| 02/2007 | Graduate Aptitude Test for Engineering (GATE), Indian Institute of Technology, India |
| 12/2006 | National Eligibility Test (NET) for Research Fellowship, UGC and CSIR, Government of India. |
| 04/2005 | Dr. Lakshmi Singh Prize for academic excellence in Delhi University |
| 04/2005 | Third rank B.Sc. (H) Zoology, in Delhi University Examination |
| 04/2005 | Certificate of Merit for First Rank in BSc. (H) Zoology, Hansraj College, Delhi |
| 04/2003 | Certificate of Merit for First Rank in BSc. (H) Zoology, Hansraj College, Delhi |
| 2001 | National Standard Examination in Biology (NSEB) India. |

PUBLICATIONS

1. De la Cruz, L., Bajaj, R., Becker, S., Zweckstetter, M (2010). The intermembrane space domain of Tim23 is intrinsically disordered with a distinct binding region for presequences, Prot. Sci, 19, 2045-54

