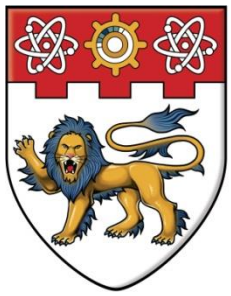


NANYANG TECHNOLOGICAL UNIVERSITY

STRUCTURES AND REGULATION OF COUPLING SUBUNIT F AND THE ARRANGEMENT OF THE SUBUNIT DF-HETERODIMER IN THE *SACCHAROMYCES CEREVISIAE* V_1V_O ATPase

SANDIP BASAK
SCHOOL OF BIOLOGICAL SCIENCES
2013

Structures and regulation of coupling subunit F and the arrangement of the subunit DF-assembly in the *Saccharomyces cerevisiae* V_1V_O ATPase

Sandip Basak

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang Technological University
in fulfillment of the requirement for the degree of
Doctor of Philosophy

2013

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor Professor Dr. Gerhard Grüber, for his expert guidance and warm encouragement. This thesis would not have been possible without the help, support and patience of my principal supervisor. I am very grateful to him for his valuable guidance, imparting thoughtful suggestions at various stages, barring which I would not have been able to complete my Ph.D thesis. He shared his wealth of experience on numerous occasions and this has provided me with a great learning experience.

I would also like to express my sincere thanks to Dr. Shovanlal Gayen to train me initially in the field of biomolecular NMR spectroscopy during the solution structure determination of peptide $F_{90\text{-}116}$. I am very thankful to Dr. Asha Balakrishna and Dr. Malathy Sony Subramanian Manimekalai for training me in the field of protein crystallography. I am very grateful to them for their valuable knowledge, suggestions and helpful discussions on crystallography. Many special thanks to Dr. Cornelia Hunke and Dr. Goran Biuković for helping me to learn the basic techniques of Molecular Biology. This has been a wonderful experience to work with them. Again, I also give my heartfelt thanks to Dr. Malathy Sony Subramanian Manimekalai for her valuable suggestions and fruitful discussions on SAXS data analysis. I would also like to take this opportunity to thank Dr. Jack Wee Lim to train me initially in the field of biomolecular 3D NMR spectroscopy during the backbone assignment and dynamic study of subunit F. I am very grateful to him for his valuable knowledge, suggestions on NMR experiments. I would like to thanks to Ms. Saw for her kind help during the ITC experiments.

I would also like to express my special thanks to my closest friend in Singapore Mr. Phat Vinh Dip for his help in either scientific or non-scientific field throughout the entire tenure of my Ph.D.

I owe very special thanks to my colleagues Lavanya, Yew Kwang, Asha and Goran for spending their valuable time to read my thesis thoroughly. Thanks to my other colleagues Dhirendra, Wilson, Vikeram, Rishi, Priya, Neelu, James, Ardina, Claudia and to all my friends at NTU. I want to thank them for all their help, support, interest and valuable hints.

Special thanks go to Prof. Dr. Manfred Roessle for his help in collecting the Small Angle X-ray (SAXS) Data of the subunit F and $F_{1\text{-}94}$ at EMBL, Hamburg Outstation (Germany).

I gratefully acknowledge the financial support rendered by the Nanyang Technological University of Singapore in the form of Research Scholarship. I am also grateful to the academic

and technical staffs at the School of Biological Sciences, who have helped me in one way or the other in my research work. I wish to express my sincere appreciation and thanks to the staff at beamline 13B1 at NSRRC (National Synchrotron Radiation Research Center) Taiwan for their help in data collection. I would also wish to acknowledge A*STAR BMRC (09/1/22/19/609) for the research support granted.

Last but not least, my heartedly gratitude goes to my parents for selfless support, love and understanding during and before my PhD study in Singapore. I thank also my sisters and all my family members who are always proud of my academic achievements and have taken all the troubles to inspire, encourage and support me to reach this stage.

To my parents for their love, endless support and encouragements!

3. Results

List of figures

Introduction

Materials and methods

Results

Discussion

List of tables

Abbreviations

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
Å	Angstrom
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
BSA	Bovine Serum Albumin
CD	Circular Dichroism
CSI	Chemical shift index
DSS	2, 2-dimethyl-2-silapentane-5-sulphonate
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
FPLC	Fast protein liquid chromatography
HSQC	Heteronuclear single quantum correlation
Hz	Hertz
IPTG	Isopropyl-β-D-thiogalactoside
ITC	Isothermal Titration Calorimetry
Kb	Kilo base pairs
kDa	kilo-Dalton
M	Molar
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
NTA	Nitrilotriacetic acid
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PDB	Protein data bank
PDB	Protein Data Bank
Pefabloc^{SC}	(4-(2-Aminoethyl)benzenesulfonylfluoride)
PMSF	Phenylmethylsulfonyl fluoride
PPM	Parts per million
RMSD	Root mean square deviation

SAXS	Small Angle X-ray Scattering
SDS	Sodium dodecyl sulphate
S. cerevisiae	*Saccharomyces cerevisiae*
TCEP	Tris(2-carboxyethyl) phosphine
TOCSY	Total correlation spectroscopy
Tris	Tris-(hydroxymethyl) aminomethane

Abstract

V-ATPases play an important role in the acidification of intracellular compartments such as lysosomes, endosomes, Golgi complexes and secretary granules. The V-ATPases are composed of at least 14 separate gene products, with many of these subunits present in multiple isoforms. The proposed subunit stoichiometry of V_1 is A_3:B_3:C:D:E_3:F:G_3:H_1 (1). The integral V_O domain contains six different subunits in a stoichiometry of a:d:c_{4-5}:c':c":e. V-ATPases exist in a dynamic equilibrium between fully assembled complexes and reversibly disassembled V_1 and V_O subcomplexes. Depending on the energy status of the cell, this equilibrium can be rapidly shifted (2). Vacuolar ATPases use the energy derived from ATP hydrolysis, catalyzed in the A_3B_3 sector of the V_1 ATPase to pump protons via the membrane-embedded V_O sector. The energy coupling between the two sectors occurs via the so-called central stalk, to which subunit F belongs. In the present study, the low resolution structure of recombinant subunit F (*VMA7*p) of the eukaryotic V-ATPase from *Saccharomyces cerevisiae* has been analyzed by small angle X-ray scattering (SAXS). The protein is divided into a 5.5 nm long egglike shaped region, connected via a 1.5 nm linker to a hook-like segment at one end. Circular dichroism spectroscopy revealed that subunit F comprises of 43% α-helix, 32% β-sheet and a 25% random coil arrangement. To determine the localization of the N- and C-termini in the protein, the C-terminal truncated form of F, F_{1-94} was produced and analyzed by SAXS. Comparison of the F_{1-94} shape with the shape of the entire subunit F showed the missing hook-like region in F_{1-94}, supported by the decreased D_{max} value of F_{1-94} and indicating that the hook-like region consists of the C-terminal residues (3). The NMR solution structure of the C-terminal peptide, F_{90-116}, was solved, showing an α-helical region between residues 103-113 (3). The F_{90-116} solution structure fitted well in the hook-like region of subunit F (3). In order to understand the structural features of F_{1-94} at the atomic level, X-ray crystallography was performed. The crystal structure of F_{1-94} reveals a Rossmann fold with alternating β-strands and α-helices (4). Elliptical shaped F_{1-94} has four β-strands which are surrounded by four α-helices. F_{1-94} contains two important loops spanning between α1-β2 ($_{26}$GQITPETQEK$_{35}$) and α2-β3 ($_{60}$ERDDI$_{64}$) which are present only in eukaryotic F subunits. Multiple sequence alignments of subunit F show that the $_{60}$ERDDI$_{64}$ loop is highly conserved among the eukaryotic V-ATPases (4). NMR spectroscopy of the entire subunit F confirmed the secondary structural features of the crystallographic structure F_{1-94} in solution as well as the C-terminal peptide, F_{90-116}. The heteronuclear NOE experiment shows that subunit F has a rigid core in the N-terminal domain, whereas α1 and α5 are more flexible in the solution (4). To understand the cross-talk between central stalk subunits with the neighboring subunits, a DF-heterodimer was generated. The DF-heterodimer binds to subunit *d* with a dissociation constant (K_d) of 52.9 µM as determined by ITC experiment (4). The DF-heterodimer

yielded crystals with a dimension of 0.13 mm x 0.10 mm x 0.04 mm, which diffracted maximum to 5 Å.

References

1. Benlekbir S., Bueler S. A. and Rubinstein J. L. (2012). *Structure of the vacuolar-type ATPase from Saccharomyces cerevisiae at 11-Å resolution*. Nat. Struct. Mol. Biol. **19**: 1356-1362.
2. Wieczorek H., Brown D., Grinstein S., Ehrenfeld J. and Harvey W.R. (1999). *Animal plasma membrane energization by prton-motive V-ATPase*. BioEssays. **21**: 637- 638.
3. Basak S., Gayen S., Thaker Y. R., Manimekalai M. S. S., Roessle M., Hunke C. and Grüber G. (2011). *Solution structure of subunit F (Vma7p) of the eukaryotic V_1V_O ATPase from Saccharomyces cerevisiae derived from SAXS and NMR spectroscopy*. Biochim. Biophys. Acta - Biomembranes. **1808**: 360-368.
4. Basak S., Lim, J., Manimekalai, M. S. S., Balakrishna, A. M. and Grüber, G. (2013) *Crystal- and NMR structures give insights into the role and dynamics of subunit F of the eukaryotic V-ATPase from Saccharomyces cerevisiae.* J. Biol. Chem. **288**: 11930-11939.

1. Introduction

1.1 Overview on eukaryotic V_1V_O ATPases (V-ATPases)

Chemical energy is very important for various vital cellular tasks, generally for metabolism purpose, transport of required substances as well as for various mechanical work like muscle movement. ATP is the major source of energy of all living cells. Chemical energy of ATP is stored in high potential phosphoanhydride bonds and removal of terminal phosphate during ATP hydrolysis liberates energy. Whenever the cell requires energy, ATP becomes converted to ADP+Pi by different ATPases. The vacuolar ATPase is one of them and is responsible for pH homeostasis and is involves in various important cellular events. V-ATPases are multi-subunit complexes which have a membrane bound region acting as a proton pump, thereby generating proton gradient across the transmembrane by using chemical energy released after the hydrolysis of ATP (1,2). Generally in eukaryotic cells, the V-ATPases are present in endomembrane systems, in some cases they are also present in the plasma membrane. V-ATPases acidify different subcellular organelles to maintain their activity like endosomes, lysosomes, Golgi-derived vesicles, synaptic vesicles, the plant vacuole, and chromaffin granules (3-5). Due to the acidification property, the enzyme involves in some crucial activities like endocytosis regulated by receptors, regulation of apoptosis and secretion of neurotransmitter *etc.* (6,7). The endocytic trafficking regulated by V-ATPase is complicated. As shown in figure 1.1, upon internalization at the plasma membrane, molecules pass through different endocytic compartments which have progressively decreasing pH. The pH of endosomal compartments is maintained by V-ATPase, which pumps protons into the lumen. In early endosome, the V-ATPase interacts with ADP-ribosylation factor-6 (Arf6) and ADP-ribosylation factor nucleotide site opener (ARNO). The process dependents on the intra-endosomal pH (8). Bafilomycin (Baf) inhibits V-ATPase activity, leading to the blockage of endocytic trafficking at early endosomes. During this blockage recruitment of endosomal COPI coat (eCOPI), which is depended on ARF1, is hampered. It also affects endosomal carrier vesicles (ECVs) formation (9). Transport of molecules in different compartments is controlled by ARF1 and ARF6. Molecules, which are not to be degraded, are recycled by recycling endosome regulated by ARF6 and transferred to plasma membrane (9). In lysosomes, the substrate degradation is carried out by cathepsins, which are synthesized in the trans-Golgi network. The activity of cathepsins depends on a low pH environment. Newly synthesized cathepsins are supplied to late endosome by mannose-6-phosphate receptor (MPR) as a carrier followed by the release of cathepsin in lysosome which then degrades the substrates (Figure 1.1).

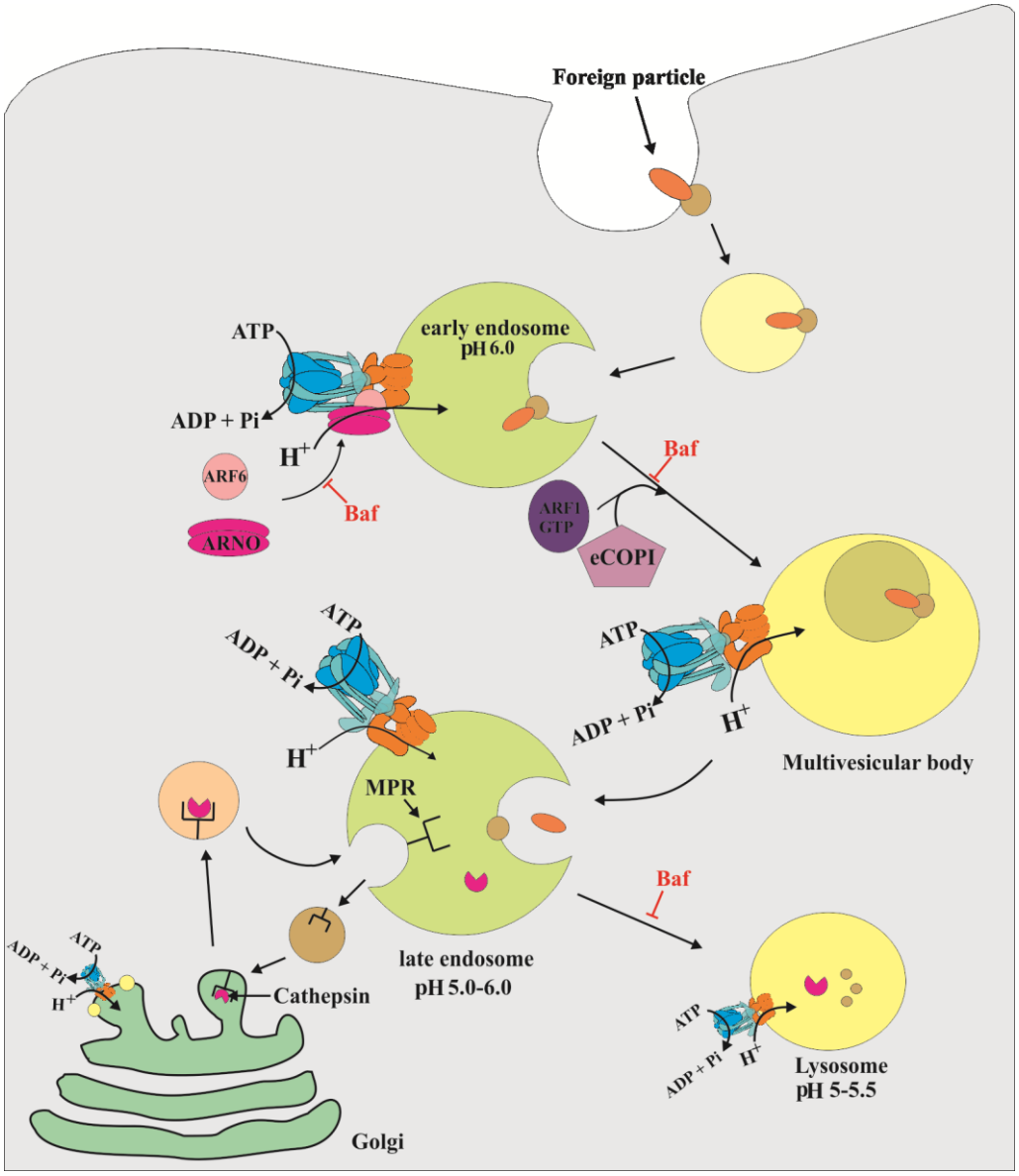


Figure 1.1: Roles of V-ATPases in substrate internalization (modified from Recchi and Chavrier, 2006 and Nishi and Forgac, 2002) (5,9). ARF1: ADP-ribosylation factor-1; ARF6: ADP-ribosylation factor-6; ARNO: ADP-ribosylation factor nucleotide site opener; Baf: Bafilomycin; eCOPI: endosomal COPI; MPR; mannose-6-phosphate receptor

Besides acidification and membrane trafficking, V-ATPases function as pH sensors in renal proximal tubular cells (10). In normal physiological condition, the V-ATPase maintains the pH of renal tubule. Maintenance of pH for sperm maturation and keeping sperms in immotile state during its passage through vas deferens and epididymis are also few of the functions that can be ascribed to V_1V_O ATPases (Figure 1.2A) (2,4,11-13). In higher organisms, V-ATPases can be located in the plasma membrane of renal intercalated cells and bone osteoclasts. Here, the enzyme pumps protons to the space outside of the cell through the plasma membrane, thereby functioning for example in urine acidification and bone remodeling (Figure 1.2B-C) (14).

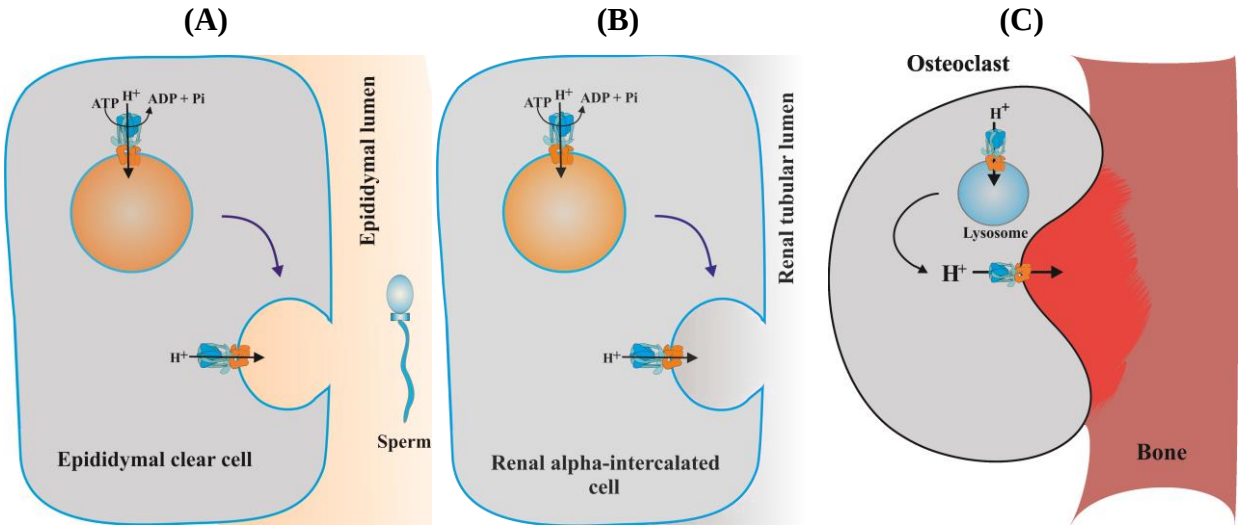


Figure 1.2: Physiological roles of V-ATPases. *(A)* V-ATPase maintains the pH in epididymal lumen, which helps in the maturation and viability of the sperm. *(B)* Maintenance of pH in renal tubular lumen by V-ATPase. Some certain mutations in subunit a4 or B1 isoforms increase acid secretion in distal tubule or collecting duct creates renal acidosis. *(C)* V-ATPase is essential for bone resorption (modified from Nishi and Forgac, 2002) (5,15).

Vacuolar ATPases are also essential components of lepidopteran insects such as the tobacco hornworm (*Manduca sexta*) larva and mosquito *(Aedes aegypti)* larva midgut. V-ATPase is responsible for the high alkalinization of the tobacco hornworm midgut. Proton motive force generated by V-ATPase activates, K^+/H^+ antiport in midgut and leads to high alkalinization with a pH higher than 11 (16-18). It has been proposed that high midgut pH in the insects helps in the breakdown of dietary tannins for their easier absorption (18). In another study, it was proved that clathrin-coated vesicles obtained from calf brain contain a proton pump, which utilized Mg^+ and ATP to generate proton motif force. It has been observed that the low pH generated by proton pump inside the vesicle is required for dissociation of receptor-ligand complex and delivered to lysosomes degradation (ligand) or recycled-back to the plasma membrane (receptors) (19). Synaptic vesicles take key responsibility during the neuronal transmission at the nerve terminals. It stores neurotransmitters followed by controlled exocytosis into the synaptic cleft. Among the various factors, responsible for the neurotransmitter accumulation in the synaptic vesicles, the generation of electrochemical gradient of protons mediated by the V-ATPase is crucial (20). Neurotransmitters are accumulated in synaptic vesicles, located at the nerve endings, and the V-ATPase has a prominent role in this accumulation process. Proton translocation against the electrochemical gradient leads to the transportation of neurotransmitters through the exchangers (antiporters) (Figure 1.3). When the presynaptic vesicles are docked to the synaptic membrane, neurotransmitters are released via membrane fusion channels, formed by the merging of two opposing V_O domains of V-ATPases (21). Under such conditions, free V_O

sectors are generated by established regulatory phenomenon of reversible disassembly, where soluble V_1 domain dissociates from membrane V_O sector to inhibit V-ATPase activity (2,22,23).

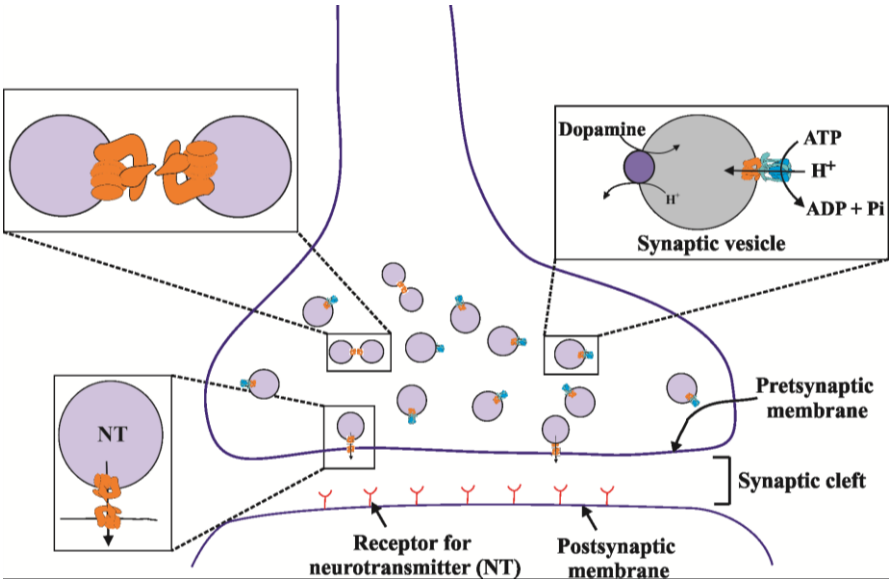


Figure 1.3: Energy coupling of uptake of transmitter by synaptic vesicles (modified from Nishi and Forgac, 2002; Morel, 2003 and Moriyama *et al.*, 1992) (5,6,20).

Recently, a diverse and unexpected function of V-ATPases in the left-right symmetry during the development in *Xenopus* has been reported (24). This mechanism is seen as an early upstream event in the process of development and is conserved in zebrafish (24). In one recent study, it has been reported that V-ATPase has also a role in Wnt signaling (nomenclature of Wnt is derived from the combination of *Drosophila* wingless (w) and the mouse homolog Int1 (nt) proteins (25)). Wnt/b-catenin signaling considered as a vital process in embryonic development, stem cell biology and even different diseases including cancer. During early central nervous system development of Xenopus, Wnt signaling is important for anterio-posterior patterning which is mediated by V-ATPase (26). V-ATPases are involved in different kind of crucial pathophysiology like cancer and diabetes (27-29). In one study, it has been shown that inhibition of V-ATPses lead to S-phase cell-cycle arrest and triggers apoptosis in cancerous epithelial cells (30). Acidity of extracellular compartment is an important factor in the cellular invasiveness, tumor proliferation and metastasis (28,29). It has been reported that in the metastatic state, cancerous cells have high density of V-ATPases compared to the cells that exhibit low metastatic rate. The V-ATPases are important regulators of intracellular pH in tumor cells, which have acidic cytosols due to high glycolytic rate (28). In the highly invasive cells, other than intracellular compartments, V-ATPase is over-expressed and translocated to the plasma membrane. Excessive expression of the plasma membrane V-ATPase gives the cells an edge in the increased proliferation and metastatic

activity (29). It has been studied that in a V-ATPase-transfected yeast cell model, the pH_{cyt} increased, because of the overexpression of proton pumps on the cell membrane causing cell proliferation and finally leads to tumorigenesis (31). Since cellular acidity is associated with apoptosis (32,33), this is triggered in many mammalian cell types due to the inhibition of V-ATPase (34). It has been reported that in poorly metastatic and angiogenic cells from diabetes mellitus model exhibit decreased plasma membrane V-ATPase and also reduced cell invasivenes (29). It is also reported that V-ATPase is involved in multi-drug resistance (MDR) (Figure 1.4) (28,35). Anti-cancer drugs like doxorubicin generally gets efflux by the cells due to the self defense mechanism, which can be inhibited by blocking the activity of V-ATPase (28). In cancerous cells, pH is always low because of high metabolism. During this time, V-ATPases are accumulated in plasma membrane and efflux proton from cells and subsequently increase the pH. As acidic pH is a critical factor to trigger cell death, inhibition of V-ATPase by specific inhibitors such as bafilomycin A_1 causes cellular apoptosis (36-38). Apoptotic death of neutrophils under *in vitro* culture condition is reported. Incorporation of G-CSF (Granulocyte Colony Stimulating Factor) can delay apoptosis via a mechanism which is involved in the up regulation of V-ATPases (33).

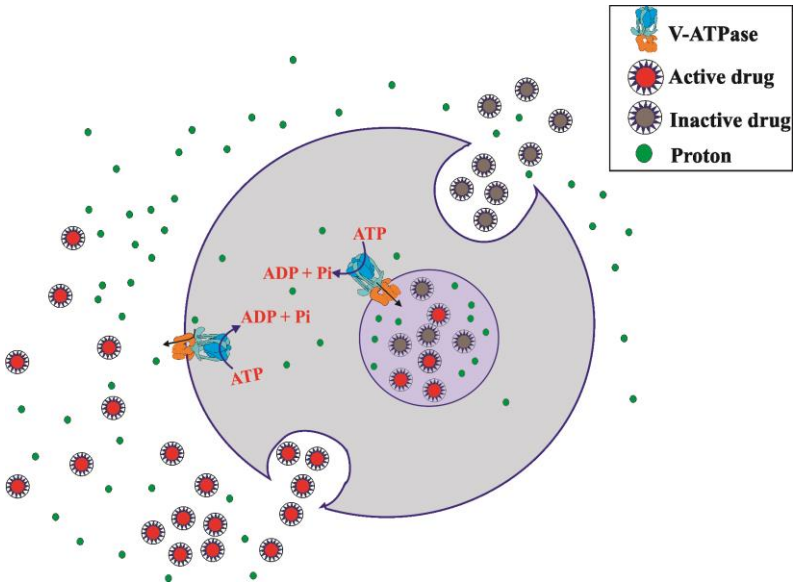


Figure 1.4: An example of multi drug resistance in cancer cells. where drug becomes inactivated followed by pump out from the cell (modified form Daniel *et al.*, 2013) (35).

1.2 Structural features of eukaryotic V_1V_O ATPase from *S. cerevisiae*

V-ATPases have mechanistic and overall structural similarities to F_1F_O- and A_1A_O ATP synthases; however, they have remarkable diversities in subunit composition, functioning as well as modes of regulation (12). V_1V_O-ATPases are solely ATP dependent proton pumps, in contrast to F/A-type ATP synthases, which produce ATP at the expense of ion motive force (imf) under physiological conditions in eukaryotes and archae, respectively (12). Also, V-ATPases are regulated by reversible disassembly phenomenon, causing dissociation of V_1V_O ATPase into the individual V_1 and V_O sectors, in response to various conditions for example glucose deprivation (23). This unique feature of V_1V_O ATPase was first observed in the internal membrane of yeast cells and apical membrane of *Manduca sexta* intestine (23,39). Later on it has been demonstrated that this reversible disassembly is very important for the regulation of lysosomal acidification of the dendritic cells (40). Such regulatory mechanism is absent in the A- and F-ATP synthases. V_1V_O ATPase of *S. cerevisiae* is composed of a ~570 kDa water-soluble hydrophilic V_1 ATPase, dressed in a hexameric fashion at the top of the enzyme. The hexameric domain is produced by the alternating arrangement of subunit A and B, around a central cavity that occupies an asymmetric seventh mass forming the central stalk (41,42). The V_1 domain is responsible for ATP hydrolysis. Upon ATP hydrolysis, energy is released and used by membrane-bound V_O domain leading to the H^+-translocation (1,2,12,43-45). The proposed subunit stoichiometry of V_1 domain is A_3:B_3:C_1:D_1:E_3:F_1:G_3:H_1 (12,46). The enzyme has an integral membrane embedded V_O domain, which is responsible for the ion translocation. The subunit stoichiometry of V_O domain is a_1:d_1:e:c_{4-5}:c'$_1$:c''$_1$ (Figure 1.5) (5,46-48). Higher eukaryotes lack c' subunit but contain accessory subunits Ac45 (49,50) and M8-9 (51,52) which are attached to the V_O domain. A heavily glycosylated V_O associated subunit M9.7 has been identified in the midgut and malpighian tubules of *Manduca sexta* (53).

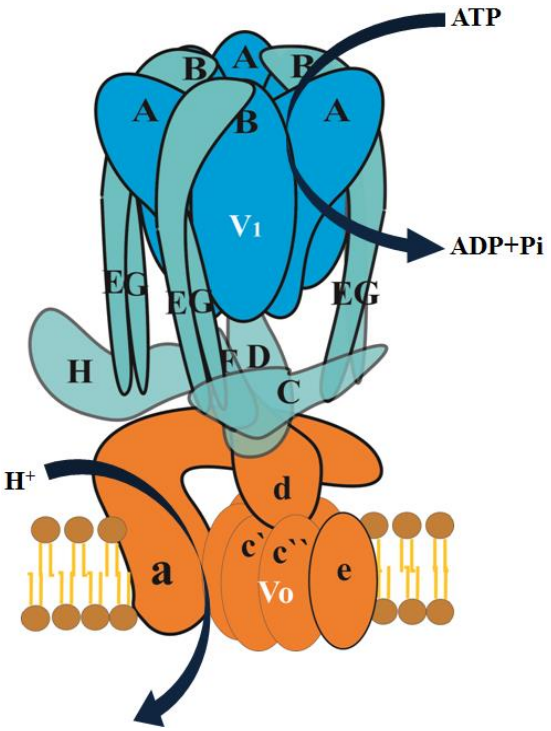


Figure 1.5: Modified schematic diagram of V-ATPases (12). ATP hydrolysed by V_1 catalytic domain of V-ATPase producing energy, which is used to transport ion to the cellular compartment across the membrane.

Table 1.1. Genes of *S. cerevisiae* V-ATPase subunits are shown along with the respective molecular weight and subunit function(s) (2,42,47,48,54-68).

Domain	Subunits	Yeast Gene	MW (kDa)	Subunit Function
V_1	A	*VMA1*	70	Catalytic site, coupling, peripheral stalk
	B	*VMA2*	60	Nucleotide binding, regulatory, actin and aldolase binding
	C	*VMA5*	44	Assembly, reversible dissociation, actin binding
	D	*VMA8*	29	Coupling process
	E	*VMA4*	33	Assembly, peripheral stalk, RAVE, aldolase binding
	F	*VMA7*	14	Coupling process
	G	*VMA10*	13	Activity, assembly, peripheral stalk
	H	*VMA13*	54	ATPase activity
V_O	*a*	*VPH1*, *STV1*	100	Proton pumping, targeting, assembly, peripheral stalk
	d	*VMA6*	40	peripheral stalk, activity
	c	*VMA3*	17	H^+-transport
	c'	*VMA11*	17	H^+-transport
	c''	*VMA16*	21	H^+-transport
	e	*VMA9*	9	unknown

Previously, the 3-D model of the *S. cerevisiae* vacuolar ATPase was derived by single particle transmission electron microscopy (14). The EM structure clearly showed the hexameric fashion of catalytic subunits (A_3B_3), where the non homologous region (NHR) of subunit A is protruding outwards (14). Recently, the cryo-EM structure of *S. cerevisiae* V-ATPase was solved at 11 Å resolution revealing the subunits arrangement in the enzyme more clearly (Figure 1.6) (56). Due to the clarity of the subunits rearrangement several *in-vitro* cross linking studies could be explained. The EM structure revealed that, there are three peripheral stalks made up of EG-heterodimers present in the V_1 domain and interestingly one unique peripheral stalk has different conformation than others. One side of the EG-heterodimer is linked to the catalytic subunit B and the other side is close to the V_O domain. In the enzyme structure subunit C, H and N-terminal domain of subunit *a* forms a collar like

region, oriented parallel to the V_O domain. Subunit C, which is parallel to the V_O domain, connects two peripheral EG-heterodimers.

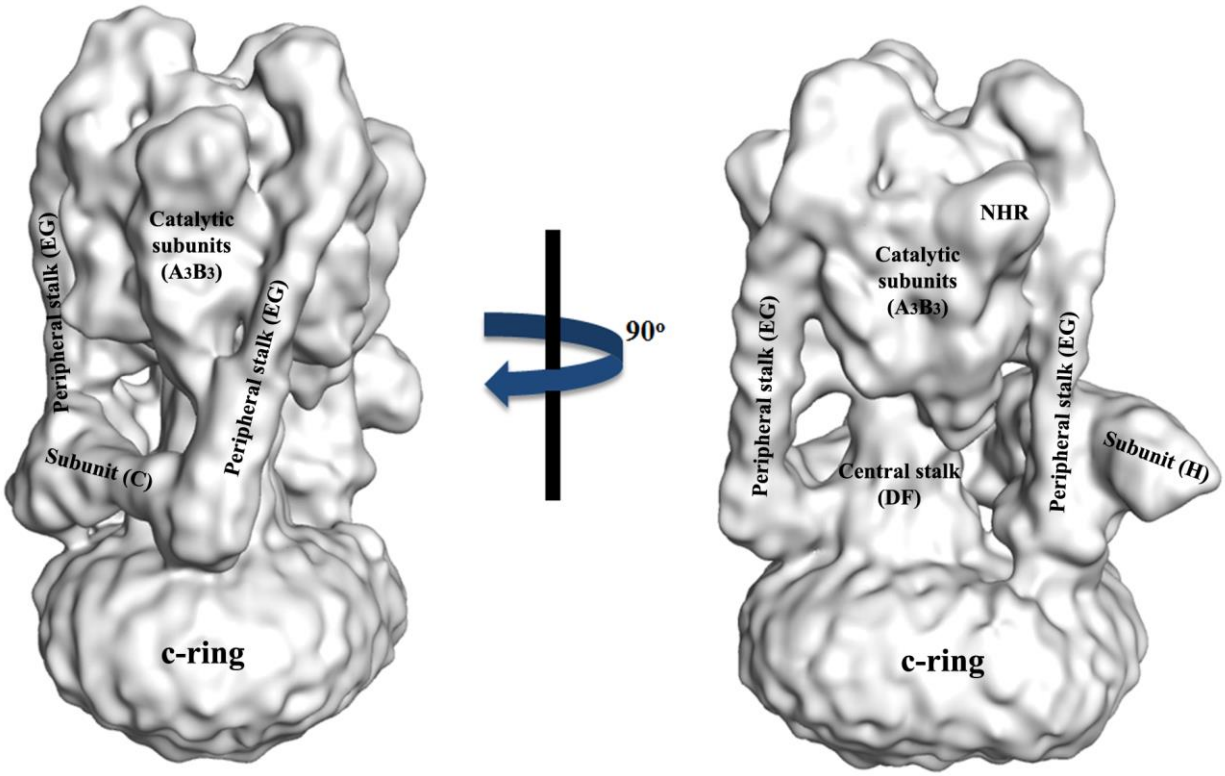


Figure 1.6: The 3-D cryo-EM structure of V_1V_O ATPase from *S. cerevisiae* solved at 11 Å resolution, shows subunits rearrangement in the intact complex (46). NHR: Non-homologous region.

1.3 High resolution insights into the subunit C, E, G and H of eukaryotic V_1V_O ATPase from *S. cerevisiae*

Inspite of subunit arrangements in *S. cerevisiae* V-ATPase being clear from the EM structure, a high resolution structure of the entire V-ATPase is still missing. There are some low as well as high resolution structures available of different isolated subunits. Atomic structures of subunit C (69,70), G (70-72), H (66) and E (70) from *S. cerevisiae* V_1 ATPase are solved so far, which belong to the stalk part of the enzyme.

The hydrated form of subunit C is a boot-shaped structure with 12.5 nm in length (73). As determined by X-ray crystallography, subunit C comprises of an upper head domain, composed of both α-helices and β-strands (residues 167-262), a globular foot domain (residues 1-55 and 320-392) and both are connected by an elongated helical neck domain (Figure 1.7) (69). Deprivation of glucose from the media or a drop in the ATP to ADP ratio eukaryotic V-ATPases undergo a reversible dissociation, where subunit C has been shown to play an important regulatory role (2,60,61). It has been revealed that subunit C is lost from V-ATPases during the process of disassembly. Furthermore, mutational analysis of subunit C have demonstrated a critical role of C, to balance the activities like V_1V_O assembly as well as ATPase activity (74). Previous study demonstrated that the C-terminus of subunit C has an ATP/ADP binding site (60). Upon binding of ATP/ADP subunit C might undergo structural alteration, regulating reversible disassembly of V-ATPase (60).

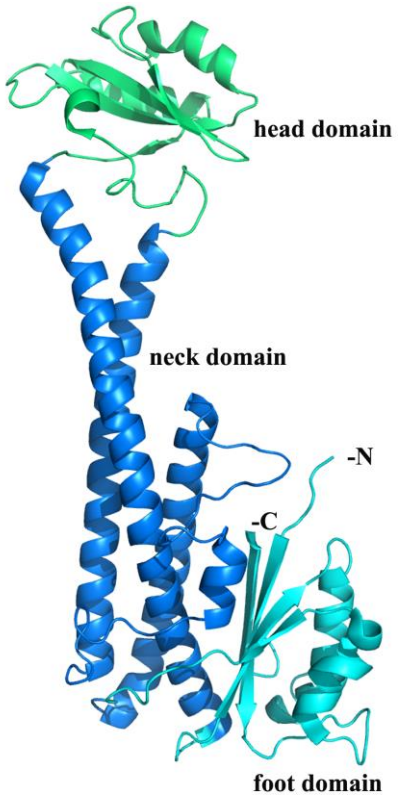


Figure 1.7: Crystal structure of subunit C (PDB ID: 1U7L (69)).

Subunit H is an α-helical subunit with a molecular weight of 54 kDa, whose N- and C-terminal domain form a shallow groove. Both domains are connected by a four-residue loop (Figure 1.8) (66). Subunit H is an essential regulatory component, believed to inhibit the V-ATPase activity of the dissociated V_1 domain during the process of reversible disassembly (65-67). It has been shown that the C-terminal part of subunit H is situated parallel to the N-terminal part of subunit *a* (46). It was hypothesized that during reversible disassembly, there is a structural rearrangement of both N- and C-terminal domain, which is mediated through the flexible linker segment regulating the ATPase activity of the enzyme (46). Subunit H is proposed to interact with the N-terminal domain of

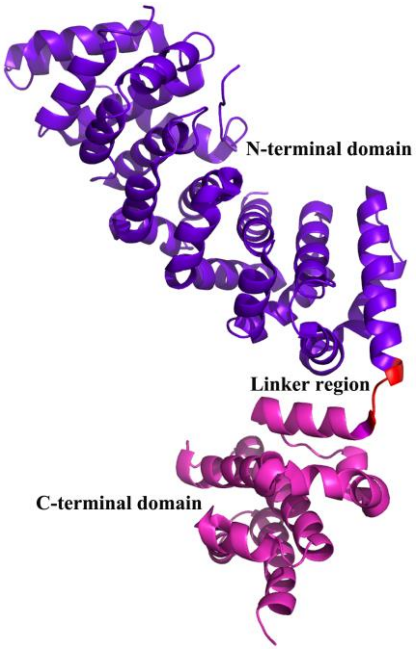


Figure 1.8: Cartoon representation of subunit H determined by X-ray crystallography (PDB ID: 1HO8 (66)).

subunit *a* and may act as a stator to prevent energy loss due to the futile rotation of V_1 domain relative to V_O (75).

Subunit G is an α-helical elongated structure and required for ATPase activity and assembly of V_1 and V_O (70-72). Mutations of G subunit can partially inhibit the reversible disassembly of V_1V_O ATPase (64). Genetic experiment and cross-linking data suggested that subunit E and G are situated close together (76). Subunit G forms a 150 Å long heterodimer with subunit E in a coiled coil fashion which has been determined recently (Figure 1.9) (70). The characteristic feature of subunit G is the presence of an unstructured glycine reach 'bulge', proposed to be important during the association of the EG-heterodimer and subunit C (70). Mutational analysis have demonstrated the crucial roles of subunit E in the assembly and catalytic activity of the A_3B_3 subcomplex (63). Subunit E has also been shown to be in close proximity to the catalytic subunits A, B and to subunit D (in nucleotide dependent manner) and has been described to be essential for the ATPase activity. Conformational changes of subunit E in the presence of Ca^{2+}ATP and -ADP are also reported in the *M. sexta* V_1 ATPase (77). The N-terminal domain of subunit E (residues 19-39) has been recently mapped to bind to subunit G (78), whereas the C-terminus of subunit E is believed to have a regulatory function in V_1V_O ATPases (63). Studies with the *M. sexta* V_1 ATPase have shown that subunit C, F and H dissociation increases the hydrolyzing activity of V_1 (79).

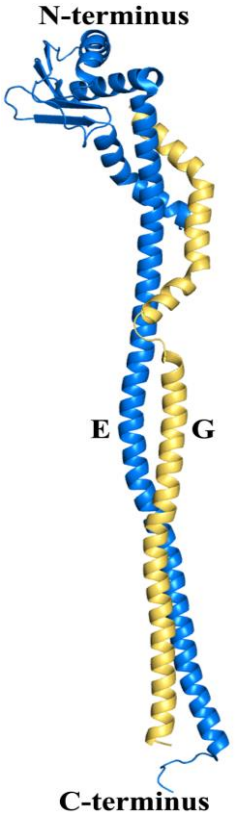


Figure 1.9: High resolution structure of EG-hererodimer determined by X-ray (PDB ID: 4DL0 (70)).

Docking of existing atomic structures of individual subunits from *S. cerevisiae* V-ATPase into the cryo-EM structure reveals many interacting interfaces (Figure 1.10) (46). Docking study has been done using Situs program (80). Initially subunit H was not fitting properly. EInemo program (81) was used to find out the optimum position of subunit H. The final overall correlation coefficient of fitted structures to the EM map was 0.8542. The C-terminal domain of subunit E is making contact with subunit B and the C-terminal part of subunit G is interacting with subunit E, not with the subunit B. The N-terminal region of the three different EG-heterodimers interact with different subunits in different manners. One peripheral stalk is attached only with the head domain of subunit C, where as the second EG-heterodimer is positioned close to the N-terminal region of subunit H and *a*. The third one is

making contact with the N-terminal domain of subunit *a* and the foot domain of subunit C (46).

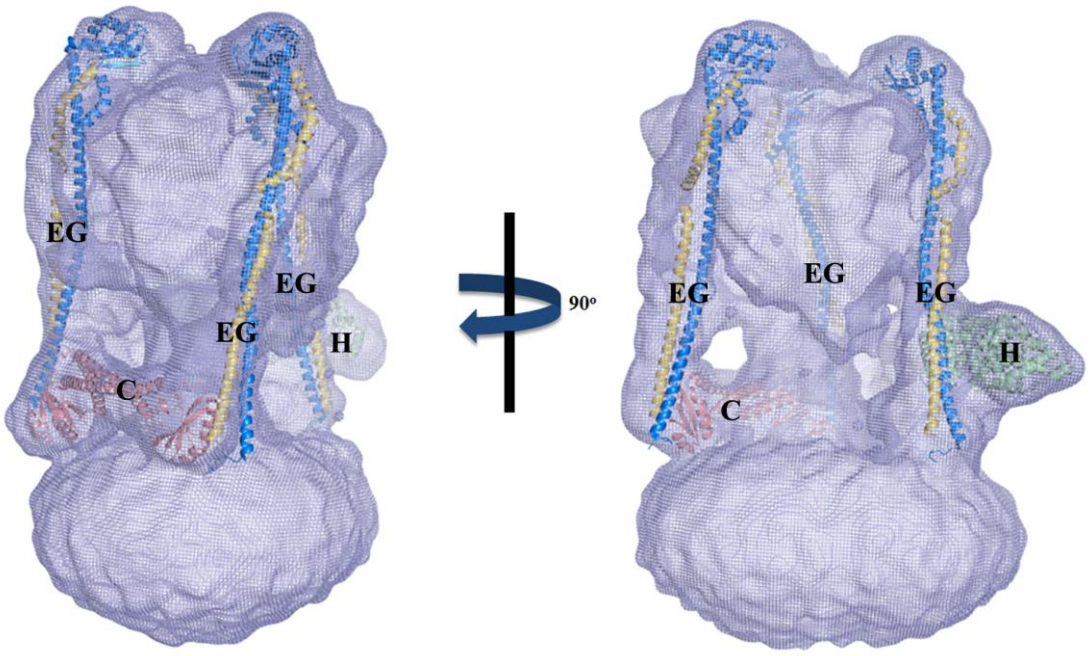


Figure 1.10: Docking of existing atomic structures of individual subunits (subunit E (Blue; PDB ID: 4DL0 (70)), G (Yellow orange; PDB ID: 4DL0 (70)), H (Green; PDB ID: 1HO8 (66)) and C (chocolate; PDB ID: 1U7L (69)) of the V-ATPase from *S. cerevisiae* into the 11 Å cryo-EM structure of the same enzyme (46).

Only one rotational position of the central stalk was captured in the cryo-EM structure although three structurally different orientations could be observed (46,82). Due to the static single orientation of the central rotor part, it is difficult to answer other probable interactions with neighbouring subunits. The uniqueness of eukaryotic V-ATPases is that it can undergo reversible disassembly in deprivation of glucose, which is not found in other ATPases. During reversible disassembly there are structural changes mainly at the stalk part (83). Thus, more structural, biochemical and biophysical work are needed to understand nature's most versatile enzyme, especially the rotary coupling central stalk region.

1.4 Outline of the thesis

During the coupling process, the "stalk" part of V-ATPases plays an important role. This part resides in the interface of an assembly of the V_1 and V_O, which contain subunits C-H and *a* and *d*, respectively. Subunit F and D are situated at the central part of the stalk and directly involved in the ATP hydrolysis and proton translocation (41). Structural alterations of subunit F have been reported during the catalysis of ATP (77,84). Subunit F interacts with subunits A, D, and E, which is nucleotide dependent (77,84). It has been proved that when Ser 381 residue of C-terminal part of subunit H is mutated to cysteine, it cross-linked to subunit in the

free V_1 ATPase, resulting in a drop of ATP hydrolysis (65,67). In a recent study, it has been demonstrated that subunit H has a flexible linker segment, allowing the C- and N-terminal domain to reorient during the reversible dissociation of V-ATPases (46). It was hypothesised that in the free V_1 domain the C-terminal domain of subunit H might cross-linked with the central rotor part which might then stop the rotation of central part, resulting drop in ATP hydrolysis (46). Due to the prime role of the central stalk, more structural information of subunit D and F from yeast V-ATPase are needed, which would throw light into the mechanistic details of the enzyme. In this thesis, subunit F and $F_{1\text{-}94}$ were over-expressed and purified. Solution X-ray scattering, X-ray crystallization technique and NMR spectroscopy, which are complementary to each other, were used to characterize subunit F. Initially, SAXS experiments were performed which showed domain features of subunit F and $F_{1\text{-}94}$. The structural superimposition of subunit F and $F_{1\text{-}94}$ showed that the major domain is made by N-terminal amino acids, whereas C-terminal amino acids are making minor domain positioning orthogonal to the major domain. As the C-terminal domain was predicted to be dynamic in solution (85), the atomic structure of the C-terminal domain of subunit F was solved by NMR, showing α helical content in solution. Thereafter, the high resolution structure of $F_{1\text{-}94}$ was solved by X-ray crystallography. The atomic structure of $F_{1\text{-}94}$ revealed the unique structural features, which are different from subunit F of the related A-ATP synthase. The structure showed two important loops, present only in eukaryotic subunit F. NMR experiments have been performed to assign the backbone amino acid residues. The secondary structure, of $F_{1\text{-}94}$ revealed by X-ray crystallography, is also confirmed by NMR spectroscopy in solution. It is also confirmed that the C-terminal helix is present in the entire subunit F. The ^{15}N-[^{1}H] heteronuclear NOE experiment was performed to identify the flexible regions in the entire subunit F. The data revealed the flexibility of the adjacent helix near to the unique loop, which might allow the movement of loop close to subunit H resulting in decrease ATP hydrolysis during disassembly. Subunit F is proposed to interact with other subunits, thus it is important to study the probable interaction at atomic level. Together with subunit F, subunit D also forms the part of central stalk. Therefore, both proteins were co-expressed. ITC experiment confirmed interaction between the subunit DF-heterodimer and subunit *d* with a K_d value of 52.9 µM. The weaker interaction reflects that during disassembly subunit DF is separated from the subunit *d*. Finally, crystals of the DF-heterodimer have been generated and diffracted to 5 Å resolution at an in-house Rigaku machine. The crystallization condition optimization is still going on.

2. Materials and methods

2.1 Materials

2.1.1 Chemicals

All the chemicals used for the study were at least or higher grade. Chemicals were purchased from the following companies:

2.1.1.1 Buffers and salts

	Sigma (St. Louis, MO, USA)
	USB (Sampscott, MA)
	Calbiochem (Darmstadt, Germany)
	Fluka (Sigma, Buchs Germany)
	Roth (Karlsruhe, Germany)
	Serva (Heidelberg, Germany)
DTT	Hoefer (San Francisco, CA, USA)
Ni^{2+}-NTA	QIAGEN (Hilden)
$Pefabloc^{SC}$	BIOMOL (Hamburg, Germany)
PMSF	Sigma (St. Louis, MO, USA)
$^{15}NH_4Cl$	Cambridge Isotopes Lab (USA)
^{13}C Glucose	Cambridge Isotopes Lab (USA)

2.1.1.2 Electrophoresis Chemicals

Agarose, SDS, Glycine, APS etc.	Bio-Rad (Hercules, CA, USA)
Antibiotics	Calbiochem, Sigma and Gibco (Invitrogen)
IPTG	Fermentas
BSA	GERBU (Heidelberg, Germany)

2.1.1.3 Molecular biology materials

Genomic DNA	*Saccharomyces cerevisiae* (AH104 strain)
Primers	1st Base and Research Biolabs (Singapore)
Peptides	NTU proteomics core facility, School of Biological Sciences, Nanyang Technological University, Singapore.
Pfu DNA polymerase	Fermentas (Glen Burnie, MD, USA)
NcoI, SacI, BamHI, NdeI, XhoI	Fermentas and New England Biolabs
T4 DNA Ligase	Fermentas and NEB
Miniprep Plasmid Kit	Qiagen (Hilden, Germany)

Nucleobond AX mediprep Kit	MN & Co (Düren, Germany)
Escherichia coli expression strains	DH5α, BL21 (DE3)
Plasmid vector	pET-9d1(86), pETDuet-1.

2.1.2 Chromatography

2.1.2.1 Ion Exchange

RESOURCE^{TM} Q (6 ml)	GE Healthcare (Uppsala, Sweden)

2.1.2.2 Gel filtration

Superdex 75 HR (10/30)	GE Healthcare (Uppsala, Sweden)

2.1.2.3 Instruments and accessories

Akta FPLC	GE Healthcare (Uppsala, Sweden)
Millex Filters (0.45 μM)	Millipore (Bradford, USA)
Syringe, needles and accessories	BD Biosciences

2.1.2.4 Protein concentration, estimation

Centriprep YM10	Millipore (Co-cork, Ireland)
Amicon ulta (3-30kDa)	Millipore (Co-cork, Ireland)
BCA Assay Kit	Pierce (Rockford, IL, USA)

2.1.3 Other instrumentation

PCR Thermocycler:	
Biometra T personal	Biometra
Biometra T gradient	Biometra
Sonoplus Sonicator	Bendelin
Micropulser Electroporator	Bio-Rad
Ultraspec 2100 Pro Spectromphotometer	Amersham Biosciences
Microcal iTC200	Microcal (Northampton, UK)

2.1.4 Computer software

Vector NTI 10.3.0	Invitrogen
SPARKY	University of California (San Francisco, USA)
Topspin 1.3	Bruker Biospin
CAYANA 2.1	Kimmo Paakkonen and Peter Günter (Japan)

TALOS	Gabriel Cornilescu *et al*, (NIH, USA)
MOLMOL	Koradi *et al*, (ETH, Zürich)
PyMOL v0.99	DeLano Scientific LLC, USA
HKL2000 package	OTWINOWSKI and MINOR, 1997
Refmac5 (CCP4 package)	MURSHUDOV *et al.*, 1997
MOLREP	VAGIN and TAPLYAKOV, 1997
PHASER	McCOY *et al.*, 2007
CNS	BRUNGER *et al.*, 1998
Coot	EMSLEY and COWTAN, 2004)
PROCHECK	LASKOWSKI *et al.*, 1993
Quantity One	Bio-Rad
CCP4	Winn *et al.*, 2011
NMRpipe	Delaglio *et al* (NIH, USA)
nmrDraw	Howard Hughes Medical Institute, USA

2.2 Methods

2.2.1 Production and purification of subunit F (*VMA7*p) and F_{1-94} of yeast V_1V_O ATPase

2.2.1.1 Induction test

Composition of 4x Lysis buffer:

250 mM Tris/HCl, pH 6.8
9.2 % SDS
40 % Glycerol
0.2 % Bromophenol Blue

Subunit F and F_{1-94} were cloned by Dr. Youg Raj Thaker, as described elsewhere (87). Five single colonies were randomly picked from agar plate and grown at 37 °C in an incubator shaker (Infors HT Minitron shaker) until the OD_{600} of the cells reached 0.6 - 0.7. Shaking was maintained at 170 – 200 rpm. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to final concentration of 0.5-1 mM, to induce protein production and kept for 2 – 3 h at 37 °C. The concentration of IPTG depends on the characteristics of the protein being produced and the conditions that have been optimized for the respective protein expression. Respective controls were left uninduced by not using IPTG under identical conditions. The same amount of cells were pelleted down by centrifugation at 13000 x g for 5 min and resuspended with 60 µl of 1x lysis buffer in presence of 10 mM DTT. Resuspended cells were heated at 95 ºC for 5 min and 15 – 20 µl of each sample was loaded onto a 17% SDS polyacrylamide gel. It has been observed that the expression was high in the induced cultures ((+) supplemented with IPTG) than uninduced cultures ((-) not supplemented with IPTG). A representative induction gel is shown below in figure 2.1.

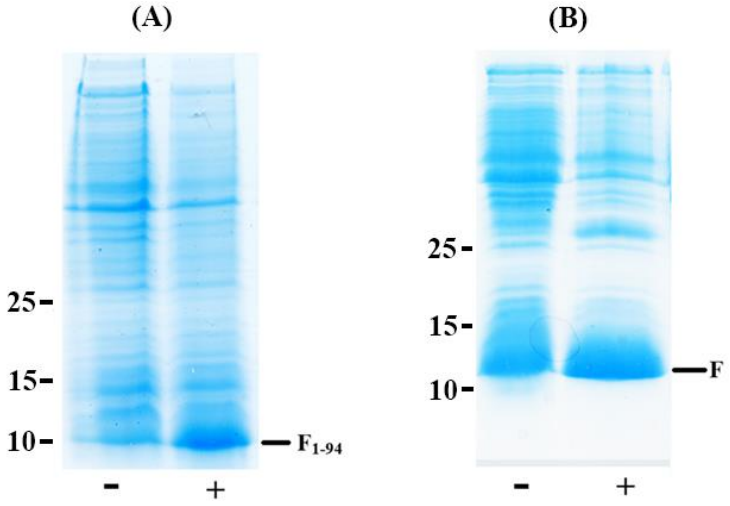


Figure 2.1: Induction test of subunit F_{1-94} *(A)* and F *(B)*. uninduced (-) and induced (+) culture of subunit F and F_{1-94}. Induction was done with 0.5 – 1 mM IPTG at 37 °C. 1x lysis buffer was used and heated at 95°C for 5 min and run on a 17% SDS gel.

2.2.1.2 Solubility test of recombinant proteins

After optimization of the IPTG concentration, temperature and shaking condition, the solubility of the produced proteins were tested in various buffers using 50 ml of culture. Cell suspensions were first divided into 5 equal fractions and then pelleted down by centrifuging at 10,000 x g for 10 min. Pellets in different fractions were suspended in different buffers such as (A) 50 mM Hepes, 300 mM NaCl, pH 7.0, (B) 50 mM Hepes, 150 mM NaCl, pH 7.5, (C) 50 mM Tris, 150 mM NaCl, pH 7.5 and (D) 25 mM Phosphate buffer, 200 mM NaCl, pH 6.8. Freshly prepared Pefabloc^{SC} (2 - 8 mM) in water and/or (1 - 2) mM PMSF (dissolved in isopropanol) were added into the corresponding buffer to inhibit proteases. Reducing agent such as DTT (1 mM) or β-mercaptoethanol (1 mM) was used in some preparations. Resuspension was sonicated at 10% power with KE 72 tip (Bandelin Sonoplus) three times for one min, each with a cooling interval of two min between each cycle. Afterwards samples were centrifuged at 10,000 x g for 7 min. Supernatant was then transferred to a fresh tube and the pellet was resuspended in the same buffer. All steps were carried out on ice. Pellet and supernatants were subsequently loaded onto a 17% SDS gel.

2.2.1.3 Purification of subunit F and $F_{1\text{-}94}$

After successful solubility tests, the proteins were purified in larger amount. The cells were grown in LB medium containing kanamycin (30 μg/ml) and shaken at 37 °C until an optical density OD_{600} of 0.6 - 0.7 was reached. To induce the expression of His_6-F, the cultures were supplemented with IPTG to a final concentration of 1 mM. Cells were grown for another 4 h at 37 °C and then harvested at 8 500 x *g* for 12 min, 6 °C. The cells were lysed by sonication for 3 x 1 min in buffer A (50 mM HEPES, pH 7.0, 300 mM NaCl, 2 mM PMSF and 2 mM Pefabloc^{SC} (BIOMOL)). The lysis step was done on ice. Precipitated material was pelleted down by centrifugation (Centrifuge 5810R, eppendorf) at 10 000 x *g* for 35 min. The supernatant was then filtered (0.45 μm; Millipore) and His-tagged protein was mixed with Ni^{2+}-NTA resin and allowed to bind to the matrix for 1.5 h at 4 °C. Subsequently the sample was passed through column containing Ni^{2+}-NTA resin and the flow through was collected. Afterwards bound protein was eluted with an imidazole-gradient (25 - 400 mM) in buffer A. Fractions containing His-tagged subunit F were identified by running an SDS-PAGE (88). Required fractions of protein were then pooled together and concentrated using Centricon YM-3 (3 kDa molecular mass cut off) spin concentrators (Millipore). Imidazole was removed by gel filtration chromatography using a Superdex 75 HR 10/30 column (GE Healthcare), where 50 mM HEPES (pH 7.0), 300 mM NaCl and 5 mM EDTA was used as a mobile phase.

F_{1-94} was purified following the same protocol described before. The purity and homogeneity of all protein samples were analyzed by SDS-PAGE (88). SDS-gels were stained with Coomassie Brilliant Blue G250 and bicinchoninic acid assay (BCA; Pierce, Rockford, IL., USA) was performed to determine protein concentrations.

2.2.2 Quantification of proteins by bicinchoninic acid (BCA) method

Bovine serum albumin (BSA) is used as a standard for protein quantification by BCA method. BCA reagent is available in a kit from Pierce (Rockford, IL, USA.) and the protein concentration was measured as per manufacture's protocol. Generally, two dilutions of protein samples were taken. To minimize any error in the measurement, each of the samples was measured in triplicates. Optical density was measured at 562 nm in a spectrophotometer. Water or buffer was used as a blank. The standard curve was then drawn from OD_{562}. From the standard curve final protein concentration was calculated.

2.2.3 Circular dichroism (CD) spectroscopy

Steady state CD spectra were measured in the far UV-light (190 – 260 nm) using a *Chirascan* spectropolarimeter (Applied Photophysics). Spectra were collected in a 60 µl quartz cell (Hellma) with a path length of 0.1 mm, at 20 °C and a step resolution of 1 nm. The readings were average of 2 sec at each wavelength and the recorded ellipticity values were the average of triplicate for each sample. CD spectroscopy of subunit F and the truncated F_{1-94} (2.0 mg/ml) were performed in a buffer of 50 mM HEPES (pH 7.0) and 300 mM NaCl. The CD spectrum of the buffer was used as a blank and subtracted from the spectrum of the protein. CD values were converted to mean residue molar ellipticity (Θ) in units of degree cm^2 $dmol^{-1}$ using the software *Chirascan Version 1.2*, Applied Photophysics. This baseline corrected spectrum was used as input for computer methods to obtain predictions of secondary structure. The following algorithms were used to analyze the CD spectrum: Varselec (89), Selcon (90), Contin (91), K2D (92) (all methods as incorporated into the program Dicroprot (93) and Neural Net (94).

2.2.4 Solution X-ray scattering experiments and data analysis of subunit F and F_{1-94}

Small angle X-ray scattering (SAXS) data for subunit F and F_{1-94} were collected by following standard procedures on the X33 SAXS camera (65,95) of the EMBL Hamburg located on a bending magnet (sector D) on the storage ring DORIS III of the Deutsches Elektronen Synchrotron (DESY). A photon counting Pilatus 1M pixel detector (67 x 420 mm^2) was used at a sample - detector distance of 2.4 m covering the range of momentum

transfer $0.1 < s < 4.5$ nm^{-1} ($s = 4p \sin(q)/l$, where q is the scattering angle and l = 0.15 nm is the X-ray wavelength). The s-axis was calibrated by the scattering pattern of Silver-behenate salt (d-spacing 5.84 nm). The scattering from buffer alone was measured before and after each sample measurement and the average of the scattering before and after each sample is used for background subtraction. A range of protein concentrations (2.5 to 7.4 mg/ml) was measured for both *S. cerevisiae* subunit F and $F_{1\text{-}94}$, to assess any concentration-dependent inter-particle effects. Both proteins have been measured in a buffer, composed of 50 mM HEPES (pH 7.0), 300 mM NaCl and 1 mM DTT. The protein as well as the buffer samples have been injected automatically using the sample-changing robot for solution scattering experiments at the SAXS station X33 (96). All the data processing steps were performed automatically using the program package PRIMUS (97). The forward scattering *I(0)* and the radius of gyration R_g were evaluated using the Guinier approximation (98) assuming that for spherical particles at very small angles ($s < 1.3/R_g$) the intensity is represented by $I(s) = I(0) \exp(-(sR_g)^2/3)$. These parameters were also computed from the entire scattering patterns using the indirect transform package GNOM (99), which also provide the distance distribution function *p(r)* of the particle as defined:

$$p(r)=2\pi\int I(s)sr\sin(sr)ds$$

The molecular mass of both proteins were calculated by comparison with the forward scattering from the reference solution of BSA. From this procedure a relative calibration factor for the molecular mass (MM) can be calculated using the known molecular mass of BSA (66.4 kDa) and the concentration of the reference solution by applying

$$MM_p = I(0)_p / c_p \times \frac{MM_{st}}{I(0)_{st} / c_{st}}$$

where $I(0)_p$, $I(0)_{st}$ are the scattering intensities at zero angle of the studied and the BSA standard protein, respectively, MM_p, MM_{st} are the corresponding molecular masses and c_p, c_{st} are the concentrations. Errors have been calculated from the upper and the lower I(0) error limit estimated by the Guinier approximation.

Low-resolution models of subunit F and $F_{1\text{-}94}$ were built by program DAMMIN (100), which represents the protein as an assembly of dummy atoms inside a search volume defined by a sphere of the diameter D_{max}. Starting from a random model, DAMMIN employs simulated annealing to build a scattering equivalent model fitting the experimental data $I_{exp}(s)$ to minimize discrepancy:

$$\chi^2 = \frac{1}{N-1}\sum_j \left[\frac{I_{\exp}(s_j) - cI_{calc}(s_j)}{\sigma(s_j)}\right]^2$$

where N is the number of experimental points, c is the scaling factor and $I_{calc}(s_j)$ and $\sigma(s_j)$ are the calculated intensity and the experimental error at the momentum transfer s_j, respectively. *Ab initio* shape models for subunit F and $F_{1\text{-}94}$, respectively were obtained by superposition of ten independent DAMMIN reconstructions for each subunit by using the program package DAMAVER (101).

2.2.5 Cloning of $F_{1\text{-}94}$ mutants of V_1V_O ATPase from *S.cerevisiae*

$F_{1\text{-}94}$ mutants were generated according to the method of Ho *et al.*, 1989 (102), by using four primers: flanking forward primer with a *NcoI* restriction site (*a*), flanking reverse primer with a *SacI* restriction site (*d*) and internal primers with site of mutation (*c* and *b*) (Table 2.1). The DNA fragments encoding the N-terminal extein (232 nucleotides) and C-terminal extein (109 nucleotides) were amplified by polymerase chain reaction (PCR) method with the primers *a*-*b* and *c*-*d*, respectively (Table 2.1). The DNA of subunit F construct (*VMA7*) of *S. cerevisiae* was used as template for the PCR. Two separate PCR reactions were setup in a total volume of 50 µl as follows:

Reagents	**Amount**
Pfu buffer (10x)	5 µl
dNTP's (2 mM)	1.5 µl
Primers (100 µM)	2 x 0.5 µl
Template (genomic DNA)	1 µl
MilliQ water	40.5 µl
Pfu DNA polymerase	1 µl

The flanking primers (*a* and *d*) hybridized at each end of the subunit $F_{1\text{-}94}$ sequence, thereby incorporating *NcoI* and *SacI* restriction sites (underlined), and an internal primer (forward primer *b* and Reverse primer *c*), that hybridized at the internal site (Table 2.1). The PCR reaction mixture was set up on ice and the lid of the PCR thermocycler (Biometra T personal) was preheated to 99 ºC, before reaction tubes were placed inside the machine. The cycle denaturation, annealing and extension steps were repeated for 30 cycles. Total time taken for the PCR program was 1h 43 min 49 sec. The following PCR thermocycler program was used for the amplification:

Cycle steps	Temperature	Time	
Lid	99 ºC		
Initial denaturation	96 ºC	3 min	
Cycle denaturation	96 ºC	30 sec	
Annealing	58 ºC	45 sec	30 cycles
Extension	72 ºC	60 sec	
Final extension	72 ºC	5 min	
End	4 ºC		

In separate PCRs, two fragments of the subunit $F_{1\text{-}94}$ gene were amplified by using primers *a* and *b* and primers *c* and *d*, respectively. The amplified fragments coding the N-terminal and C-terminal exteins were fused together via a second PCR using the primers *a* and *d*. The overlap allowed one strand from each fragment to act as a primer on the other, and extension of this overlap by flanking primers (*a* and *d*) in second PCR reaction resulted in the product. To check the amplification and purity of the desired product, 5 µl of PCR amplified product was applied onto the analytic agarose gel (1.5%). If the product was found sufficiently pure, the remaining reaction mixture was applied onto preparative agarose gel and purified by gel extraction kit (QIAGEN) according to the manufacturer's protocol. The purified PCR product was finally eluted in 50 µl volume of elution buffer (Qiagen). The primers (*b* and *c*) contain the mutation. A flow chart summarizing the procedures involved in site directed mutagenesis is given in Figure 2.2.

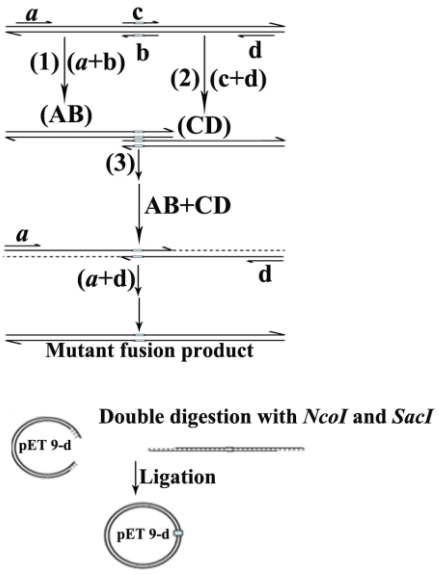


Figure 2.2: A flow chart showing the procedures involved in site directed mutagenesis using overlap extension polymerase chain reaction method (102). The same strategy was used to generate the mutants of subunit $F_{1\text{-}94}$.

After amplification, the PCR product was stored at -20 °C. 5 µl of the sample was applied onto an analytical agarose gel (1.5%) to check the quality of the PCR amplified product. The gel was soaked in ethidium bromide solution for 5 min to observe the PCR product under UV light. The size of the PCR product was confirmed with the appropriate DNA marker. In order to purify PCR the product, the remaining reaction mixture was applied onto a preparative agarose gel and the portions corresponding to the correct PCR product were cut out immediately and purified by gel extraction kit (QIAGEN) as per the manufacturers' protocol. The product was finally eluted in 30 - 50 µl volume of water or EB buffer (Tris-HCl buffer of pH 8.0). Purified DNA was double digested overnight at 37 ºC (SW22 Shaking Water Bath, Julabo, Germany) by taking 1 µl each of *Nco*I and *Sac*I enzymes. After overnight digestion, the reaction mixture was again purified by enzyme reaction purification kit from QIAGEN. Subsequently, ligation reaction between vector (Figure 2.3) and insert (amplified PCR product) was setup at room temperature. The reagents used were as follows:

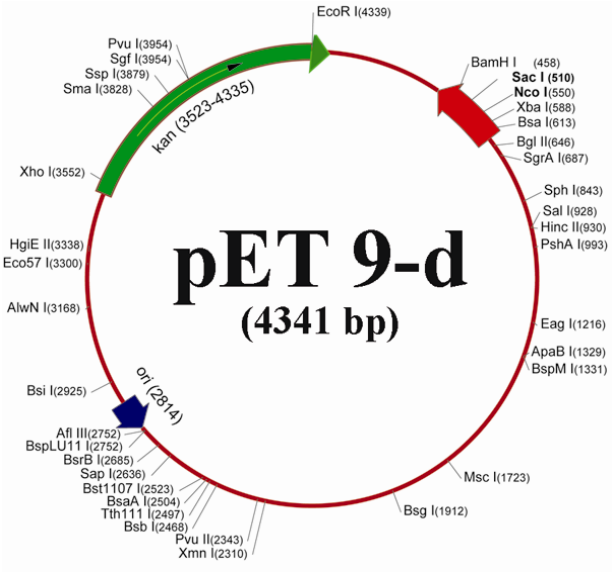


Figure 2.3: Modified pET-9d1 (+) (86) vector showing full map and unique multiple cloning sites (MCS) that has been used to clone various genes.

Ligase buffer (10x)	1.5 µl
Vector (V)	50-100 ng
Insert (I)	Variable (1-5 times of vector)
T4 Ligase	1 µl
MilliQ water	variable
Total	15 µl

Before setting up the reaction, ligase buffer was vortexed to completely dissolve $MgCl_2$ precipitate. The ligation mixture was kept at room temperature for 1.5-2 hrs. After 2 hrs, the ligation product was purified by butanol precipitate method (103). In the reaction mixture, 80 µl Milli Q water was added to make a final volume of 100 µl. 1 ml of butanol was then added into the mixture and mixed thoroughly. Then the mixture was centrifuged at

13,000 x *g* (MiniSpin® plus, eppendorf) for 10 min. Supernatant was discarded gently and the pellet was resuspended in 500 µl of 70% ethanol and centrifuged again for 10 min. Supernatant was removed and the pellet was re-dissolved in 6 µl of elution buffer (10 mM Tris-HCl, pH 8.5) (EB buffer, QIAGEN) or Milli Q water. 6 µl of ligation mixture was used for the transformation into *E. coli* DH5α cells. On the next day, several colonies were picked from the transformation plate containing appropriate antibiotic marker and then grown at 37 °C overnight. The plasmid DNA was isolated using a standard protocol (QIAGEN mini-prep kit). Purified plasmid DNA was then double digested for 2.5 hrs and applied onto an 1.5% agarose gel, in order to confirm the ligated insert. Size of insert and vector were compared with appropriate controls and markers. After confirming the presence of the insert, purified vector was sent for sequencing. Verified plasmid was finally transformed into *E. coli* BL21 (DE3) cells for protein production and they were grown on 30 µg/ml kanamycin-containing Luria-Bertoni (LB) agar plates. *E. coli* BL21 expression strains were purchased from Novagene (Darmstadt, Germany).

A similar approach has been taken to obtain different mutant constructs. The sequences of flanking primers (*a* and d) and internal primers (b and c) used for the mutants are enlisted in table 2.1. The primers (b and c) contain the mutation.

Table 2.1: Primer sequences used for the generation of different $F_{1\text{-}94}$ mutants. The primers were synthesized at *1st Base Pte Ltd*, Singapore.

Constructs		**Primer sequence (5'-3')**
$F_{1\text{-}94}$I8M	Primer *a*	GTAACTAAACCATGGCTGAGAAACGTACTCTTATGGCT
	Primer d	ATCGAGCTCTTAAATTTCTAAAATAGCAGGGAACGC
$F_{1\text{-}94}$I11M	Primer *a*	GTAACTAAACCATGGCTGAGAAACGTACTCTT
	Primer d	ATCGAGCTCTTAAATTTCTAAAATAGCAGGGAA
	Primer c	TGATGGCTGACGAAGATACTACAACTGGTTTA
	Primer b	TAAACCAGTTGTAGTATCTTCGTCAGCCATCA
$F_{1\text{-}94}$L20M	Primer *a*	GTAACTAAACCATGGCTGAGAAACGTACTCTTATAGCTGTG
	Primer d	ATCGAGCTCTTAAATTTCTAAAATAGCAGGGAACGC
	Primer c	GAAGATACTACAACTGGTATGTTGTTAGCCGGGATTG
	Primer b	CAATCCCGGCTAACAACATACCAGTTGTAGTATCTTC
$F_{1\text{-}94}$L67M	Primer *a*	GTAACTAAACCATGGCTGAGAAACGTACTCTTATAGCTGTG
	Primer d	ATCGAGCTCTTAAATTTCTAAAATAGCAGGGAACGC
	Primer c	GACGATATTGCCATCATGCTAATCAACCAACATATC

	Primer b	GATATGTTGGTTGATTAGCATGATGGCAATATCGTC
F_{1-94}L68M	Primer *a*	GTAACTAAACCATGGCTGAGAAACGTACTCTTATAGCTGTG
	Primer d	ATCGAGCTCTTAAATTTCTAAAATAGCAGGGAACGC
	Primer c	CGATATTGCCATCCTTATGATCAACCAACATATCGC
	Primer b	GCGATATGTTGGTTGATCATAAGGATGGCAATATCG
F_{1-94}I69M	Primer *a*	GTAACTAAACCATGGCTGAGAAACGTACTCTTATAGCTGTG
	Primer d	ATCGAGCTCTTAAATTTCTAAAATAGCAGGGAACGC
	Primer c	GATATTGCCATCCTTCTAATGAACCAACATATCGC
	Primer b	GCGATATGTTGGTTCATTAGAAGGATGGCAATATC
F_{1-94}I73M	Primer *a*	GTAACTAAACCATGGCTGAGAAACGTACTCTTATAGCTGTG
	Primer d	ATCGAGCTCTTAAATTTCTAAAATAGCAGGGAACGC
	Primer c	TTCTAATCAACCAACATATGGCGGAAAACATAAGAG
	Primer b	CTCTTATGTTTTCCGCCATATGTTGGTTGATTAGAA

2.2.6 Peptide synthesis

The C-terminal peptide F_{90-116} of subunit F from *S. cerevisiae* was synthesized by Liberty Automatic Microwave Peptide Synthesizer (CEM) at the Division of Chemical Biology and Biotechnology, School of Biological Sciences, Nanyang Technological University, Singapore. The peptide was then purified by reversed phase high pressure liquid chromatography on a Dynamax C-18 column (Varian Inc.), eluted with a linear 5–100% gradient of acetonitrile in 0.04% aqueous trifluoroacetic acid. The identity of the purified peptide was confirmed by MALDI-TOF MS (4800 MALDI TOF/TOF, Applied Biosystems/MDS Sciex). The purity of the peptide was analysed by electrospray ionization-MS.

2.2.7 NMR spectroscopy of peptide F_{90-116}

All NMR experiments were performed on a Bruker DRX 600 MHz spectrometer equipped with a cryoprobe. For structure determination, appropriate amounts of peptide F_{90-116} was dissolved in 25 mM phosphate buffer, pH 6.5. TOCSY and NOESY spectra of the peptide were recorded with mixing times of 80 and 300 ms, respectively, at a temperature of 25 ºC. TopSpin (Bruker Biospin) and Sparky suite (104) of programs were used for spectra processing, visualization and peak picking. Standard procedures based on spin-system identification and sequential assignment were adopted to identify the resonances (105). Inter proton distance was obtained from the NOESY spectra. NOESY peaks were categorized as

strong, medium and weak based on the signal intensity and were translated into distance constraints as 3.0 Å, 4.0 Å and 5.0 Å, respectively. Dihedral angle restraints as derived from TALOS (106) were employed to generate the three dimensional structure of the peptide in the CYANA 2.1 package (107). In total 100 structures were calculated and an ensemble of 10 structures with lowest total energy was chosen for structural analysis.

2.2.8 Crystallization of $F_{1\text{-}94}$

Crystallization of $F_{1\text{-}94}$ was attempted using vapour diffusion method. Sparse matrix screens from Hampton research, Molecular dimensions and Emerald biosystems were used for the initial screening. Both hanging and sitting drops were set up in VDX plates (Hampton Research) by mixing 2 µl of the purified $F_{1\text{-}94}$ protein (3-8 mg/ml) in buffer B (50 mM HEPES, pH 7.0, 300 mM NaCl, 5 mM EDTA) with 2 µl of the precipitant solution and incubated at 18 °C (108). Initial crystals were seen in Hampton research crystal screen 1 in both hanging and sitting drops. Once these crystals were confirmed to be protein crystals by testing in the in-house Rigaku machine, further optimization of the crystallization condition was done in order to decrease the nucleation and promote further crystal growth. Cryoprotectant solution was also optimized by testing different percentages of glycerol and final cryoprotectant solution comprised of 30 % (v/v) polyethylene glycol 4000 (PEG 4000), 0.1 M $MgCl_2$, 6xH_2O, 0.1 M Tris/HCl (pH 8.8), 25 % (v/v) glycerol. The crystals were quickly dipped in the cryoprotectant solution and were flash-frozen in liquid nitrogen at 100 K. Optimized crystals were initially tested in the in-house Rigaku machine. Finally crystals were measured at the National Synchrotron Radiation Research Center (NSRRC), Hsinchu, Taiwan for data collection.

Attempt was also taken to crystallize the entire subunit F by setting up drops in different crystallization screen but no crystal was obtained.

2.2.9 Production, purification and crystallization of selenomethionyl $F_{1\text{-}94}$ and its mutant forms

Following digestion with *NcoI* and *SacI*, the final PCR product of mutant was ligated into the pET9d1-His_6 vector (86). The resulting coding sequence included an N-terminal MKHHHHHHP-tag. The mutation was verified by DNA sequencing. The verified plasmids were finally transformed into *E. coli* BL21 (DE3) cells (Stratagene), which were grown on 30 mg/ml kanamycin-containing Luria-Bertani (LB) agar plates. The selenomethionine proteins were produced in *E. coli* strain BL21 (DE3) cells by growing a 5 ml culture overnight at 37 ºC

in LB medium containing kanamycin. The cells were pelleted by centrifugation and were washed twice with M9 minimal medium. The cell suspension was used to inoculate pre-warmed (310 K) M9 minimal medium and the cells were grown at 37 ºC. When the OD_{600} reached 0.6, the minimal medium was supplemented with 100 mg/l lysine, phenylalanine and threonine, 50 mg/l isoleucine, leucine and valine, and 50 mg/l selenomethionine. This medium was then shaken for a further 15 min. Thereafter, protein production was induced by the addition of isopropyl β-D-1-thiogalactopyranoside to a final concentration of 1 mM and growth was continued at 20 ºC overnight. The selenomethionyl F_{1-94} was purified in 50 mM HEPES, pH 7.0, 300 mM NaCl, 5 mM EDTA and 1mM DTT buffer. Crystallization drops of selenomethionine substituted F_{1-94} protein proteins were setup in the final optimized crystallization condition of native F_{1-94} (see section 2.2.8). The protein produces reproducible crystals. Due to the lack of suitable search models, molecular replacement did not work. Therefore, crystallization drops of all mutant proteins were setup, which produced either small crystals or precipitation except F_{1-94}I69M mutant.

2.2.10 Crystallization of selenomethionyl F_{1-94}I69M mutant (SeMetF_{1-94}I69M)

The hanging drops of selenomethionyl F_{1-94}I69M mutant protein were setup in VDX plates (Hampton Research) and the crystals could be successfully reproduced in the condition of 30 % (v/v) polyethylene glycol 4000 (PEG 4000), 0.1 M $MgCl_2$, 6xH_2O, 0.1 M Tris/HCl (pH 8.8), 25 % (v/v) glycerol (109). In the cover slip 2 µl of the protein solution was mixed with 2 µl of the mother solution and the crystal drops were equilibrated at 18 ºC. The protein produced small crystals under the same condition. Therefore, the crystallization conditions were optimized. Good diffraction quality crystals appeared within a day in 30% PEG 4000, 0.05 M $MgCl_2$, 0.1 M Tris-HCl pH 8.8, 10% glycerol with a 5 mg/ml protein concentration which were tested in the in-house Rigaku machine and diffracted maximum to 3.5 Å.

2.2.11 Data collection and processing

The native crystals were quick-soaked in a cryoprotectant solution containing 25% glycerol in mother liquor and flash-cooled in liquid nitrogen. The SeMetF_{1-94}I69M mutant crystals were transferred to stabilizing solution (32% PEG 4000, 0.05 M $MgCl_2$, 0.1 M Tris-HCl pH 8.8, 10% glycerol) and the crystals were cryoprotected in the same stabilizing solution containing 22% glycerol and flash-cooled in liquid nitrogen (109). A single-wavelength data set was collected from a native F_{1-94} crystal in-house at 100 K on a Rigaku R-AXIS IV image-plate detector with a Rigaku RA-Micro 7 HFM rotating copper-anode

generator (Rigaku/MSC). 360 frames were collected with an oscillation range of 1^{o} and an exposure time of 5 min per frame. For the SeMetF_{1-94}I69M mutant protein crystal, a three-wavelength multiwavelength anomalous diffraction (MAD) data set was collected at 100 K on beamline 13B1 at the National Synchrotron Radiation Research Center (NSRRC), Hsinchu, Taiwan using an ADSC Quantum 315 CCD detector. A single crystal was used to collect data at appropriate inflection, peak and high remote wavelengths based on the selenium absorption spectrum (109). Data were collected as a series of 0.5^{o} oscillation images with 5 s exposure time, covering crystal rotation ranges of 360^{o} for the peak wavelength and 180^{o} for the inflection and high remote wavelengths. The inflection and high remote wavelength data sets were initially collected, followed by the peak wavelength data set. All diffraction data were indexed, integrated and scaled using the HKL-2000 suite of programs (110). All the crystals belong to orthorhombic space group $C222_1$ and have similar unit cell parameters.

2.2.12 Structure determination of F_{1-94}

The crystals of recombinant mutant protein SeMetF_{1-94}I69M of *S. cerevisiae* V-ATPase (*VMA7*p) were obtained by hanging drop method and diffracted to 2.3 Å. The crystal belongs to the orthorhombic $C222_1$ space group (unit cell parameters a = 47.21 Å, b = 160.26 Å and c = 102.49 Å) with 4 molecules in the asymmetric unit. Phasing was carried out by three wavelengths MAD technique. Analysis with SHELXC (111) indicated a maximum resolution of 2.9 Å for substructure determination and initial phase calculation. All the four Se-sites were found using SHELXD program (112) with a correlation coefficient of 43.62%. The correct hand for the substructure and initial phases were calculated after density modification using the program SHELXE (113). The phases were improved by density modification and phase extension to 2.44 Å resolution using the program RESOLVE (114). Automated model building was carried out by Buccaneer program (115), which gave the model of the SeMetF_{1-94}I69M mutant protein with 312 residues for four molecules. The phases were extended up to 2.3 Å and iterative cycles of model building and refinement were carried out using programs COOT (116) and REFMAC5 (117) of the CCP4 suite to built the final model. Refinement was done until convergence and the geometry of the final model was validated with PROCHECK (118). The final figures were drawn using the program PyMOL (119) and structural comparison analysis was carried out by using the SUPERPOSE program (120) as included in the CCP4 suite.

2.2.13 Nuclear magnetic resonance (NMR) spectroscopy

NMR spectroscopy is one of the principal techniques used to study the structural and functional aspects of macromolecules such as proteins. The technique enables to determine the solution structure of proteins and to study protein-ligand and protein-protein interactions. As hydrogen is one of the most receptive and abundant NMR active nuclei besides other isotopes such as ^{15}N, ^{13}C, ^{31}P *etc.*, it can be observed in magnetic resonance. Any nuclei having mass number, which exhibits net intrinsic magnetic momentum and angular momentum, can be defined as NMR active. All the NMR experiments were collected on Bruker Avance 600/700 MHz machines at SBS NMR core facility.

2.2.13.1 ^{15}N single and ^{13}C ^{15}N double labeling of proteins

Reagents for Minimal Media (M9):

42 mM Na_2NPO_4
22 mM KH_2PO_4
8.5 mM NaCl
1 g/l $^{15}NH_4Cl$
0.1 mM $CaCl_2$
2 mM $MgSO_4$
10 g D-Glucose
30 µM $FeCl_3$
5 ng/l Thiamin

For NMR experiments, either single (^{15}N) or double (^{13}C ^{15}N) labelled recombinant proteins were produced in *E. coli* BL21 (DE3) cells and purified according to the established protocols mentioned in section 2.2.1.3. The clones were grown in M9 minimal media to get labelled protein. 30 ml of overnight culture was pelleted down at low centrifugal force at room temperature and washed with minimal media and then transferred to large volume of minimal media with OD_{600} of 0.1. The cultures were grown in incubator shaker at 37 ºC till OD_{600} of 0.6 - 0.7 was achieved. The protein production was induced with IPTG to final concentration of 1 mM overnight at 20 ºC. Cells were harvested and frozen in liquid nitrogen and stored at -80 ºC till purification.

2.2.13.2 One dimensional (1D) ^{1}H and multi-dimensional (2D, 3D) ^{13}C-^{15}N NMR spectroscopy

1D NMR spectrum of various proteins was collected at temperatures ranging from 283 K to 308 K on core facility Avance 600 MHz instruments (Bruker, Billerica, MA). Unlabelled or labelled protein sample in phosphate buffer, pH 6.8 was used in presence of 10 % D_2O (v/v) to record the spectrum. For 2D experiments like HSQC (Hetero nuclear Single

Quantum Coherence), ^{15}N labelled proteins in 90% H_2O and 10% D_2O in 25 mM phosphate buffer (pH 6.8) was used to collect data. Different parameters including temperature, buffer, and protein concentration were optimized before making final measurements. Pulse was also calibrated accordingly. ^{15}N-^{13}C labelled samples were utilized to collect 3D spectra such as CBCA(CO)NH, HNCACB in 10% (v/v) D_2O lock signal. Baseline corrections were applied wherever necessary. The proton chemical shift was referenced to the methyl signal of DSS (2, 2-dimethyl-2-silapentane-5-sulphonate (Cambridge Isotope Laboratories)) as an external reference to 0 ppm.

2.2.13.3 NMR data collection and processing

All NMR experiments were performed at 20 °C on 700 MHz NMR spectrometer (Bruker Avance) equipped with 5 mm triple-resonance ($^1H/^{15}N/^{13}C$) single axis gradient probes and a cryoprobe. 0.5 mM of F subunit was prepared in 50 mM Tris at pH 7.5, 200 mM NaCl, 5 mM EDTA and 5 % D_2O. D_2O was added for the locking of the sample. Appropriate pulse calibrations and other parameters including temperature, buffer and protein concentration were optimized before making final measurements. To assign the backbone of F subunit, TROSY-based experiments such as 2D ^1H-^{15}N TROSY-HSQC, 3D-TROSY-HNCO, 3D-TROSY-CBCA(CO)NH and 3D-TROSY-HNCACB were performed with improved signals (~ 10-50 % increase) unlike non-TROSY experiments, showing signal broadening for many residues. In all the above experiments, spectral widths were generally 9.6 kHz for protons, 2.8 kHz for ^{15}N, while for ^{13}C dimensions they were 6-12 kHz. All the three-dimensional experiments made use of pulsed-field gradients for coherence selection and artifact suppression, and utilized gradient sensitivity enhancement schemes. Quadrature detection in the indirectly detected dimensions was achieved using either the States/TPPI (time-proportional phase incrementation) or the echo/anti-echo method. Baseline corrections were applied wherever necessary. The proton chemical shift was referenced to the methyl signal of 2, 2-dimethyl-2-silapentane-5-sulphonate (Cambridge Isotope Laboratories) as an external reference to 0 ppm. The ^{13}C and ^{15}N chemical shifts were referenced indirectly to DSS. All the NMR spectra were processed either using nmrPipe/nmrDraw (121) or Bruker Avance spectrometer in-built software Topspin. Peak-picking and data analysis of the Fourier transformed spectra were performed with the SPARKY program (104).

2.2.13.4 Resonance assignment

In the initial stage, protein molecules must be assigned in order to investigate nuclear magnetic resonance data for analysis of any kind of interaction or dynamic studies. Scalar couplings, via dipolar coupling and via chemical shift are useful information for sequence specific assignment of protein backbone. Depending on the type of sample whether it is unlabeled proteins or isotopically labeled proteins (122), different strategies are employed to assign the resonance. In NMR, a one-dimensional (1D) spectrum is useful to investigate dispersion of resonance line of protein and also it gives an idea about the folding of the protein. But by doing one-dimensional (1D) experiment it is impossible to resolve all resonance lines in a spectrum for protein. Assignments can be done using the sequential assignment approach based on homonuclear two-dimensional (2D) correlation spectroscopy (COSY) or total correlation spectroscopy (TOCSY) (123) and 2D nuclear Overhauser effect spectroscopy (NOESY). The TOCSY experiment is basically used to identify the resonance position within each amino acid spin system and the NOESY experiment is used for the sequential assignment of the spin systems present in the amino acids. For a protein/peptide NMR spectrum, amide hydrogens resonances are generally the best dispersed and observed between 6-12 ppm. The side chain resonance assignments are generally best done by examining the peaks in the backbone amide regions. This process is known as the sequential assignment strategy developed by Wüthrich *et al* (124).

For larger molecules (>10 kDa) the sequential assignment strategy by using above assignment approach is not useful, because overlaps of the NMR signals make the assignment complicated. For that it requires the addition of a third dimension by the replacement of naturally abundant isotope ^{14}N with ^{15}N and ^{12}C by ^{13}C. Hence, assignment is based on triple resonance experiments allowing further decrease in the overlap of NMR signals. There are six triple resonance experiments commonly used in chemical shift assignments of the backbone atoms of the protein: HNCA, HN(CO)CA, HNCACB, CBCA(CO)HN, HNCO and HN(CA)CO. In HNCO experiment, magnetization is transferred from HN(i) proton via the N(i) atom to the directly attached CO(i-1) carbon atom and return the same way to the HN(i) nucleus which is directly detected. This experiment gives the chemical shifts of ^{13}CO of the previous amino acid. The details of transfer of magnetization as well as an example of assignment of the same four residues of subunit F is shown in figure 2.4 in case of HNCO experiment.

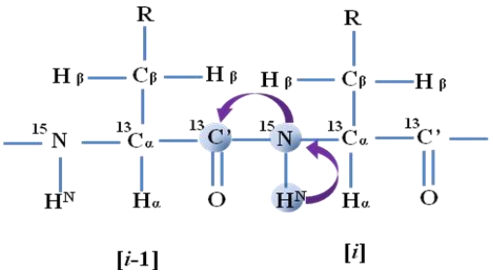


Figure 2.4: Transfer of magnetization in the HNCO experiment. The HNCO experiment correlates the amide $^{1}H^{N}$ and ^{15}N chemical shifts of one amino acid with the ^{13}CO chemical shift of the preceding residue, by using the one-bond ^{15}N–^{13}CO J coupling (15 Hz) to establish the sequential correlation (122).

HNCACB experiment correlates the $^{13}C\alpha$ and $^{13}C\beta$ resonances with the amide ^{1}H and ^{15}N resonances of the same and previous residue whereas the CBCA(CO)NH experiment correlate the $^{13}C\alpha$ and $^{13}C\beta$ resonances with the amide ^{1}H and ^{15}N resonances of previous residue (122). These two experiments take into account both $^{13}C^{\alpha}$ as well as $^{13}C^{\beta}$ values for the sequential assignments of residues. They are extremely useful in case of degeneracy of $^{13}C^{\alpha}$ values. In addition, $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ values can provide information about the type of amino acid. The details of transfer of magnetization as well as an example of assignment of the same four residues of subunit F is shown in figure 2.5 in the case of HNCACB and CBCA(CO)NH experiments.

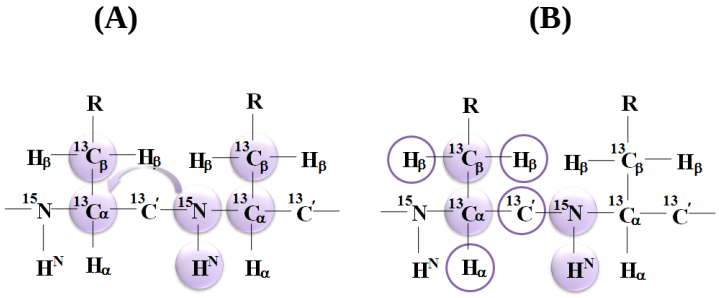


Figure 2.5: Transfer of magnetization in the HNCACB *(A)* and CBCA(CO)NH *(B)*.

2.2.13.5 Backbone dynamics

The backbone ^{1}H-^{15}N steady state heteronuclear NOE provides information about the motion of individual N-H bond vectors. Those motions are faster than the overall tumbling of the molecules (i.e. in the pico- to nanosecond time scale) and show a decreased NOE intensity relative to the average observed for the majority of the residues. This heteronuclear steady-state ^{15}N-^{1}H NOE values were obtained from the ratios of peak intensities in the saturated spectrum to those in the unsaturated spectrum. In order to suppress time- or temperature-dependent effects, the spectra were acquired in an interleaved mode with incremented

relaxation delays. All experiments were recorded with 256 t_1 increments of 2,048 data points in t_2. The spectral widths were set to 9,586.15 Hz (^{1}H), 2,406 Hz (^{15}N) at 700 MHz. Errors were estimated by evaluating the standard deviation of the NOE, σ_{NOE}.

$$\sigma_{NOE}/NOE = (\sigma I_{sat}/I_{sat})^2 + (\sigma I_{unsat}/I_{unsat})^2)^{1/2}$$

where σI_{sat} and σI_{unsat} are the standard deviation of the noise in the spectra.

2.2.14 Cloning and production of subunit H, H_{8-353}, $H_{354-478}$ and subunit C from *S.cerevisiae* V-ATPase

The genes encoding subunit H**,** H_{8-354} and $H_{354-487}$ of *S.cerevisiae* V-ATPase were amplified using *S.cerevisiae* genomic DNA as a template for PCR amplification. The amplified products were digested with *NcoI* and *SacI* restriction enzymes subsequently ligated with pET9d1-His_6 vector (86). The vector containing the individual gene was transformed into *E. coli* cells (BL21 (DE3)) and grown on 30 μg/ml kanamycin containing LB agar-plates. Recombinant protein was expressed by IPTG supplemention to a final concentration of 1 mM. The induction of individual protein was confirmed by SDS-PAGE gel. Cloning of subunit C was done as described previously (125). The names of subunits and corresponding primers for cloning are summarized below.

Constructs		**Primer sequence (5'-3')**
Subunit H	Forward Primer	GATCCATGGGCGCAACCAAAATTTTAAT
	Reverse Primer	ATGCGAGCTCTTATTTGAAGGTATATCCAATGATT
Subunit H_{8-353}	Forward Primer	GATCCATGGACAGTACTCATTTCAATGAGA
	Reverse Primer	CATCGAGCTCTTACAATTCTTGGTACTCG
Subunit $H_{354-478}$	Forward Primer	ACCCCATGGCAACCTCCTTCGAT
	Reverse Primer	ATGCGAGCTCTTATTTGAAGGTATATCCAATGAT
Subunit C	Forward Primer	CATGCCATGGCTACTGCGTTATATA
	Reverse Primer	CGGGATCCTTATAAATTGATTATATACAT

2.2.14.1 Purification of subunit H, $H_{354-478}$, C and *d* from *S. cerevisiae* V-ATPase

The cells containing His_6-tagged subunit H and $H_{354-487}$ were sonicated separately. The supernatant was collected after centrifugation which was then incubated with Ni^{2+}-NTA beads. Bound protein was eluted by an imidazole gradient (0-400 mM). Fractions between the

imidazole concentrations of 75 mM to 200 mM were pooled together and passed through the size exclusion column (Superdex75™ 10/300 GL), which was equilibrated with buffer 50 mM Tris/HCl, pH 7.5, 150 mM NaCl and 1 mM DTT. The fractions containing the respective proteins were pooled together and concentrated using Centriprep YM-3 (3 kDa molecular mass cut-off) spin concentrator and 1 µl of the respective concentrated protein sample was loaded onto the SDS-PAGE gel to check the purity of the protein. Subunit C and *d* were purified as described previously (125,126).

2.2.15 Binding studies with NMR

NMR spectroscopy is extensively used to identify the important residues for protein protein interaction or protein ligand interaction (Figure 2.6) (127,128). ^{1}H-^{15}N HSQC spectrum of single labeled protein was used as starting point at final optimized condition. The ligand protein/peptide was then added in increasing amount at the same temperature and buffer conditions. A series of ^{1}H-^{15}N HSQC spectra were recorded at each concentration of the ligand protein/peptide. 100 - 500 µM of ^{15}N labeled protein without ligand was used for the reference spectrum. Experiments were performed on Bruker Avance 700 machine using Topspin software (Bruker BioSpin) for acquisition and processing of spectra. Spectra were overlapped to monitor chemical shift changes, further analysis were done in SPARKY (129).

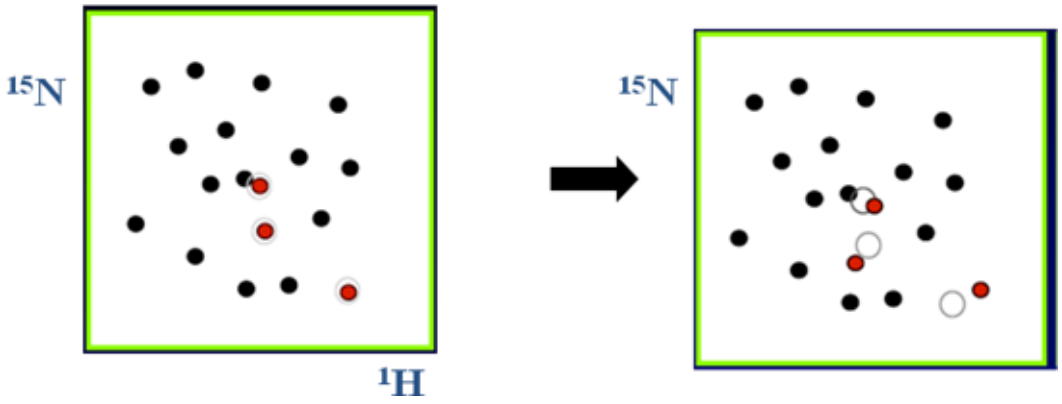


Figure 2.6: Illustration of peak position change (red color) in ^{1}H-^{15}N HSQC spectrum upon addition of binding partner (122).

2.2.16 Cloning of the subunit DF-heterodimer of *S. cerevisiae* V-ATPase

The genes encoding subunit F (*VMA7*) and subunit D (*VMA8*) of *S. cerevisiae* V-ATPase were amplified by using genomic DNA as a template. In PCR reactions forward primer 5′-AACATATGCATCACCATCATCACCACATGTCTGGTAATAGA-3′ and reverse primer 5′-CCTCTCGAGTCAGAATATAACATCGTCTTCTTGATCAGCAA-3′

incorporating *NdeI* and *XhoI* (underlined) restriction sites, were used to amplify subunit D gene whereas for subunit F gene amplification forward primer 5′-AAGGATCCCATGGCTGAGAAACGTACTCTT-3′ incorporating *BamHI* and reverse primer 5′-GTCCTGAGCTCTTACTCACCGAACAACTTTC-3′ incorporating *SacI* were used. Six additional histidine codons were incorporated in the forward primer for *VMA8*. A total volume of 50 µl reaction mixture was prepared to setup the polymerase chain reaction and the appropriate concentrations of constituents are mentioned below:

Reagents	**Amount**
Pfu buffer (10x)	5 µl
dNTP's (2 mM)	1.5 µl
Primers (100 µM)	2 x 0.5 µl
Template (genomic DNA)	1 µl
MilliQ water	40.5 µl
Pfu DNA polymerase	1 µl

All reagents were accurately measured and mixed properly inside a PCR tube, which has to be kept on ice all the time. Before, reaction tubes were placed inside the machine PCR thermocycler (Biometra T personal) and the lid was preheated to 99 ºC, followed by amplification by using appropriate program. DNA denaturation, annealing and extension were repeated for another 29 cycles. Total time taken for the PCR program was variable for each gene. The annealing temperature for *VMA7* and *VMA8* genes was 60 ºC and 54 ºC, respectively.

Cycle steps	**Temperature**	**Time**	
Lid	99 ºC		
Initial denaturation	96 ºC	3 min	
Cycle denaturation	96 ºC	30 sec	
Annealing	Variable	45 sec	30 cycles
Extension	72 ºC	60 sec	
Final extension	72 ºC	5 min	
End	4 ºC		

After amplification, PCR products were stored at -20 °C. 5 µl of PCR amplified products were applied onto an analytical agarose gel (1%) to analyze the quality. To identify the correct PCR products, the gel was soaked in ethidium bromide solution for 5 min followed by washing with water and observed under UV light. A DNA marker (100 bp) was used to confirm the right products. Preparative agarose gel was made to purify PCR products and the portions corresponding to the correct PCR products were cut out immediately and purified by gel extraction kit (QIAGEN) described in section 2.2.5. Water or EB buffer (Tris-HCl buffer of pH 8.0) was used to elute products. 1 µl of gel extracted PCR products was applied onto an 1% analytical agarose gel to estimate the purity of samples. Both purified DNA (*VMA7* and *VMA8*) were then single digested. The *VMA7* gene was digested by *BamHI* and *SacI* enzymes overnight at 37 ºC (SW22 Shaking Water Bath, Julabo, Germany) and the *VMA8* gene was digested by *NdeI* and *XhoI* enzymes. After overnight digestion, the reaction mixtures were again purified by enzyme reaction purification kit from QIAGEN. The vector was also single digested with respective restriction enzymes for one gene and then treated with alkaline phosphatase at 37 ºC (SW22 Shaking Water Bath, Julabo, Germany) for 30 min. Ligation reaction between pETDuet-1 vector (Figure 2.7) and one insert (amplified PCR product, *VMA7*) was setup at room temperature. The reagents used were as follows:

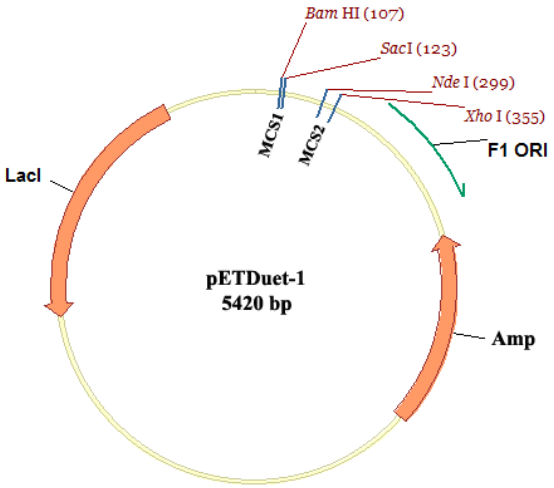


Figure 2.7: pETDuet-1 vector (Novagen) showing full map and two multiple cloning sites (MCS1 and MCS2), has been used to clone *VMA7* and *VMA8* genes.

Ligase buffer (10x)	2 µl
Vector (V)	50 ng
Insert (I)	Variable (1-5 times of vector)
T4 Ligase	1 µl
MilliQ water	variable
Total	20 µl

$MgCl_2$, which is present in ligase buffer, has a tendency to precipitate during the storage. Before setting up the reaction, ligase buffer has to be vortexed to completely dissolve

the precipitate. The ligation mixture was kept at room temperature for 1 h, followed by the purification of the ligation product by the butanol precipitate method as mentioned in section 2.2.5. The ligated product was transformed into *E. coli* DH5α cells. Cells were grown at 37 ºC for 1 h and plated on LB-agar containing carbenicillin as a marker. On the following day, five colonies were picked from the transformation plate, containing carbenicillin antibiotic marker, and grown at 37 °C overnight. The plasmid DNA was isolated using a standard protocol (QIAGEN mini-prep kit). To confirm incorporation of expected gene, purified plasmid DNA was double digested with restriction enzymes for 2.5 hrs and applied onto 1% agarose gel. Size of insert and vector were compared with the DNA markers (1 Kb). After confirming the presence of insert, the purified construct was sent for sequencing. Verified plasmid was then transformed into *E. coli* BL21 DH5α cells for second gene (*VMA8*) ligation. The same process has been carried out, in order to get vector with both genes inside and sent again for sequencing. The *VMA7* and *VMA8* genes have been cloned inside MCS1 and MCS2, respectively. Verified plasmid was finally transformed into *E. coli* BL21 (DE3) cells for protein production and they were grown on 100 μg/ml carbenicillin containing Luria-Bertoni (LB) agar plates.

2.2.17 Production and purification of the DF-heterodimer of *S. cerevisiae* V_1V_O ATPase

Induction and solubility tests (Figure 2.8A-B) were done in a similar way as described for subunit F, mentioned in section 2.2.1. After successful solubility tests, cells containing DF-heterodimer were scaled up for further purification. Cells were grown in LB medium containing carbenicillin (100 μg/ml) and shaken at 37 °C until an optical density OD_{600} of 0.6 was reached. IPTG was added with final concentration of 1 mM, in order to induce the expression of DF-heterodimer and grown for another 4 hrs. Cells were harvested at 8 500 x *g* for 12 min, 6 °C followed by purification, and opened by sonication in buffer A (50 mM TrisHCl, pH 8.5, 200 mM NaCl, 2 mM PMSF and 2 mM Pefabloc^{SC} (BIOMOL)). After sonication cell debris was pelleted down by centrifugation at 10 000 x *g* for 35 min. The supernatant was separated from pellet and filtered through 0.45 μm Millipore filter paper. The His-tagged protein complex was mixed with Ni^{2+}-NTA resin, which was equilibrated with 50 mM TrisHCl, pH 8.5, 200 mM NaCl. The mixture was allowed to bind to the matrix for 1.5 h at 4 °C. After which the lysate-resin mixture was poured in column and the flow through was collected. Bound protein was eluted with an imidazole-gradient (25 - 200 mM) in buffer A. Fractions containing His-tagged DF protein were identified by SDS-PAGE (88). Required fractions of protein were pooled together and diluted with 50 mM TrisHCl, pH 8.5 to make a

final NaCl concentration of 50 mM. The sample was applied onto anion exchange column (Resouce Q, 6 ml) and subsequently protein fractions were eluted using NaCl salt gradient. The DF protein fractions were confirmed by running SDS-PAGE gel and pooled together followed by concentration using Centricon YM-3 (10 kDa molecular mass cut off) spin concentrators (Millipore). The protein complex was injected on to a gel filtration Superdex S200 column (GE Healthcare), where 50 mM TrisHCl (pH 8.5), 200 mM NaCl and 5 mM EDTA was used as a mobile phase. The purity and homogeneity of all protein samples were analyzed by SDS-PAGE (88). SDS-gels were stained with Coomassie Brilliant Blue G250 and bicinchoninic acid assay (BCA; Pierce, Rockford, IL., USA) was performed to determine protein concentrations.

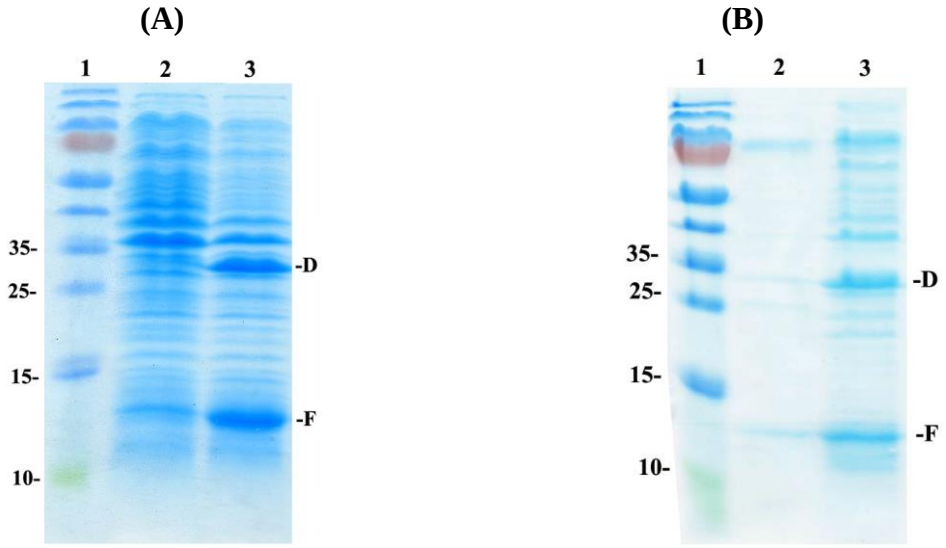


Figure 2.8: Induction and solubility test of DF-heterodimer. *(A)* Induction test of DF-heterodimer. Uninduced and induced proten are shown in *lane 2 and 3*, respectively. *Lane 1* correcponds to protein ladder. *(B)* Representative picture of solubility test of DF-heterodimer. *Lane 1*, *2* and *3* are corresponding to protein ladder, protein in pellet and protein in supernatant, respectively.

2.2.18 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) was used to study the macromolecular interactions in the solution. It permits determination of binding affinities and also the binding constant of protein-protein interaction (130). Thermodynamic parameters such as entropic and enthalpic components of such binding reactions can be resolved using ITC. Subunit interactions of the V-ATPase were monitored by isothermal titration calorimetry. The proteins were equilibrated in the same buffer (50 mM Tris-HCl, pH 7.5, 200 mM NaCl, 5 mM EDTA, 1 mM TCEP), filtered and degassed before titration. The protein was then loaded in the syringe and titrated in 1.5 µl injections against variable concentration of protein placed in the

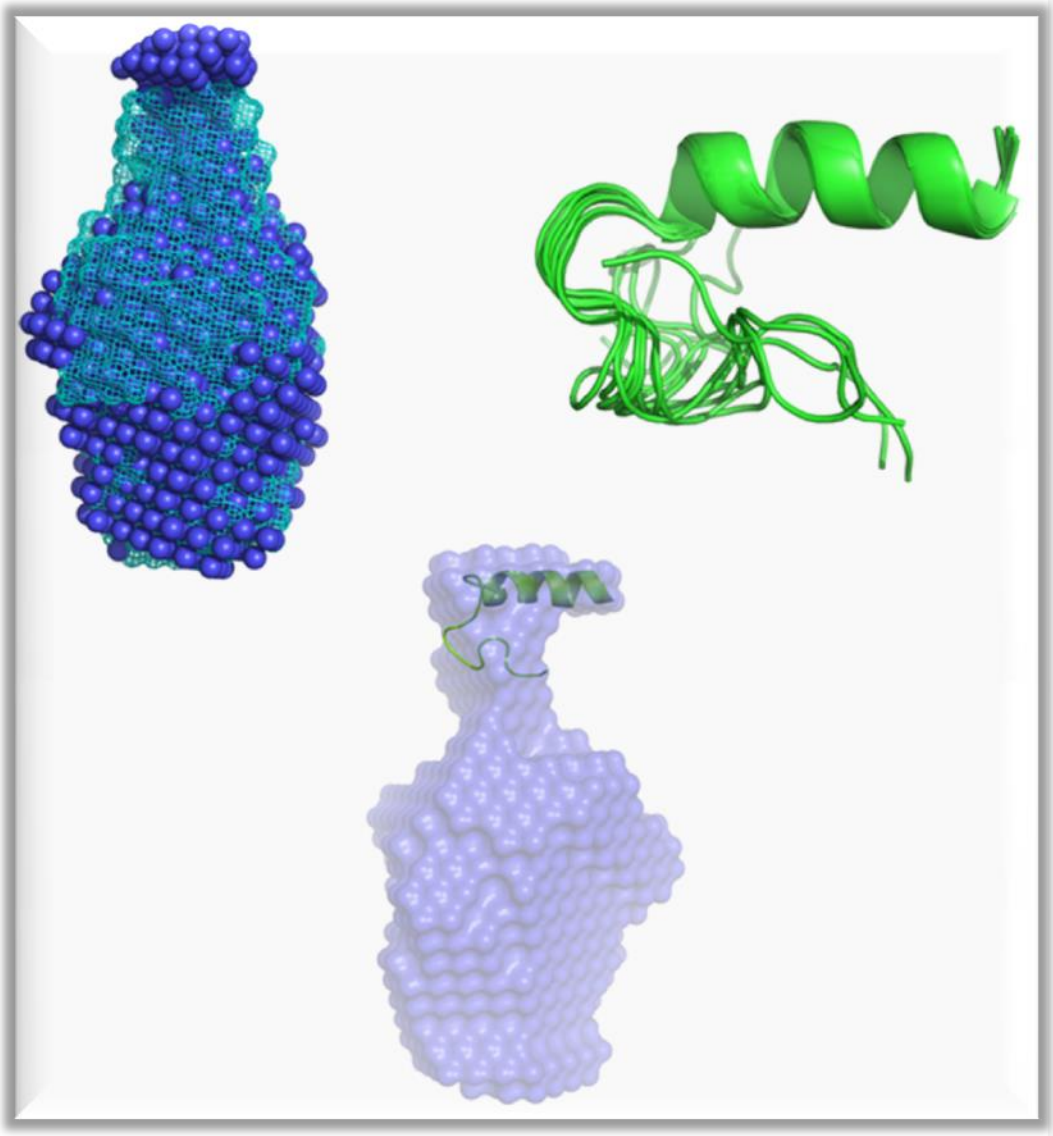
sample cell in a MicroCal iTC200 microcalorimeter (Microcal, Northampton, UK). All experiments were performed under identical conditions at 20 ºC. The binding curve was determined by least squares method and the binding isotherm was fitted using Origin v7.0 (Microcal) assuming a single-site binding model.

2.2.19 Crystallization of the subunit DF-heterodimer

Crystallization of the DF-heterodimer was attempted using vapour diffusion method. Sparse matrix screens from Hampton research and Emerald biosystems were used for the initial screening. Hanging drops were set up by mixing 1 µl of the purified DF protein (5 mg/ml and 8 mg/ml) in buffer B (50 mM TrisHCl, pH 8.5, 250 mM NaCl, 5mM EDTA) with 1 µl of the precipitant solution and incubated at 18 °C (108). Initial crystals were seen in Hampton research crystal screen 1 condition 11 (0.1 M Sodium citrate tribasic dihydrate pH 5.6, 1.0 M Ammonium phosphate monobasic). Once these crystals were confirmed to be protein crystals, further optimization of the crystallization condition was done in order to increase the size of crystals. Cryoprotectant solution was also optimized by testing different percentages of glycerol and final cryoprotectant solution comprised of 0.1 M sodium citrate tribasic dehydrate, pH 5.6, 1.1 M ammonium phosphate monobasic, 25 % (v/v) glycerol. The crystals were quickly dipped in the cryoprotectant solution and were flash-frozen in liquid nitrogen at 100 K. Optimized crystals were tested in the in-house Rigaku machine and subsequently analyzed at National Synchrotron Radiation Research Center (NSRRC), Taiwan.

3. Results

3.1 Structural characterization of subunit F domains of the *S. cerevisiae* V_1V_O ATPase in solution by SAXS and NMR spectroscopy



3.1.1 Introduction

Eukaryotic V-ATPases are multi-subunit enzyme complexes which pump proton in an ATP-dependent manner and play crucial physiological roles in various intracellular compartments, which lead to several cellular events such as endocytosis and intracellular transport (131). V_1V_O-ATPases consist of a hydrophilic V_1 domain, which is water-soluble and responsible for ATP hydrolysis, and a membrane bound V_O domain, which is responsible for translocation of proton. The stalk part consists of the V_1 subunits C-H and subunit *a* and *d* of V_O, and has an important role in energy coupling. The stalk subunits form the structural and functional interface between the two domains. Two important subunits F and D, which belong to central stalk, and are directly involved in ATP hydrolysis in the A_3B_3 hexameric headpiece of the V_1 domain (41). A secondary structure prediction (Figure 3.1.1) shows that subunit F is composed of an alternative arrangement of α-helices and β-strands.

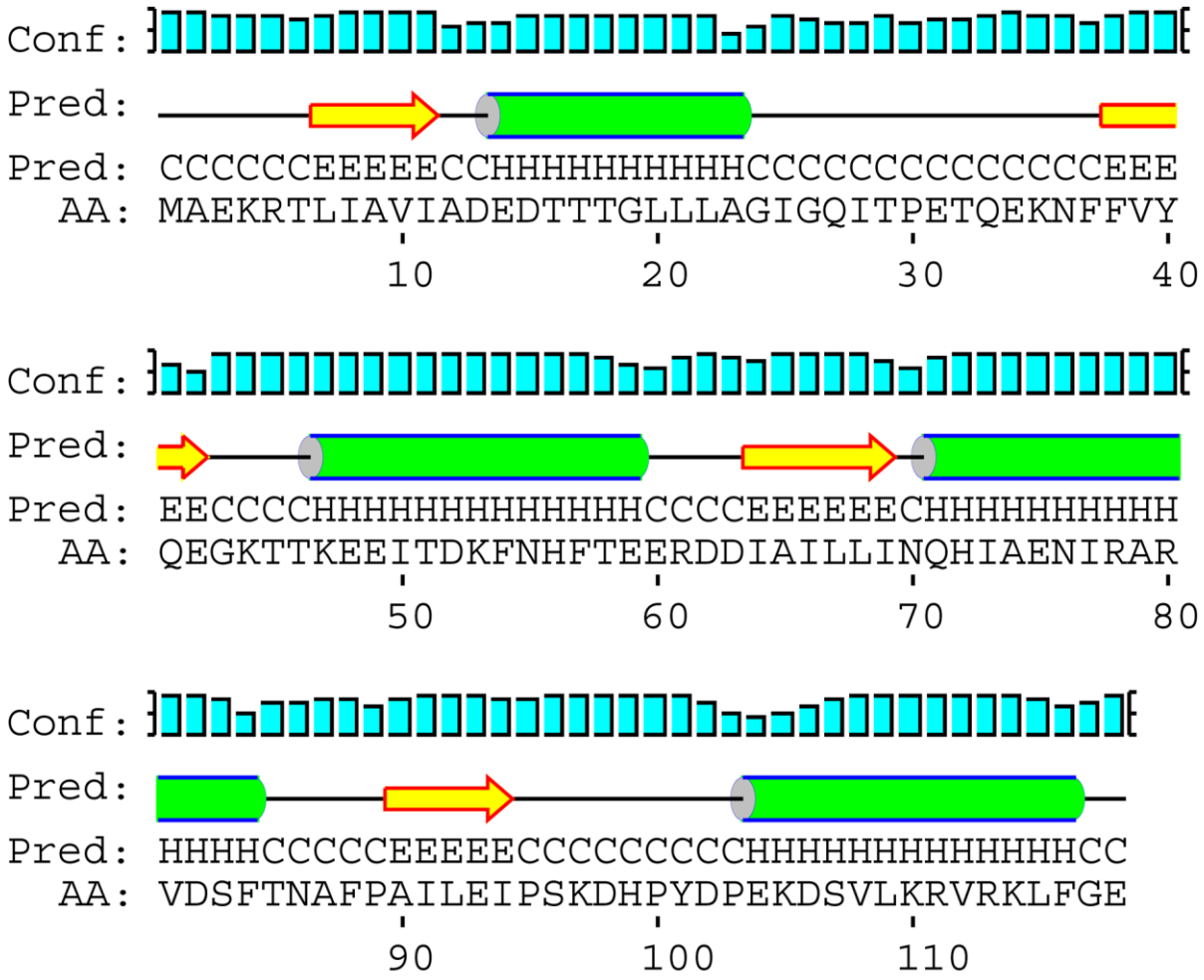


Figure 3.1.1: Secondary structure prediction of subunit F by psipred, where α-helix= , β-strand= , coil= and confidence of prediction=.

A recent peptide array analysis revealed that several peptides of subunit F can interact with other stalk subunits, indicating complicated arrangements of stalk subunits in the entire complex (132). Due to the central role of subunit F, structural insight of this protein is very important. In this current study, complementary techniques such as SAXS and NMR spectroscopy were utilized to solve and characterize the low, as well as high resolution structures of subunit F and its C-terminal peptide F_{90-116}, respectively. The domain features of subunit F and F_{1-94} are characterized by SAXS experiment. Superimposition of these two structures was enabled to identify the both termini of the subunit F.

3.1.2 Production and purification of subunit F

The 14 kDa His-tagged protein, subunit F was overexpressed using conditions mentioned in section 2.2.1.1. The protein was soluble in the respective buffer mentioned in section 2.2.1.2, which represent subunit F of the *S. cerevisiae* V_1V_O ATPase and was confirmed by applying the sample onto the SDS-PAGE gel. In the first step of purification, affinity column chromatography technique was carried out, where Ni^{2+}-NTA matrix was used to bind the protein. Subsequently the bound protein was eluted by a gradient of imidazole (0-400 mM) which was prepared in a buffer containing 50 mM HEPES, pH 7.0, 300 mM NaCl, 2 mM PMSF and 2 mM Pefabloc^{SC}. Each fraction was collected in a separate tube containing EDTA and then applied onto the SDS-PAGE gel to check for the presence of desired protein. Protein eluting between 75 to 200 mM concentration of imidazole was pooled together, concentrated and then loaded onto a gel filtration column (Superdex^{TM} 75 10/300 GL) (Figure 3.1.2), to obtain pure and monodispersed protein. Purity of the protein was ascertained by applying the sample onto an SDS-PAGE gel. The gel was stained with

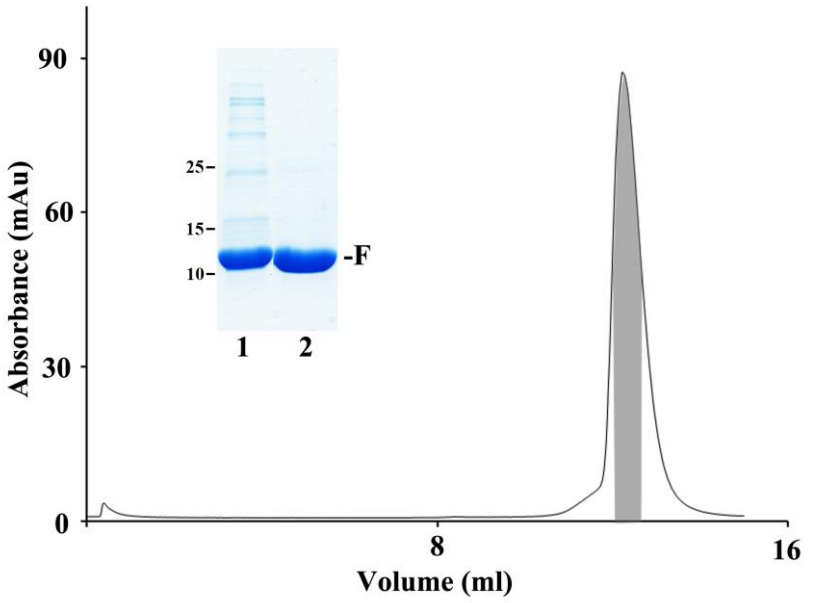


Figure 3.1.2: Purification of *S. cerevisiae* subunit F (*VMA7*p). After Ni^{2+}-NTA purification the protein was loaded onto Superdex^{TM} 75 10/300 GL column. Insert in the figure shows the SDS gel with two lanes where *lane 1* represent the protein after Ni^{2+}-NTA purification and *lane 2* represent the eluted protein after Superdex^{TM} 75 purification (87).

Coomassie Brilliant Blue G250 dye to visualize protein (Figure 3.1.2). BCA and BioSpec-nano (Shimadzu Scientific Instruments) were used to determine protein concentration (87). Matrix-assisted laser desorption/ionization (MALDI) mass spectrometry was performed and showed, that the subunit F has a total mass of 13,685.82 Da, thus confirming mass calculated from the sequence of the protein.

3.1.3 Estimation of secondary structure content in subunit F

Circular dichroism spectroscopy was carried out to estimate the percentage of secondary structure content of subunit F. The spectrum was obtained between 190-260 nm wavelength and showed one maxima at 192 nm and two minima at 208 and 222 nm (Figure 3.1.3). This pattern is typical for an α-helical protein. A minimum at 215 nm is expected for proteins with β-sheet. Due to the overlapping α-helical region, the minimum of β-sheet is not clear. The percentage of average secondary structure content was determined to be 43%, 32% β-sheet and 25% random coil, which is in line with the predicted secondary structure calculated from the amino-acid sequence of subunit F. The molar ellipticity ratio at 208 nm and at 222 nm is calculated to be 0.98 (87).

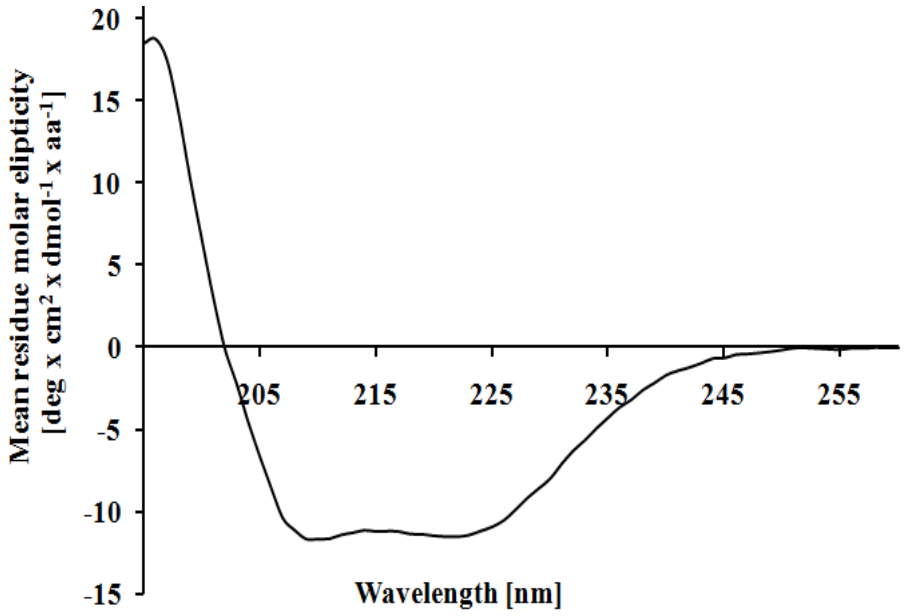


Figure 3.1.3: CD spectrum of subunit F showing far UV-CD region (87).

3.1.4 Solution shape of *S. cerevisiae* subunit F determined by SAXS

The first aim of this project was to analyze domain features of subunit F from eukaryotic V-ATPases. Henceforth, SAXS experiments were conducted to obtain a scattering pattern of subunit F in solution. The final experimental composite scattering curve is shown in figure 3.1.4A. Further Guinier plots analysis revealed the good experimental data quality without

any protein aggregation in the solution in the concentration range of 2.5-7.4 mg/ml. Subunit F displayed a radius of gyration (R_g) of 2.28 ± 0.3 nm and a maximum dimension (D_{max}) of 7.7 ± 0.3 nm (Figure 3.1.4B). Forward scattering of reference protein bovine serum albumin (BSA: 66.4 ± 2 kDa) was also obtained to compare the scattering of subunit F to calculate the molecular weight of sample protein. The calculated molecular mass of subunit F is 14 ± 2 kDa, which is in agreement with the result from gel filtration chromatography mentioned in chapter 3.1.2 and implied that subunit F is a monomer in solution at the 2.5-7.4 mg/ml concentration range. Distance distribution function analysis suggested the existence of one major domain in subunit F, producing a principal maximum ($p(r)$) of around 2.8 nm (Figure 3.1.4B) (87). It is also observed that a separate protuberant domain present in the structure produced a shoulder from 5.8 nm to 7.7 nm. The gross structural shape of subunit F was calculated from the scattering data as shown in figure 3.1.4A. The calculated and experimental scattering curves of the protein are fitting nicely in the entire scattering range (87). The respective fit is shown in figure 3.1.4A which has a discrepancy of χ^2 = 1.349. In total, ten independent models were reconstructed by the program DAMMIN (100) and each generated a reproducible shape. Finally all shapes were averaged and are shown in figure 3.1.5. Subunit F contains two domains; a large egg-like major domain, connected to a small hook-like domain via a 1.5 nm long linker part. The dimensions of major egg-like domain is about 5.5 x 3.3 nm, whereas the hook-like region is around 0.89 nm in length and 2.23 nm in width (87).

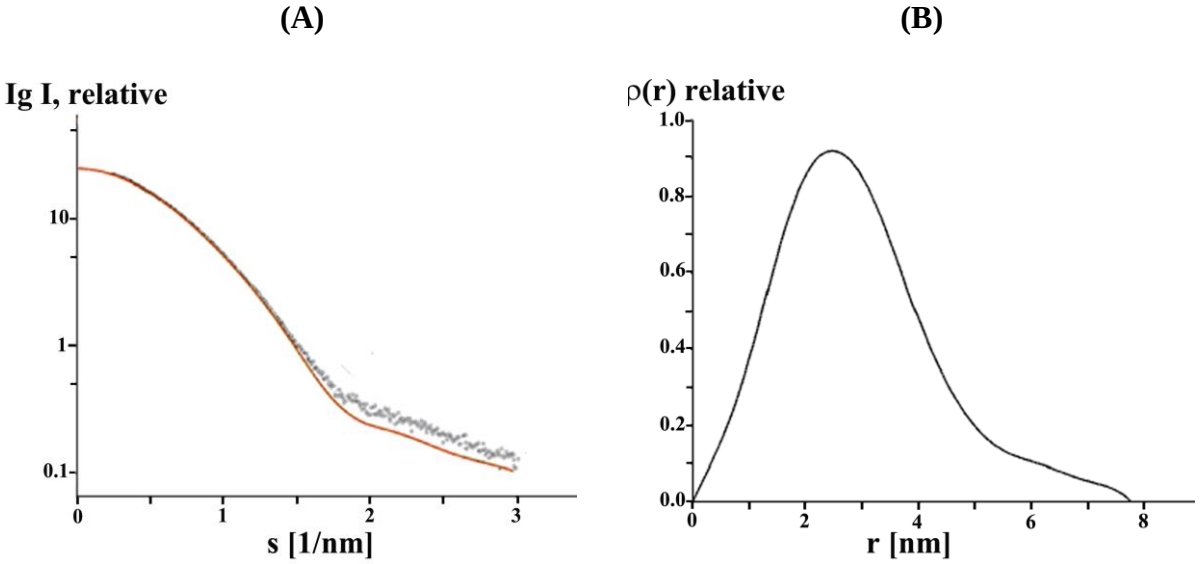


Figure 3.1.4: SAXS experiment of *S. cerevisiae* subunit F. *(A)* Scattering curve obtained from experiment (●) and the fitting curve (—) of subunit F. *(B)* Distance distribution functions of subunit F. For clarity the scattering curve unit of subunit F is displaced by a logarithmic unit (87).

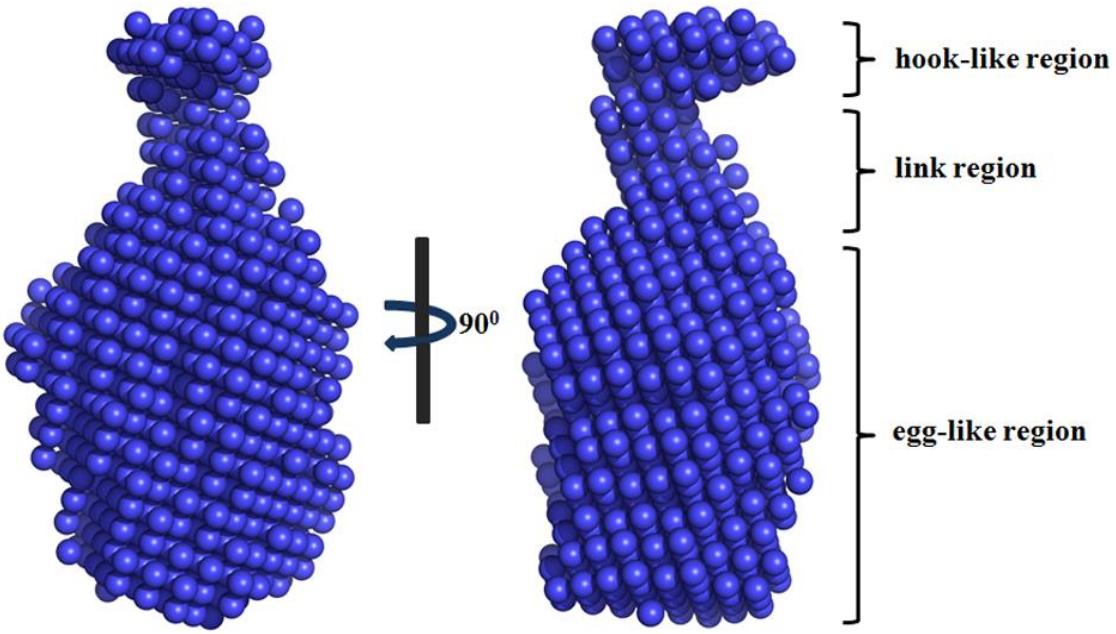


Figure 3.1.5: Picture showing domain features of subunit F solved by SAXS, rotated 90º along the Y axis (87).

3.1.5 Production and purification of $F_{1\text{-}94}$

Based on the previous study, the N- and C-terminal domain of subunit F were shown to have globular and extended conformations, respectively. The extended C-terminal region is speculated to be flexible and dynamic in nature (85,87). Based on the secondary structure prediction, $F_{1\text{-}94}$ construct was generated and $F_{95\text{-}118}$ was thought to be C-terminal flexible region which might have an effect on the dynamic stability of N-terminal region. Therefore, the generation of $F_{1\text{-}94}$ shall enable to find out the effect of the C-terminal domain on the entire structure as well as the correct positions of both termini in the low resolution shape of subunit F. The gene encoding $F_{1\text{-}94}$ was cloned by PCR method which is reported recently (87). The over-expressed recombinant protein $F_{1\text{-}94}$ was purified in two steps. Firstly, metal chelate affinity chromatography (Ni^{2+}-NTA) technique was used, where the protein was eluted using imidazole-gradient (0-400 mM). The protein fractions, eluted between 75-200 mM imidazole-gradient, were pooled together and applied onto the gel-filtration (Superdex^{TM} 75 10/300 GL) column for further purification (Figure 3.1.6). Protein quality was checked by an SDS-PAGE gel. The protein was then analysed by mass spectrometry and showed a mass of 11,155.94 Da (87), confirming the presence of protein $F_{1\text{-}94}$.

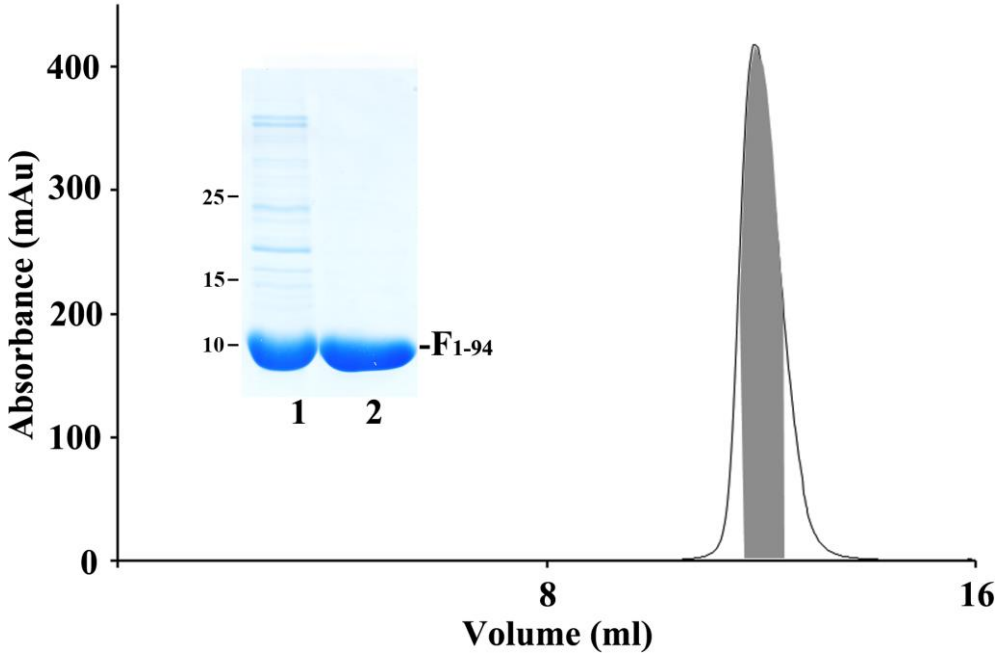


Figure 3.1.6: Purification of F_{1-94} from *S. cerevisiae* V-ATPase. *(A)* Elution profile of F_{1-94} protein, applied onto gel filtration Superdex S75 column as a final step of purification. Insert in figure shows SDS gel where *lane 1* represents protein purified after Ni^{2+}-NTA affinity chromatography and *lane 2* represents protein purified after gel chromatography by using Superdex 75 column (87).

Circular dichroism spectroscopy was utilized to confirm the presence of secondary structure in F_{1-94}. Average secondary structure composition was calculated, showing 40% α-helix, 32% β-sheet and 28% random coil present in the structure (Figure 3.1.7).

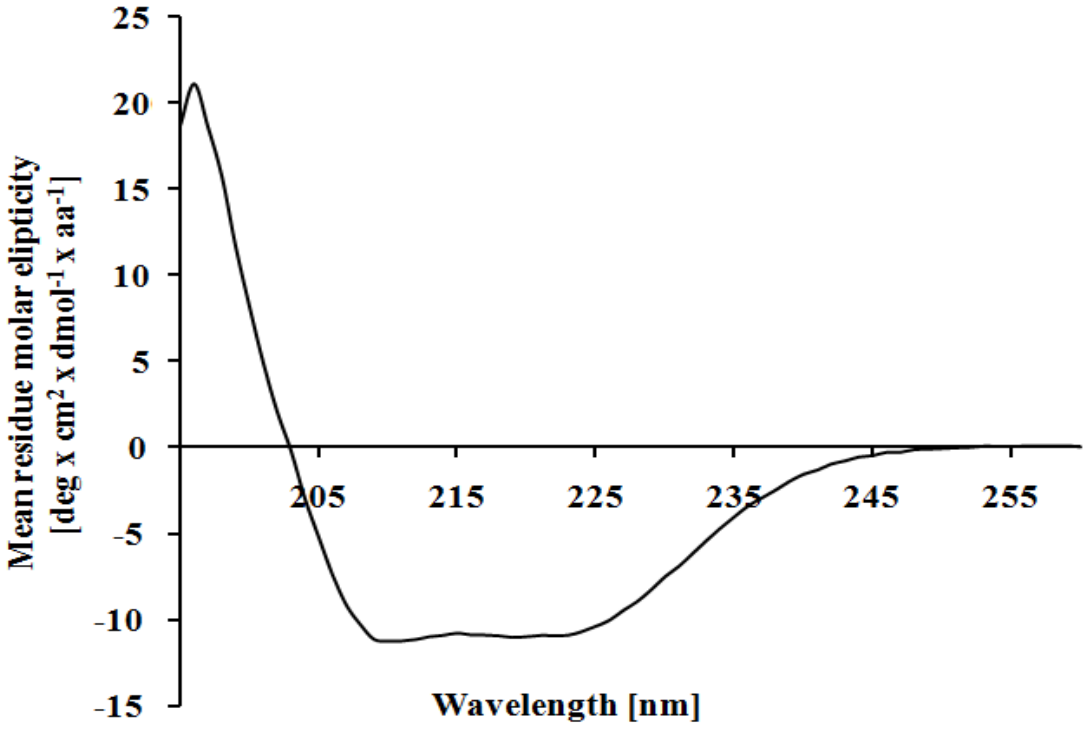


Figure 3.1.7: CD spectrum of recombinant protein F_{1-94} (87).

3.1.6 Low resolution shape of $F_{1\text{-}94}$ in solution

SAXS experiments were performed to further investigate $F_{1\text{-}94}$. The main aim of doing these experiments was to look into the effect of C-terminus truncation on the tertiary structure of the protein as well as to identify the position of termini of subunit F. The scattering curve generated from the SAXS data of $F_{1\text{-}94}$ in solution is presented in figure 3.1.8A. The solution shape of $F_{1\text{-}94}$ was calculated, producing a good fit to the experimental scattering curve in the entire range of scattering with a discrepancy of $\chi 2 = 1.175$ (Figure 3.1.8A). Comparing the SAXS-data of both subunit F and $F_{1\text{-}94}$, it was found that the R_g value of the truncated protein reduced slightly by 0.19 ± 0.02 nm, whereas the D_{max} of $F_{1\text{-}94}$ decreased drastically (D_{max} = 7.0 ± 0.3 nm) (Figure 3.1.8B) (87). The calculated molecular mass of $F_{1\text{-}94}$ reflected that $F_{1\text{-}94}$ is monomeric in solution. Figure 3.1.9 shows the low resolution molecular shape of the $F_{1\text{-}94}$ determined *ab initio*. $F_{1\text{-}94}$ also has an egg-like domain similar to the entire F subunit but the hook-like region is missing. An interesting observation is that the truncation of C-terminal region does not alter the shape and dimensions of the larger major domain in the case of both proteins. Superimposition of the major domain of subunit F and $F_{1\text{-}94}$ allowed to locate the C-terminus of subunit F, which is situated on the upper hook-like domain (Figure 4.1.2).

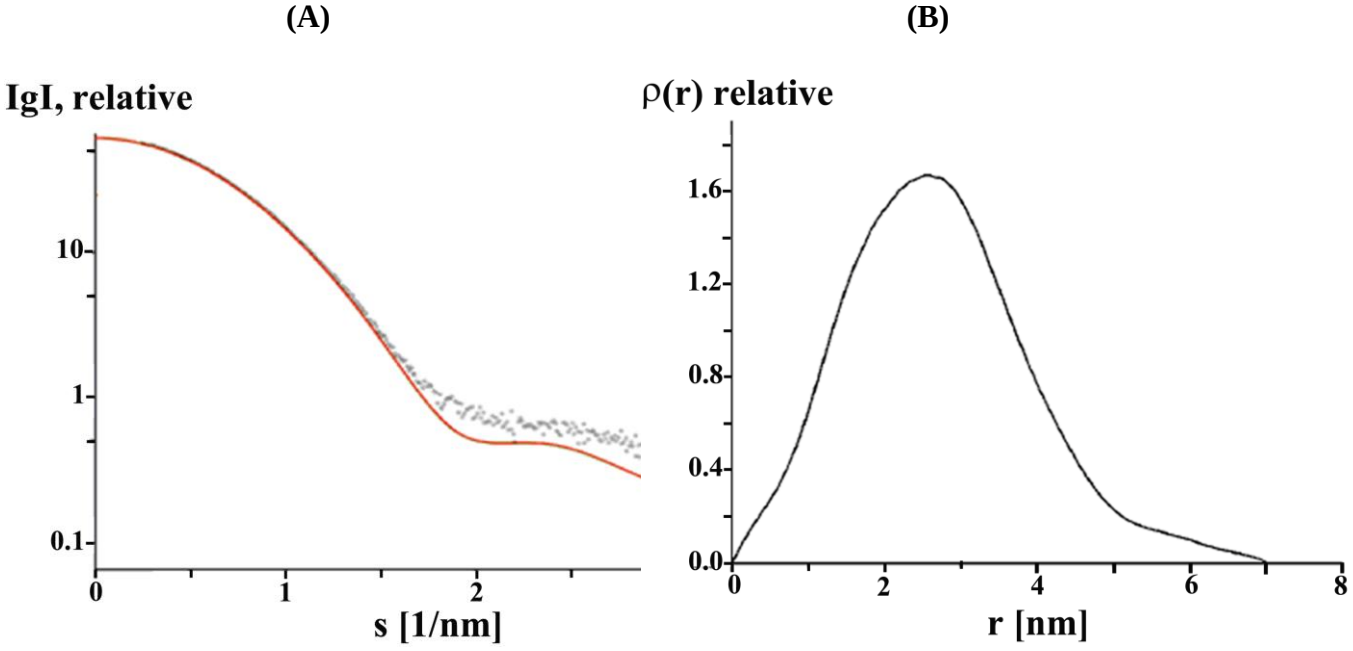


Figure 3.1.8: SAXS experiment of $F_{1\text{-}94}$. *(A)* Scattering curve obtained from experiment (●) and the fitting curve (-) of $F_{1\text{-}94}$. *(B)* Distance distribution curve of $F_{1\text{-}94}$. For clarity the scattering curve unit of $F_{1\text{-}94}$ is displaced by a logarithmic unit.

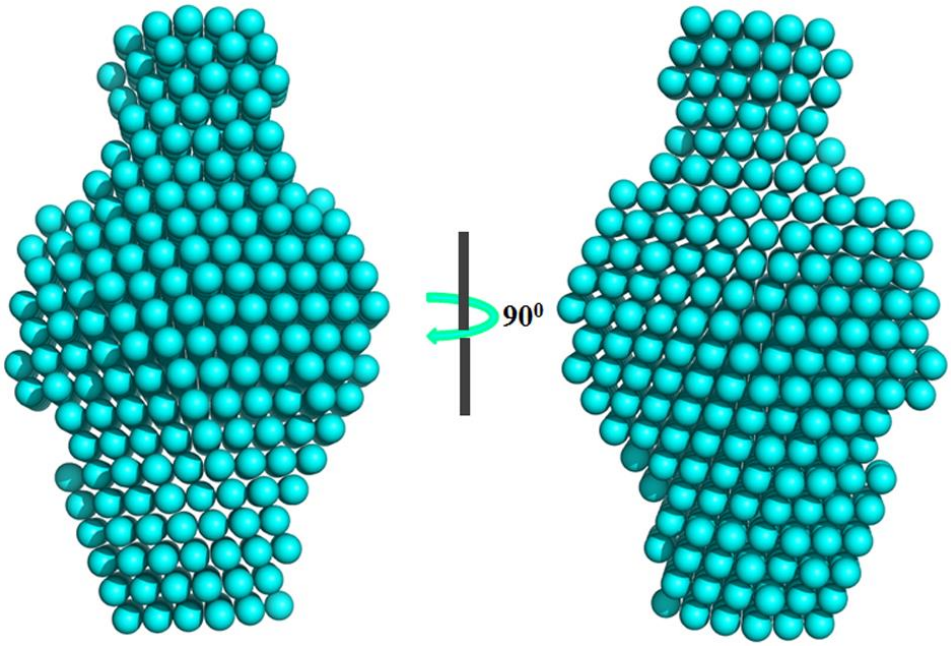


Figure 3.1.9: Low resolution solution shape of $F_{1\text{-}94}$, rotated 90º along the Y axis (87).

3.1.7 Secondary structure of the C-terminus of subunit F ($F_{90\text{-}116}$)

According to the secondary structure prediction, the C-terminal region has a tendency to form a helix between residues E_{104}-F_{116}, which is predicted to be dynamic in solution. In order to investigate the C-terminal domain of subunit F, the solution structure of a 27-residues peptide, $F_{90\text{-}116}$ ($_{90}$AILEIPSKDH PYDPEKDSVL KRVRKLF$_{116}$), was determined by NMR. The peptide was synthesized by microwave technique and then purified by reverse phase HPLC at the Division of Structural Biology and Biochemistry, SBS, as described in section 2.2.6. Circular dichroism spectroscopy was performed to estimate the percentage of secondary structure presence in the peptide. The spectrum was measured in a range of 190-260 nm wavelength (Figure 3.1.10A). The peptide contains 45% α-helix and 55% unstructured features. The formation of α-helix was triggered by Trifluoroethanol (TFE). This result is in line with the predicted secondary structure based on the amino-acid sequence of $F_{90\text{-}116}$ (Figure 3.1.10B). The molar ellipticity ratio at 208 nm and at 222 nm is 0.98 (87).

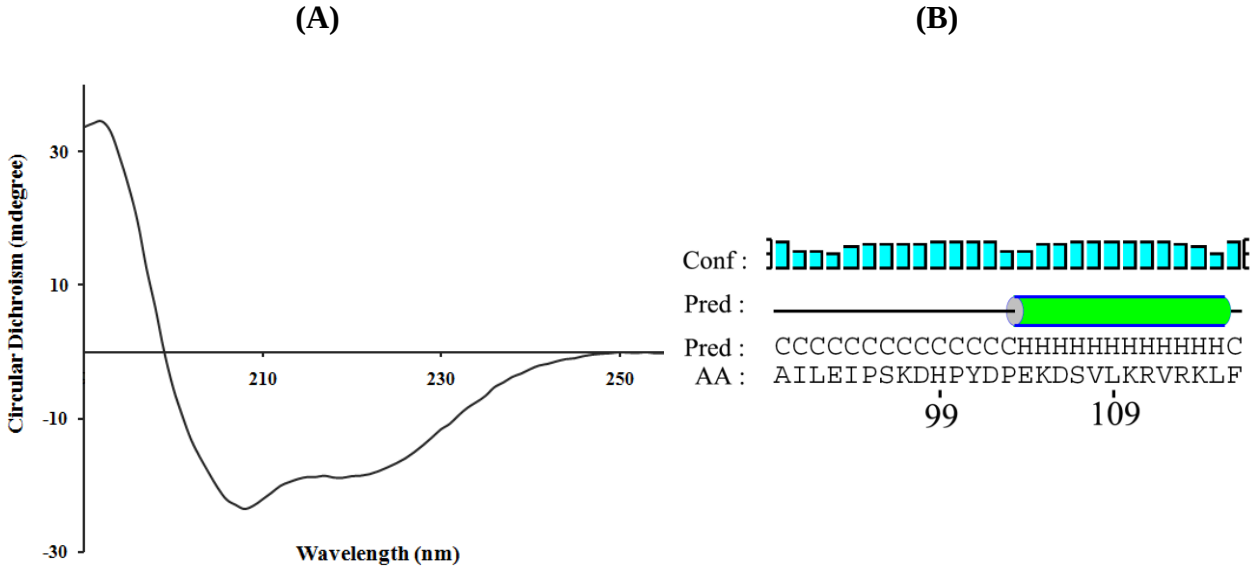


Figure 3.1.10: CD spectrum of peptide $F_{90\text{-}116}$. *(A)* Spectrum showing Far UV-CD region of peptide $F_{90\text{-}116}$ in the presence of TFE. *(B)* Prediction of Secondary structure of peptide $F_{90\text{-}116}$ by psipred software, where α-helix= , β-strand= , coil= and confidence of prediction=.

3.1.8 NMR structure of $F_{90\text{-}116}$ in solution

It is very essential to solve the peptide $F_{90\text{-}116}$ structure to gain deeper insight about the structural behaviour of the hook-like region in solution (87). NMR spectroscopy was used to acquire homonuclear ^{1}H-^{1}H TOCSY and ^{1}H-^{1}H NOESY experiments in solution and standard sequential assignment procedures were followed to assign residues of the peptide (Figure 3.1.11A) (105). Based on HA chemical shifts information, an online software PREDITOR (87) was used to predict the secondary structure content, showing an α-helical stretch at the C-terminus of the peptide spanning between P_{103}-R_{113} (Figure 3.1.11C). A representative cross peaks assignment for the NH-NH region is shown in figure 3.1.11A. Based on the assigned NOESY spectrum, NH–NH, Hα–NH (i, i+3), Hα–Hβ (i, i+3) and Hα–NH (i, i+4) connectivity were drawn which is presented in figure 3.1.11B. Analysis of NOE connectivity data supports the formation of an α-helical stretch in the C-terminus (87). 100 structures were calculated and among them 10 structures which have the lowest energy were chosen for further analysis. The calculated overall RMSD of 10 ensembled structures for the backbone atoms and heavy atoms was 0.303 Å and 1.46 Å, respectively (Figure 3.1.12A-B). All structures had energies lower than -100 kcal/mol. There were no NOE and dihedral violations greater than 0.3 Å and 5º, respectively. Table 3.1.1 represents the summary of the statistics for 10 lowest energy structures. The calculated structure has a total length of 27.30 Å with a helical region (residues P_{103}-R_{113}, 14.9 Å) flanked with N- and C-terminal unstructured regions (residues A_{90}-D_{102} and K_{114}-F_{116}, respectively) (Figure 3.1.12B). Figure 3.1.13C shows the surface

electrostatic potential of F_{90-116} generated by Pymol program (87). The charged distribution showed that the helix is positive-charged which is contributed by basic residues like K_{110}, R_{111}, R_{113} and K_{114} (87).

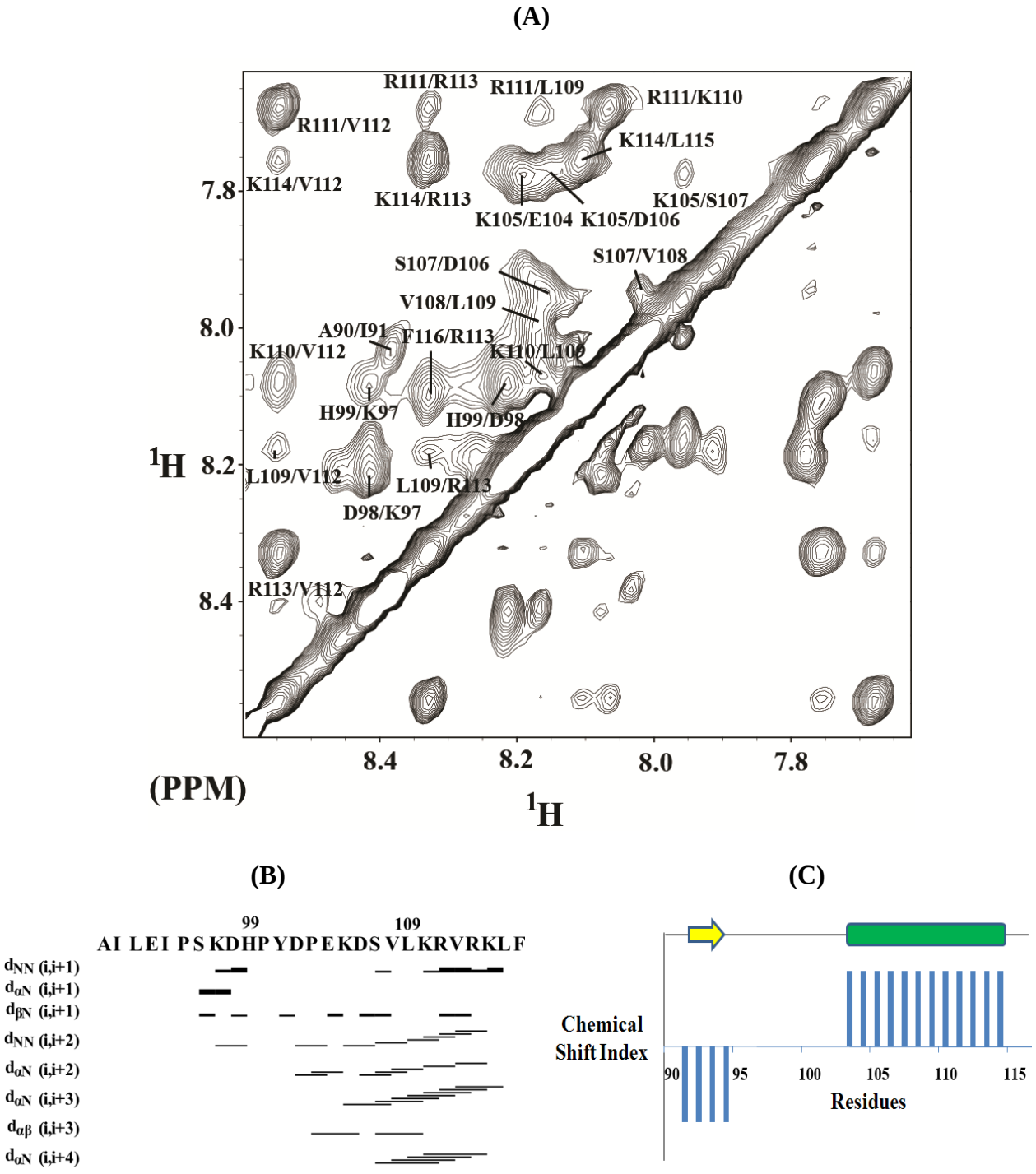


Figure 3.1.11: 2D homonuclear NMR experiment of peptide F_{90-116} (87). *(A)* Assignment of cross-peaks at HN-HN region of the NOESY spectrum of peptide F_{90-116}. Sparky 3.1 program was used for peak picking and based on TOCSY and NOESY spectra, cross peaks were identified. *(B)* The NOE connectivity plot of F_{90-116}. *(C)* Prediction of secondary structure using Hα chemical shifts of F_{90-116} by online program PREDITOR (87).

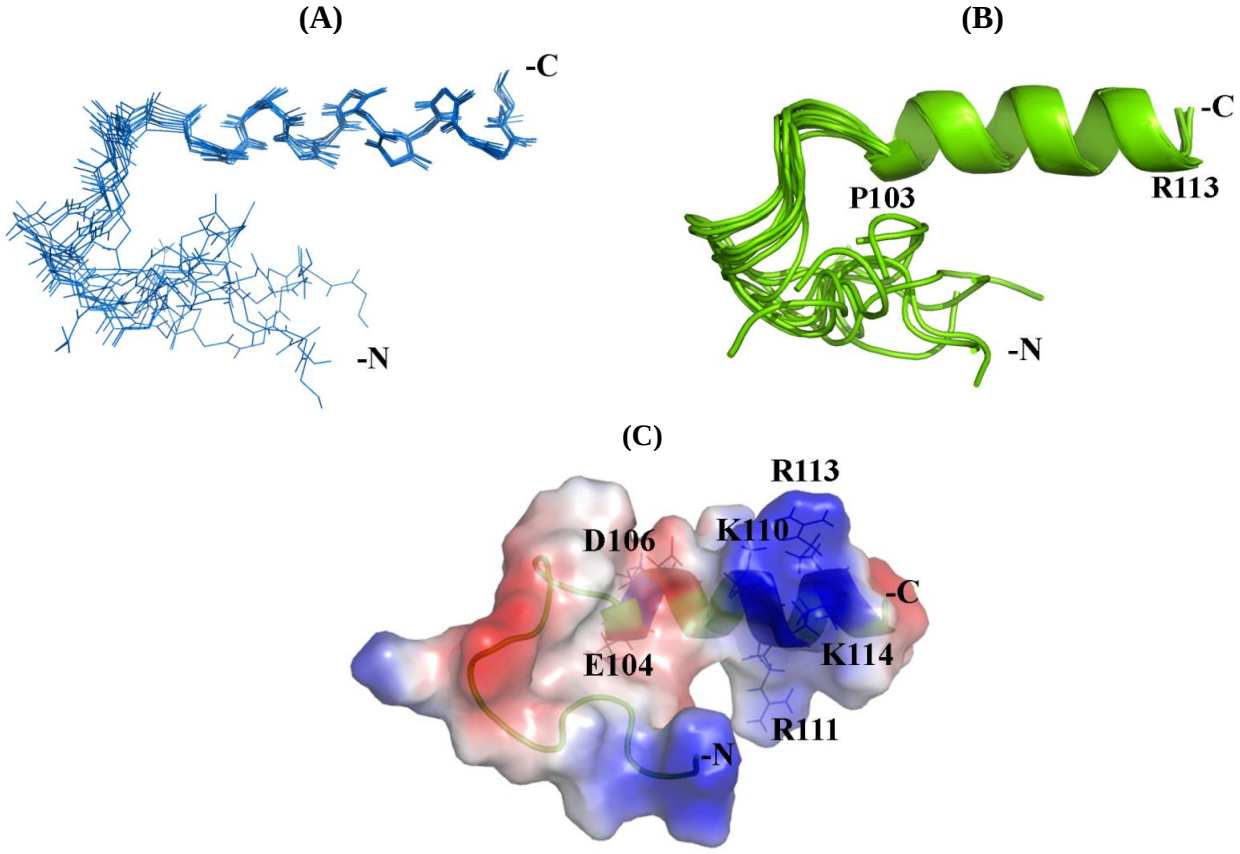
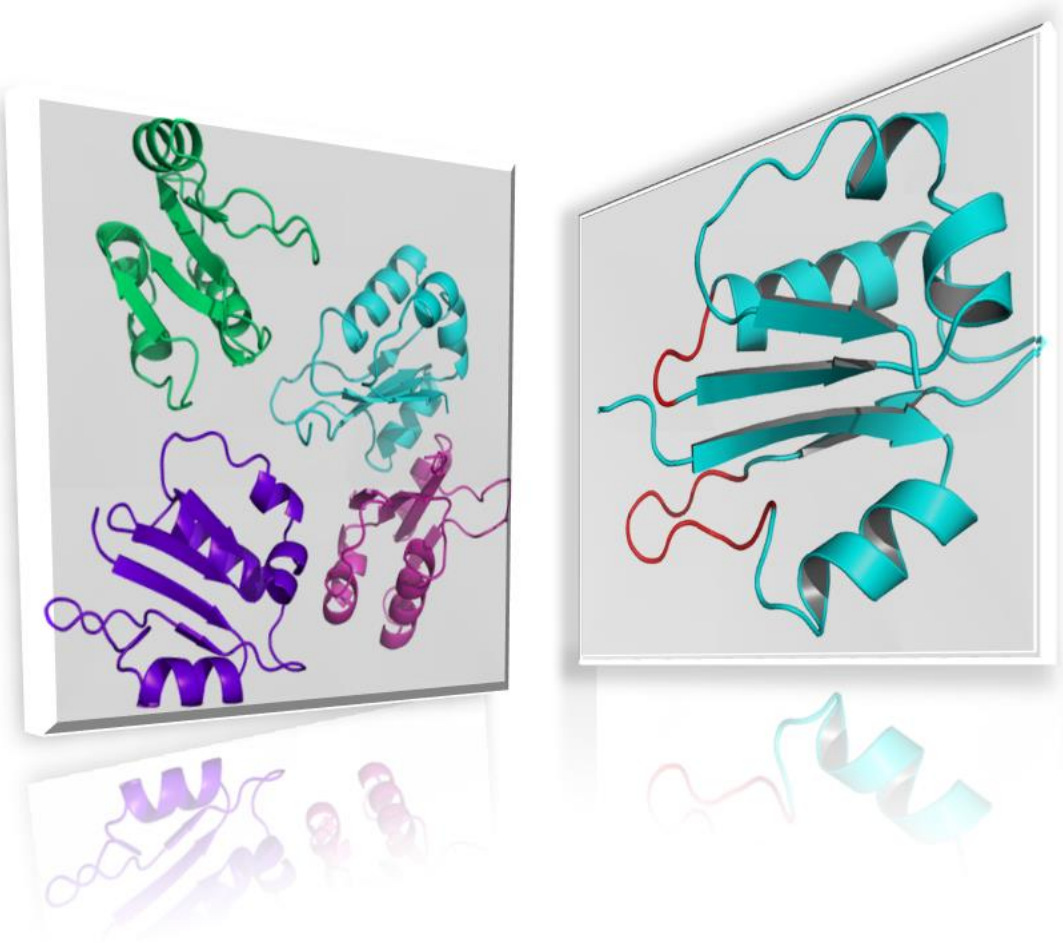


Figure 3.1.12: Solution structure of F_{90-116} by NMR. Line *(A)* and cartoon representation *(B)* of superimposed 10 lowest energy NMR structures of F_{90-116}. *(C)* Surface electrostatic charge of F_{90-116}, showing positive (blue) and negative (red) potentials (87).

Table 3.1.1: NMR structural statistics of peptide F_{90-116} (87).

Total number of residues	27
Total number of NMR restraints	410
Intraresidual (|i–j|=0)	51
Short range (|i–j|≤ 1)	69
Medium-range (2 ≤|i–j|≤ 5)	94
Long-range (|*i–j*| > 5)	2
Dihedral angle constraints	30
Ramachandran plot statistics (%)	
Most favoured regions	58.2
Additionally allowed regions	30.5
Generously allowed regions	11.4
Disallowed region	0
Structural precision for well ordered region	
rmsd backbone (P_{103}-F_{116})	0.303 Å
rmsd heavy atoms	1.46 Å

3.2 The crystallographic structure of subunit F, $F_{1\text{-}94}$ of V_1V_O ATPase from *S. cerevisiae*



3.2.1 Introduction

V-ATPases are known to catalyze ATP into ADP leading to release of energy which is used for proton translocation through the transmembrane domain into the intracellular vesicles. As a part of central stalk, subunit F plays an important role during the coupling event. Subunit F is proposed to undergo structural alterations and interact with subunit A, D and E during catalysis in nucleotide-dependent manner (41,45,133). Among all ATPase family, V-ATPases have a unique feature called reversible disassembly. During this process V_1 part dissociate from V_O part causing structural rearrangement in the stalk region. Due to the dynamic rearrangement of stalk region, high resolution structural information is required which might help us to understand the complex mechanism of this part. With the intention to get a high resolution structure, crystallization studies were done for both subunits F and F_{1-94}. Among these, F_{1-94} produced crystals and diffracted to 2.64 Å resolution. Due to the lack of a suitable structural model, molecular replacement did not work, which directed us to generate selenomethionine protein. Initially F_{1-94} had two methionine residues at the very N-terminal region which were not enough to produce selenium signal during the fluorescence scanning of the selenomethionyl protein crystals. To increase the number of methionine residues, F_{1-94}I69M mutant was produced and finally selenomethionyl F_{1-94}I69M protein was crystallized to solve the structure.

3.2.2 Crystallization of F_{1-94}

Crystallization drops for subunit F_{1-94} were setup in both sitting and hanging drops, using 3 mg/ml and 5 mg/ml protein concentrations. Initial crystals were observed in crystal screen 1 of Hampton Research condition 6 (30% (v/v) polyethylene glycol 4000 (PEG 4000), 0.10 M Tris/HCl, pH 8.5 and 0.20 M $MgCl_2$, 6xH_2O) at 18^o C. In both sitting and hanging drops, numerous tiny crystals were observed after quick nucleation and the crystals looked multiple when observed under microscope (Figure 3.2.1A). These crystals were small in size (0.17 mm x 0.07 mm x 0.01 mm). Few crystals which did not appear multiple under the microscope were tested in the in-house machine and diffracted maximum to 7 Å resolution machine (Figure 3.2.1B). After indexing the diffraction data by HKL2000 suite program, it showed higher unit cell parameters (a = 545 Å, b = 177 Å, c = 295 Å). To improve the quality of the F_{1-94} crystals and diffraction resolution, a series of optimizations were carried out by varying crystallization parameters.

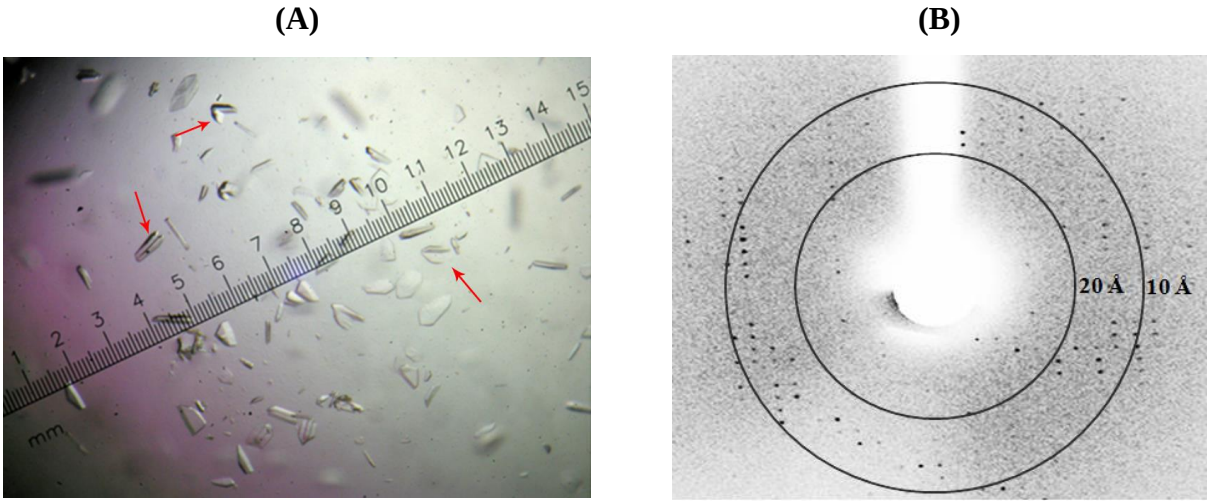


Figure 3.2.1: Initial F_{1-94} crystals, which were crystallized in the condition of 30 % (v/v) PEG 4000, 0.10 M Tris/HCl, pH 8.5, 0.20 M $MgCl_2$, 6xH_2O at 18° C temperature. *(A)* Crystals of F_{1-94}. Red arrows indicate the presence of few multiple crystals observed under microscope. *(B)* The best diffraction pattern of F_{1-94} crystal before optimization. Circles show resolution shell for 20 Å and 10 Å, respectively.

Different concentrations of glycerol (5-25%) were added to reduce the rate of nucleation, resulting in uniformly shaped crystals with a dimension of 0.3 mm x 0.04 mm x 0.02 mm (Figure 3.2.2A-C). Different concentrations of PEG 4000 (20-35%) were also tested in the grid screen but the protein crystallized only with 30% PEG 4000 (Figure 3.2.3A-C). Varying magnesium chloride hexahydrate ($MgCl_2$, 6xH_2O) concentrations (0.01-0.4 M) in the grid screen had a major effect on crystal size and rate of nucleation (Figure 3.2.4A-C). Probably due to salting in effect at 0.1 M $MgCl_2$, 6xH_2O concentration, the nucleation rate was considerably reduced resulting in bigger sized crystals with a dimension of 0.4 mm x 0.07 mm x 0.03 mm.

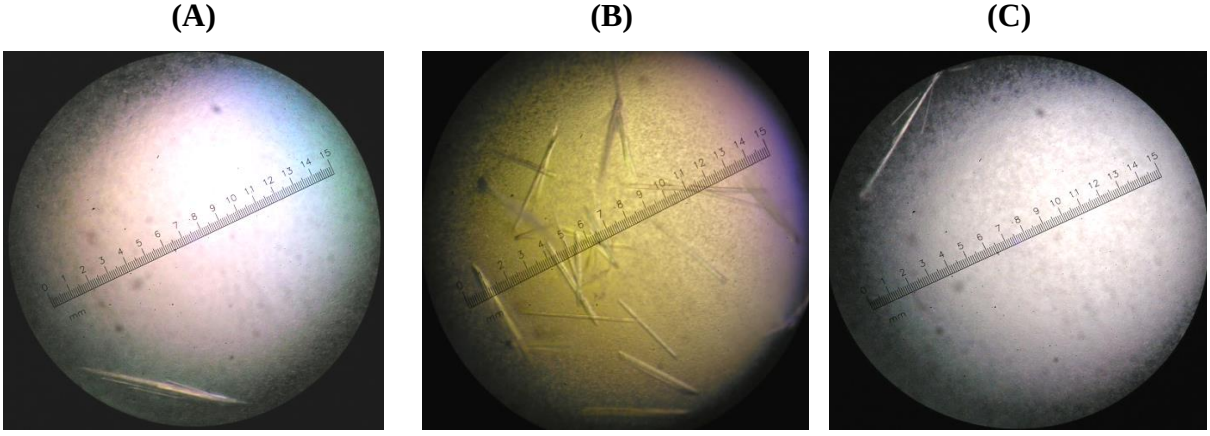


Figure 3.2.2: Selected pictures of F_{1-94} crystals at different glycerol concentrations. Formation of F_{1-94} crystals in the condition of 30% (v/v) PEG 4000, 0.20 M $MgCl_2$, 6xH_2O and 0.10 M Tris/HCl, pH 8.5 at 5% glycerol *(A)*, 10% glycerol *(B)* and 15% glycerol *(C)*.

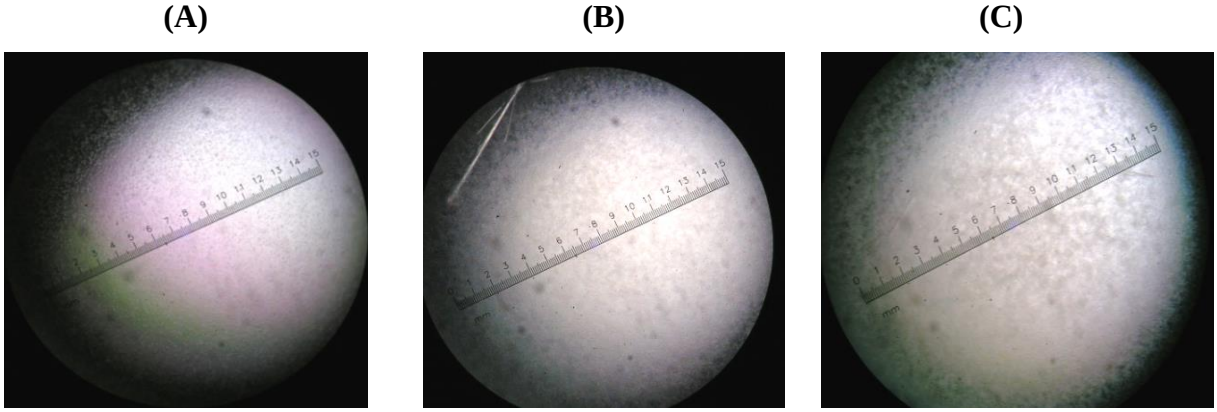


Figure 3.2.3: Representative pictures of F_{1-94} crystals where percentage of PEG 4000 was varied. Crystallization drop in the condition of 0.20 M $MgCl_2$, $6xH_2O$, 0.10 M Tris/HCl, pH 8.5 and 10% glycerol at 25% PEG 4000 *(A)*, 30% PEG 4000 *(B)* and 35% PEG 4000 *(C)*.

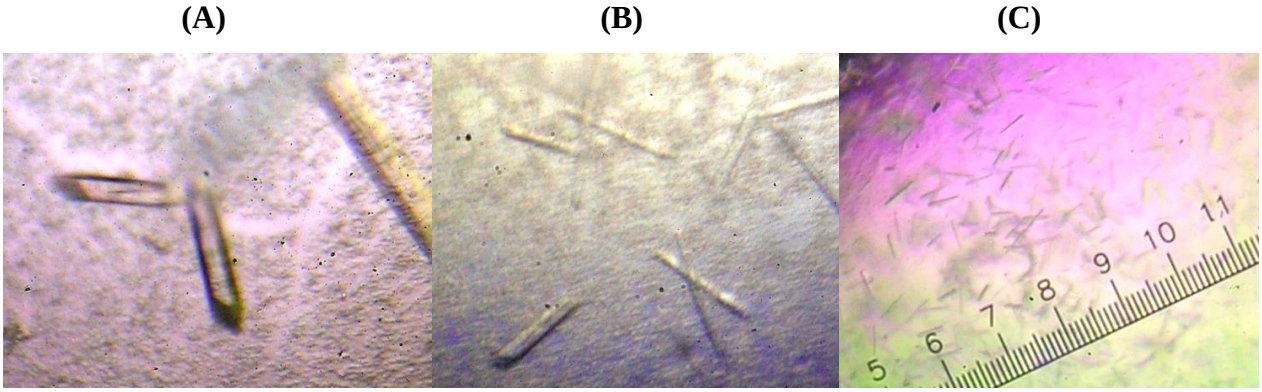


Figure 3.2.4: Selected pictures of crystals yielded after varying $MgCl_2$, $6xH_2O$ concentration. *(A)* Crystals in the condition of 30% PEG 4000, 0.10 M Tris/HCl, pH 8.5 and 10% glycerol at 0.1 M $MgCl_2$, $6xH_2O$ *(A)*, 0.25 M $MgCl_2$, $6xH_2O$ *(B)* and 0.4 M $MgCl_2$, $6xH_2O$ *(C)*.

Variations of pH (7.5-9.0) were done to improve the crystal quality. At pH 8.8 of 0.10 M Tris/HCl buffer, multiple crystals were not observed under the microscope (Figure 3.2.5A-C). Thereafter, nucleation rate was slower down using paraffin and silicon oil yielded microcrystalline precipitation (Figure 3.2.6A-C).

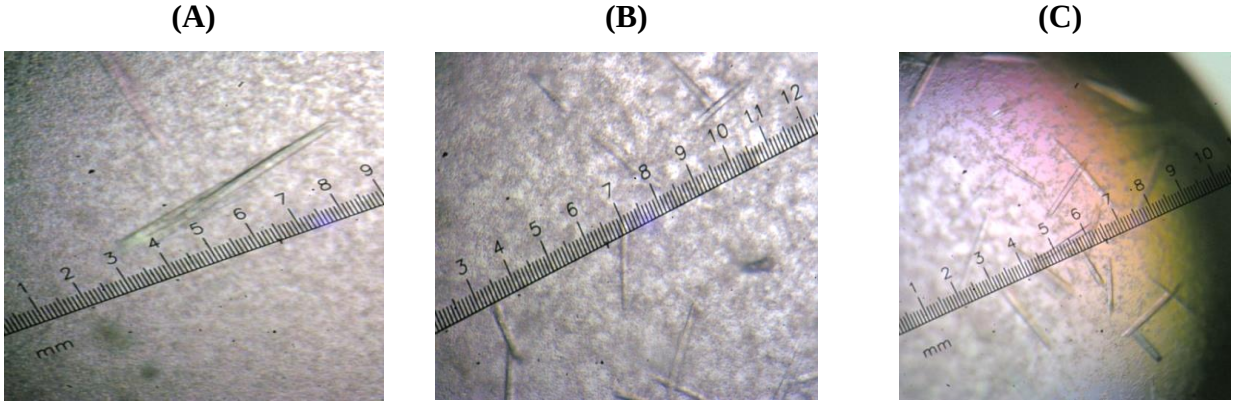


Figure 3.2.5: Selected crystals pictures where pH of 0.1 M Tris/HCl buffer was changed. Crystals appeared in 30% PEG 4000, 0.10 M $MgCl_2$, $6xH_2O$, 10% glycerol and 0.10 M Tris/HCl at pH 8.2 *(A)*, pH 8.4 *(B)* and pH 8.8 *(C)*

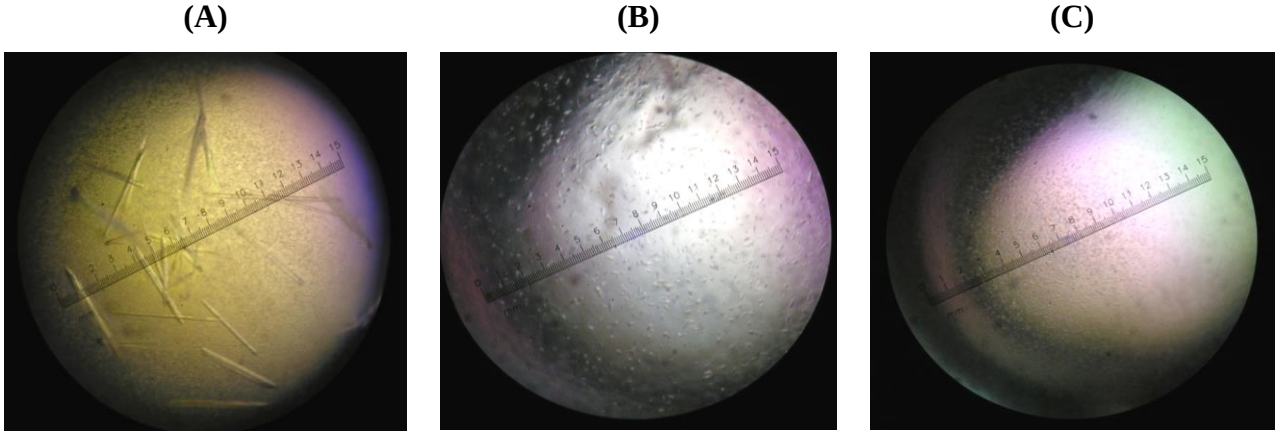


Figure 3.2.6: Crystal pictures of F_{1-94} after covering the reservoir solution by paraffin and combination of paraffin and silicon oil. Crystals were formed in the condition of 30% (v/v) PEG 4000, 0.20 M $MgCl_2$, 6xH_2O, 0.1 M Tris/HCl, pH 8.8, 10% glycerol without covering reservoir solution *(A)*, reservoir solution was overlaid by paraffin *(B)* and reservoir solution was covered by mixture of paraffin and silicon oil *(C)*.

List of other parameters, which were attempted during optimization are given below:

- Different grades of PEG e.g. PEG 400, PEG 1500, PEG 2000, PEG 3350, PEG 5000, PEG 6000, PEG 8000.
- Different types of chloride salt e.g. CsCl, LiCl, NH_4Cl, $NiCl_2$, 6xH_2O and KCl.
- Different types of magnesium salt e.g. $(CH_3COO)_2Mg$, 4H_2O, $Mg(CHOO)_2$, 2H_2O, $Mg(NO_3)_2$, 6xH_2O and $MgSO_4$, H_2O.
- Addition of 2-methyl-2,4-pentanediol (MPD) instead of PEG 4000.
- Addition of basic amino acid e.g. arginine.
- Variations of protein concentration.
- Variations of temperature.
- Micro and macro seeding.
- Additive screening.

Based on the grid screens, crystal drops were set up in the condition of 30 % (v/v) PEG 4000, 0.10 M of $MgCl_2$, 6xH_2O, 0.10 M Tris/HCl, pH 8.5 and 10 % (v/v) of glycerol which yielded better diffracting crystals (Figure 3.2.7B-C). The optimized crystals were harvested and tested in in-house Rigaku machine.

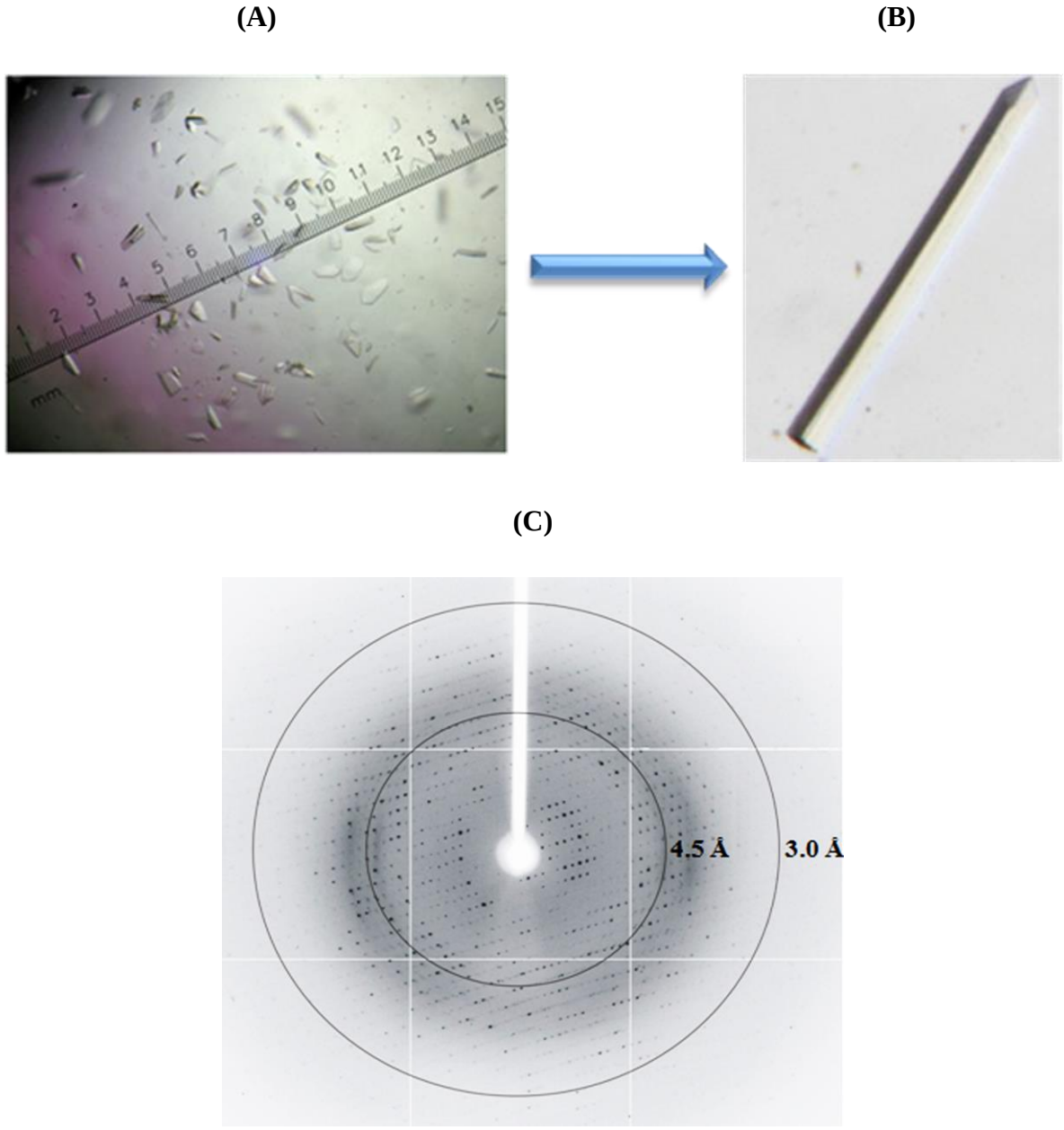


Figure 3.2.7: Optimization of F_{1-94} crystals. *(A)* Crystals in conditions of 30% PEG 4000, 0.20 M $MgCl_2$, 6xH_2O, 0.10 M Tris/HCl (pH 8.5). *(B*) Pictures of a crystal after optimization using 30% PEG 4000, 0.10 M $MgCl_2$, 6xH_2O, 0.10 M Tris/HCl (pH 8.8) and 10% glycerol. (C) Representative diffraction image of F_{1-94} of V-ATPase from *S. cerevisiae*. Circles showing resolution shell for 4.5 Å and 3.0 Å, respectively (109).

A single-wavelength data set of a F_{1-94} crystal was collected at the in-house Rigaku machine where temperature was maintained at 100 K (Figure 3.2.7C). A total of 360 frames were collected and the exposure time was 5 min per frame with an oscillation range of 1º. HKL2000 suite of program (110) was employed to index, integrate and scale all diffraction data. The summary of data processing statistics is shown in Table 3.2.1.

Table 3.2.1: Summary of data collection parameters and processing statistics of F_{1-94} crystal (109).

Wavelength (Å)	1.54178
Space group	$C222_1$
Unit cell parameters (Å)	
a =	47.206
b =	160.264
c =	102.494
Molecules in asymmetric unit	4
Solvent content (%)	46.03
Resolution limits (Å)	30.0-2.64 (2.73-2.64)
No. of images	360
No. of reflections	26881
Unique reflections	11788
Multiplicity	6.4 (6.3)
Completeness (%)	98.9 (99.0)
R_{merge} (%)*	10.2 (28.0)
<I//σ(I)>	13.7 (7.1)

*$R_{merge} = \Sigma\Sigma_i|I_h - I_{hi}|/\Sigma\Sigma_i I_h$, where I_h is the mean intensity for reflection h.

The native F_{1-94} crystal diffracted to 2.64 Å resolution. The crystal belongs to the orthorhombic space group $C222_1$ having unit cell parameters of *a* = 47.21 Å, *b* = 160.26 Å and *c* = 102.49 Å. A representative diffraction pattern is shown in figure 3.2.2C. Cell content analysis showed 46.03% solvent content and 2.28 $Å^3$ Da^{-1} Matthews coefficient (V_m) (134), considering four molecules in one asymmetric unit (109). Initially molecular replacement was tried to solve the phase where the structures of subunit F of the related A-ATP synthase from *P. furiosus* (PDB ID: 2QAI) and from *A. fulgidus* (PDB ID: 2I4R) were used as initial model. However, no final solution for structure determination could be found. Hence, my focus was directed towards obtaining crystals from selenomethionine substituted F_{1-94}.

3.2.3 Production, purification and crystallization of selenomethionyl F_{1-94} protein

Multi-wavelength anomalous diffraction (MAD) technique is a fundamental approach for the phasing of protein crystal structures (135). A single protein crystal is required to contain one or more heavy atoms that cause anomalous diffraction from the incoming X-rays at the wavelength used for the diffraction experiment. The most suitable heavy atom for this

purpose is selenium. Selenomethionine is most commonly incorporated into the protein for collecting MAD data. As there was no suitable model for molecular replacement, selenomethionine substituted F_{1-94} protein was produced to solve phasing problems during structure determination. Two methionine residues are present in the His-tagged F_{1-94} protein, one is at the very N-terminus His-tagged region (MKHHHHHHP), whereas the second methionine is at the very first amino acid of the encoded F_{1-94} protein. The details about the production and purification procedure of selenomethionine substituted F_{1-94} protein are given in section 2.2.9. Crystallization setup for selenomethionine substituted F_{1-94} was done using hanging-drop vapor diffusion method at 18° C at optimized solution condition, containing 30% (v/v) PEG 4000, 0.10 M Tris/HCl, pH 8.8, 0.10 M $MgCl_2$, 6xH_2O and 10% glycerol. In the optimized condition, 2 µl protein solution was mixed together with an same volume of mother solution and equilibrated at 500 µl volume of reservoir solution. The nucleation process initiated within two days leading to the formation of crystals. Figure 3.2.8A-B shows an image of selenomethionine substituted F_{1-94}, which could be successfully crystallized and diffracted to 3.8 Å resolution.

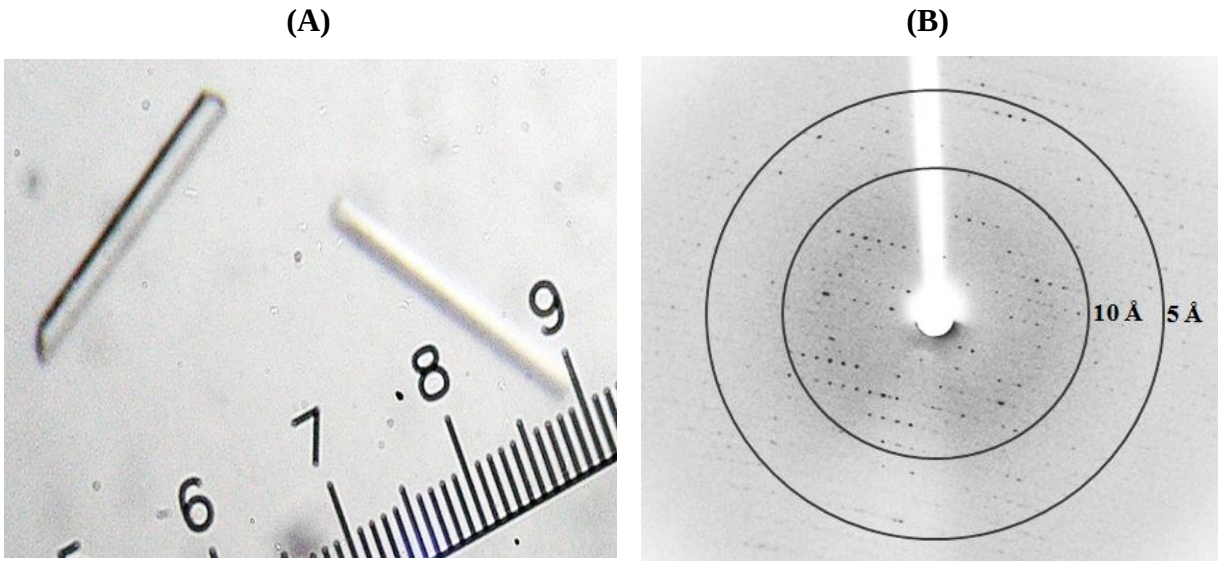


Figure 3.2.8: Crystals of selenomethionine substiuted F_{1-94} protein in the condition of 30% (v/v) PEG 4000, 0.1 M $MgCl_2$, 6xH_2O, 0.1 M Tris/HCl (pH 8.8) and 10% glycerol. *(A)* Crystals of selenomethionine substiuted F_{1-94}. *(B)* Diffraction pattern of selenomethionine substiuted F_{1-94} crystal. Circles showing resolution shell for 10.0 Å and 5.0 Å

A fluorescence scan was performed at the absorption edge of selenium with selenomethionyl native $F_{1–94}$ crystals which produced very weak signal (Figure 3.2.9). The data set that was collected and processed resulted in anomalous signal which is not useful for phase

determination. The most probable reason for the weak signal is attributed to the fact that there are only two methionine residues situated at the very N-terminus end of the protein.

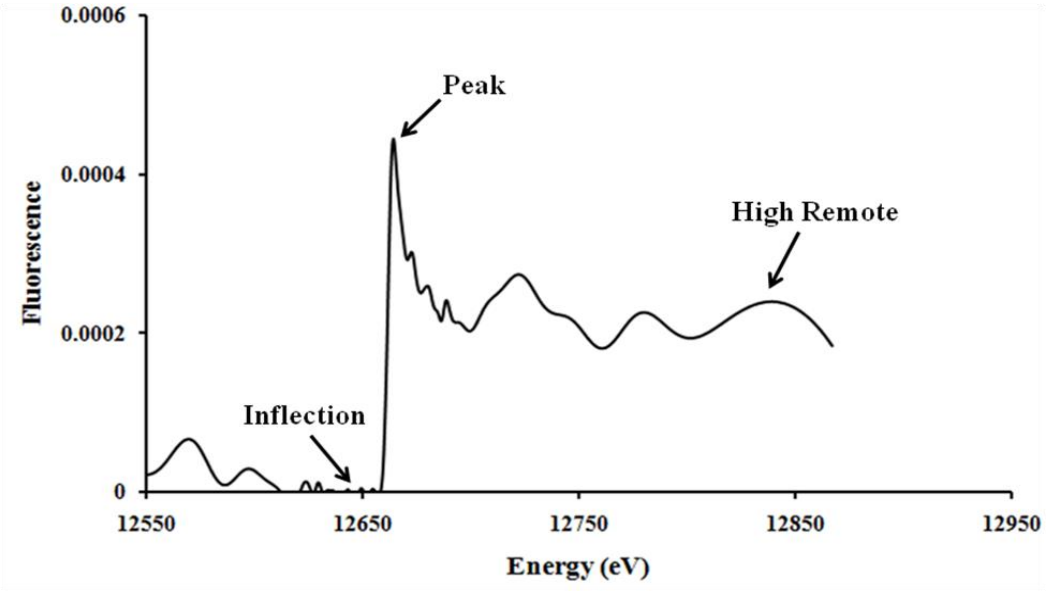


Figure 3.2.9: Fluorescence scanning of SeMet $F_{1\text{-}94}$ protein crystal.

3.2.4 Cloning, expression and purification of $F_{1\text{-}94}$ mutants

In an effort to improve the speed and accuracy of protein structure determination by X-ray crystallography, the methionine residues should be spread throughout the whole protein structure (136). Two methionine residues are present in the His-tagged $F_{1\text{-}94}$ protein, one at the very N-terminus of the His-tagged region and the second methionine is the very first amino acid of the encoded $F_{1\text{-}94}$ protein. To increase the number of methionine in the structure of $F_{1\text{-}94}$ protein, mutagenesis was done. Leucine residue of $F_{1\text{-}94}$ was mutated into methionine with the idea that this mutation would not impact significantly on the structure of the whole protein, because methionine is roughly the same size as leucine and flexible enough to fairly occupy the same volume (137,138). Methionine is the only moderately conserved amino acid most frequently replaced by leucine or isoleucine in the course of evolution (139). After multiple sequence alignment analysis, several different mutants, $F_{1\text{-}94}$I8M, $F_{1\text{-}94}$I11M, $F_{1\text{-}94}$L20M, $F_{1\text{-}94}$L67M, $F_{1\text{-}94}$L68M, $F_{1\text{-}94}$I73M, $F_{1\text{-}94}$L20, 67M and $F_{1\text{-}94}$I69M were designed and generated (Figure 3.2.10). Mutations were done in the structural region to minimize flexibility of selenomethionine residue, which might enhance fluorescence signal of the selenium atom. Initial focus was on those leucine residues which are present in the structural region. Based on this, $F_{1\text{-}94}$L20M, $F_{1\text{-}94}$L67M and $F_{1\text{-}94}$L68M mutants were designed. After comparing with the crystal structure of homologous subunit F of *E. hirae* A-ATP synthase (F_{Eh}), $F_{1\text{-}94}$I8M, $F_{1\text{-}}$

$_{94}$I11M and $F_{1\text{-}94}$I73M mutants were designed where the mutated residue in F_{Eh} is exposed outside.

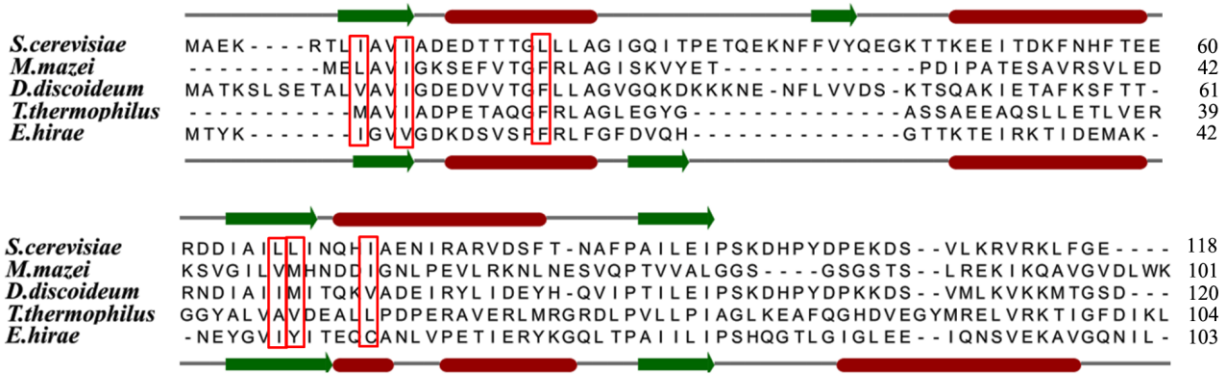


Figure 3.2.10: Sequence alignment of subunit F from eukaryotic V-ATPases. The secondary structure elements of the eukaryotic $F_{1\text{-}94}$ from *S. cerevisiae* and the A-ATP synthase subunit F from *E. hirae* are shown above and below, respectively. Red boxes indicate the position of amino acid where methionine mutation was done.

The mutants were successfully created by overlap extension PCR method with primers *a*-*b* and *c*-*d*, respectively, as mentioned in section 2.2.5. The two amplified DNA fragments, produced after PCR 1, were then annealed together by doing second PCR reaction using the flanking primers *a* and *d* (Figure 3.2.11). The fused DNA fragments were digested with *Nco*I and *Sac*I restriction enzymes and successfully ligated into pET-9d1 vector (86).

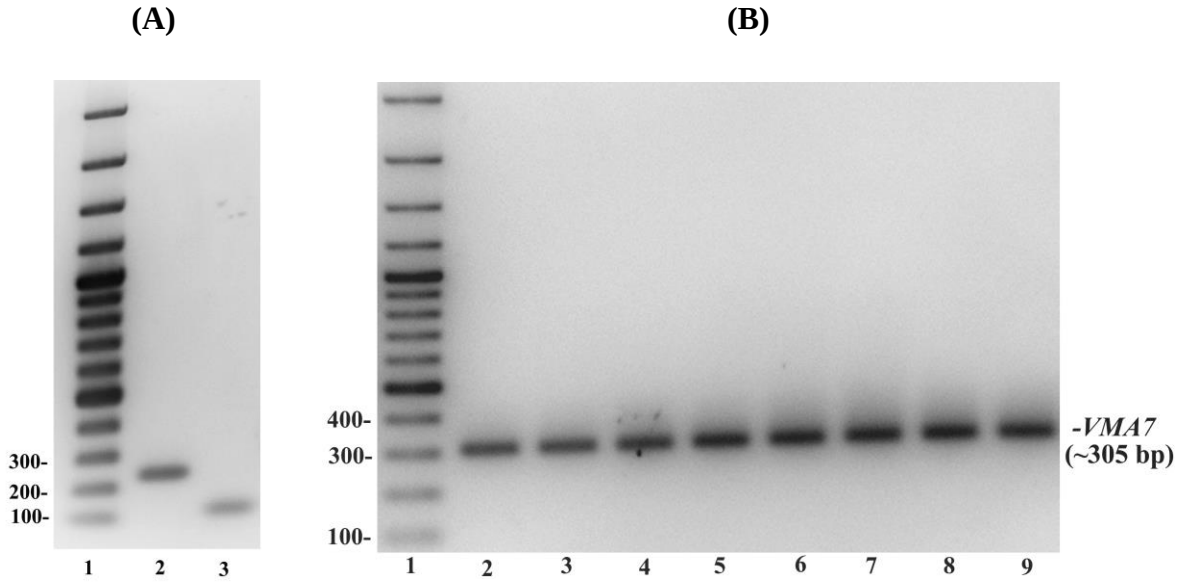


Figure 3.2.11: *(A)* A representative picture of PCR amplification of $F_{1\text{-}94}$ mutants. PCR I products; Marker DNA (100 bp, Fermentas) (*lane 1*); larger fragment (*lane 2*); smaller fragment (*lane 3*). *(B)* PCR II; final products; Marker DNA (100 bp, Fermentas) (*lane 1*); $F_{1\text{-}94}$I8M (*lane 2*); $F_{1\text{-}94}$I11M (*lane 3*); $F_{1\text{-}94}$L20M; (*lane 4*); $F_{1\text{-}94}$L67M (*lane 5*); $F_{1\text{-}94}$L68M (*lane 6*); $F_{1\text{-}94}$I73M (*lane 7*) *and* $F_{1\text{-}94}$L20, 67M (*lane 8*).

Induction and solubility tests were done after successful cloning of $F_{1\text{-}94}$ mutants. All mutant proteins were soluble in 50 mM Hepes, pH 7.0, 300 mM NaCl buffer. The selenomethionine substituted $F_{1\text{-}94}$ mutant proteins were overexpressed and purified following the same protocol developed for wild type (WT) selenomethionyl subunit $F_{1\text{-}94}$ (Figure 3.2.12).

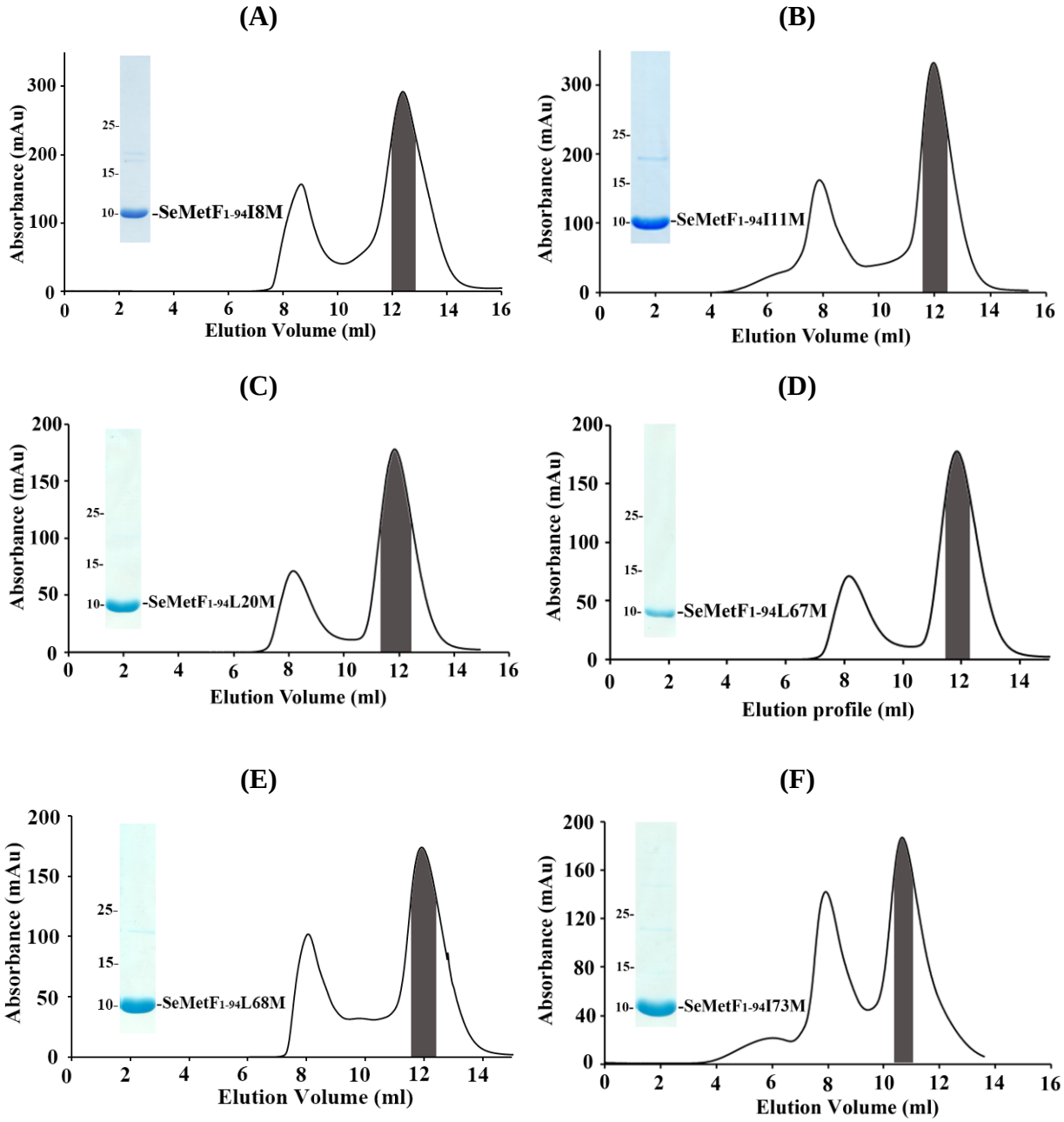


Figure 3.2.12: Superdex 75 elution profile of $F_{1\text{-}94}$ mutants; *(A)* SeMet $F_{1\text{-}94}$I8M, *(B)* SeMet $F_{1\text{-}94}$I11M, *(C)* SeMet $F_{1\text{-}94}$L20M, *(D)* SeMet $F_{1\text{-}94}$L67M, *(E)* SeMet $F_{1\text{-}94}$L68M, *(F)* SeMet $F_{1\text{-}94}$I73M. Insert shows respective SDS-PAGE gel.

Crystallization drops of mutants were setup in the optimized condition of native $F_{1\text{-}94}$ crystal, resulting either in microcrystals or precipitation (Figure 3.2.13).

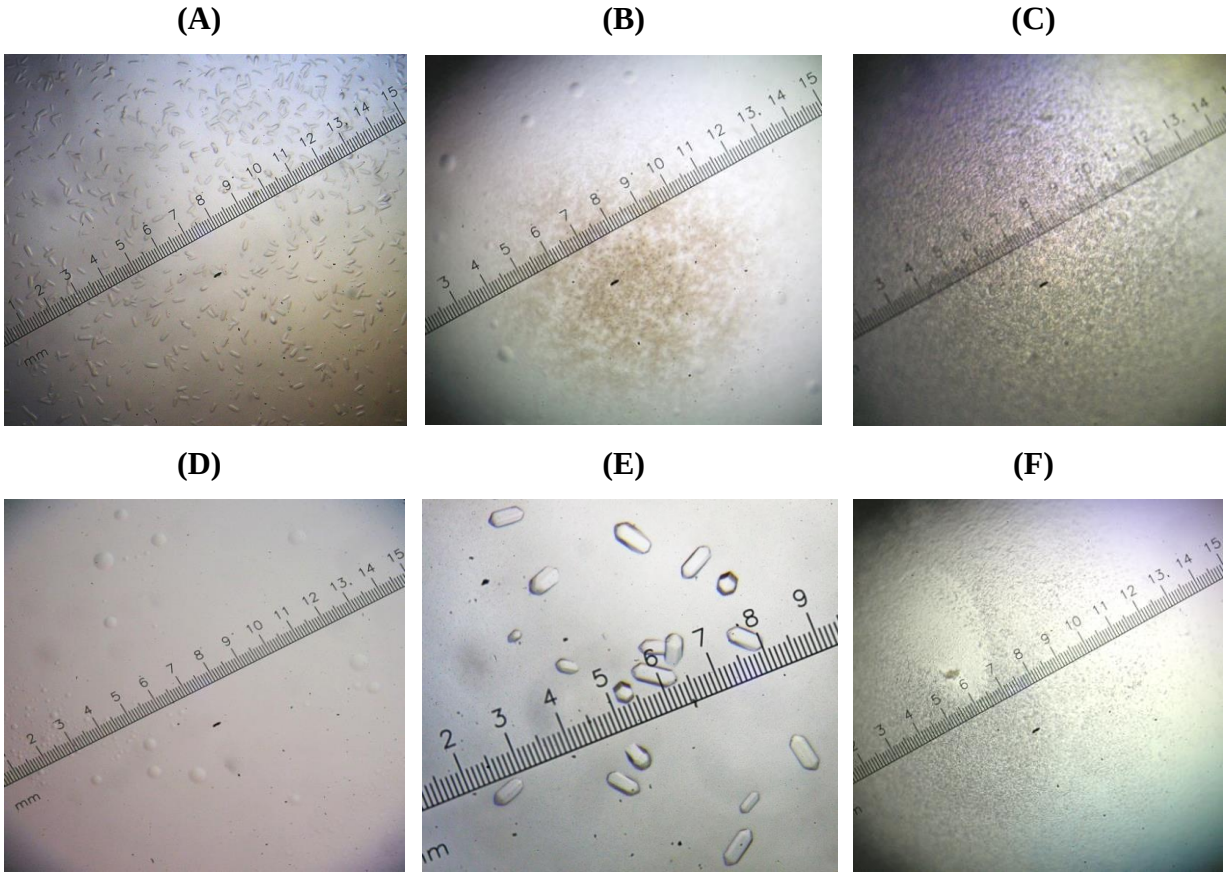


Figure 3.2.13 : Crystal drops of selenomethionyl F_{1-94} mutants. All drops were set up in the condition with 30% (v/v) PEG 4000, 0.10 M $MgCl_2$, 6xH_2O, 0.10 M Tris/HCl, pH 8.8 and 10% glycerol *(A)* Microcrystals of SeMet F_{1-94}I8M. *(B)* microcrystalline precipitation of SeMet F_{1-94}I11M. *(C)* Precipitation of SeMet F_{1-94}L20M. *(D)* Clear drop of SeMet F_{1-94}L67M. *(E)* Microcrystal of SeMet F_{1-94}L68M. *(F)* Microcrystalline precipitation of SeMet F_{1-94}I73M.

Secondary structure prediction shows that I69 is positioned near to the end of beta strand (Figure 3.2.14). Mutation in this position might have less influence on the tertiary structure of F_{1-94}. Therefore, F_{1-94}I69M mutant was generated, cloned and transformed into BL21 (DE3) for overexpression. SeMetF_{1-94}I69M mutant was purified and crystal drops were set up in the optimized condition of native F_{1-94}, yielding crystals, which were further optimized.

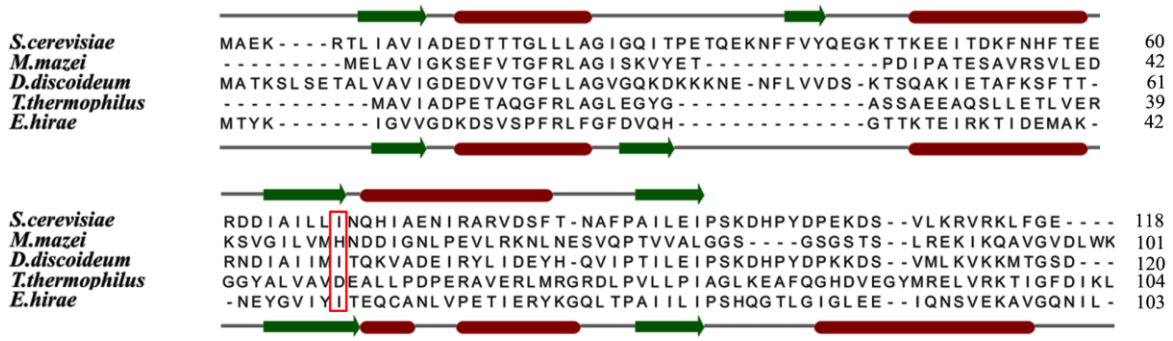


Figure 3.2.14: Multiple sequence alignment of subunit F from eukaryotic V-ATPases. Secondary structure elements of the eukaryotic F_{1-94} from *S. cerevisiae* and the A-ATP synthase subunit F from *E. hirae* are shown above and below, respectively. Red boxes indicate the position of isoleucine which mutated to methionine

3.2.5 Production, purification and crystallization of selenomethionyl F_{1-94}I69M mutant

Selenomethionyl form of F_{1-94}I69M was produced as in wild type selenomethionyl F_{1-94}, which yielded pure protein (Figure 3.2.15A). Crystal drops were setup in the optimized condition of native F_{1-94} and produced crystals which were smaller in size (Figure 3.2.15B) and thus needed further optimization.

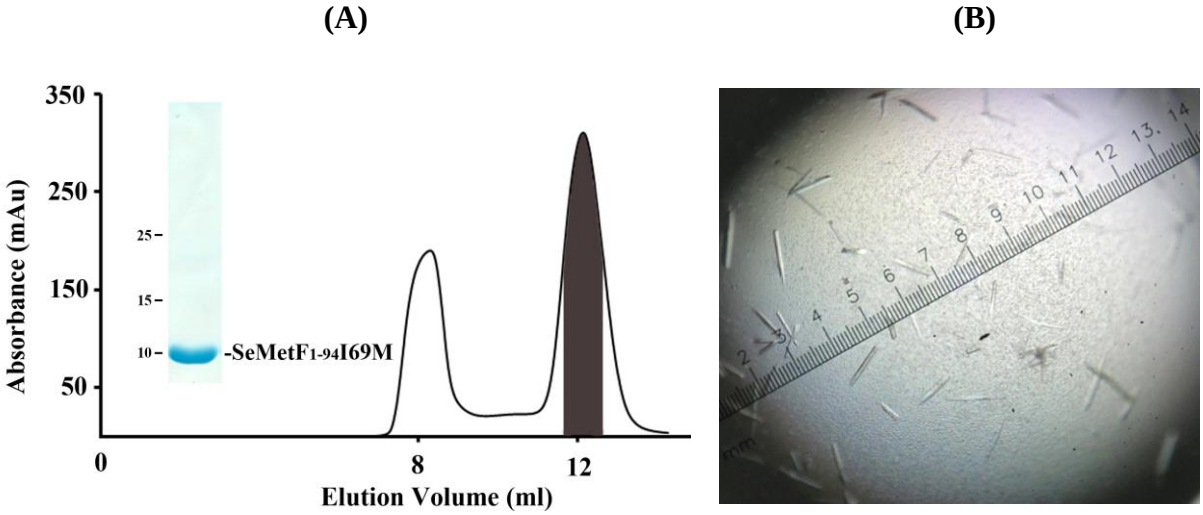


Figure 3.2.15: *(A)* Purification chromatogram of SeMetF_{1-94}I69M protein. After purification by Ni^{2+}-NTA chromatography, the ptotein was loaded onto the gel-filtration S75 column. Insert shows SDS-PAGE gel of protein purified after gel filtration chromatography. *(B)* Crystals of selenomethionyl F_{1-94}I69M mutant appeared in the condition with 30% (v/v) PEG 4000, 0.10 M $MgCl_2$, 6xH_2O, 0.10 M Tris/HCl (pH 8.8) and 10% glycerol.

Different grid screens were used to optimize the crystallization condition. At the gel filtration step (Superdex^{TM} 75 column), only 20-30% of apical peak fractions (Figure 3.2.16) were pooled together to get highly pure protein leading to better diffraction quality crystals. Bigger sized crystals (0.3 mm x 0.07 mm x 0.05 mm) were obtained at 291 K in the condition with 30% PEG 4000, 0.05 M $MgCl_2$, 0.10 M Tris-HCl pH 8.8, 10% glycerol, where the concentration of protein was 5 mg/ml. During harvesting, crystals cracked leading to poor diffraction when observed in the in-house X-ray machine (Figure 3.2.17A-B).

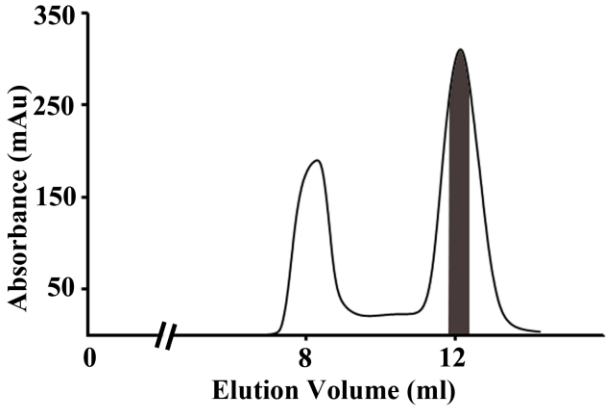


Figure 3.2.16: Gel filtration chromatogram of SeMetF_{1-94}I69M mutant protein. Only 20-30% peak fractions pooled together to get pure homogeneous protein.

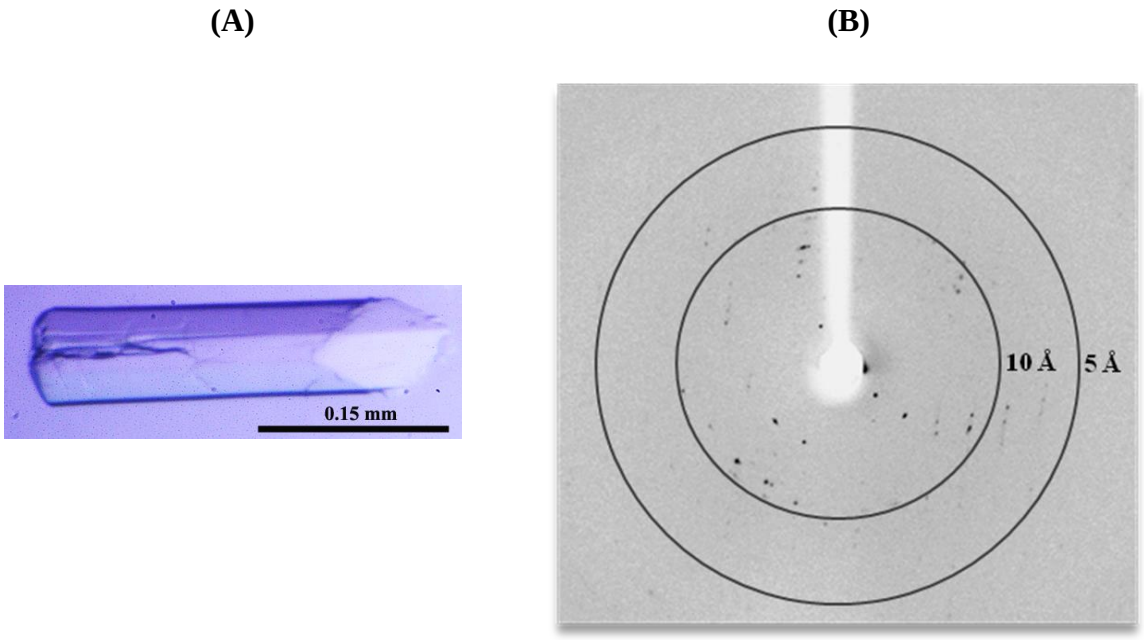


Figure 3.2.17: *(A)* Crystal of SeMetF_{1-94}I69M protein which diffracted to 8 Å due to the crack. *(B)* Diffraction pattern of the cracked crystal. Circles showing resolution shell for 10 Å and 5.0 Å.

To avoid this problem, crystals were first washed in a stabilization solution containing 32% (v/v) PEG 4000, 0.10 M Tris/HCl, pH 8.8, 0.10 M $MgCl_2$, 6xH2O and 10% glycerol and subsequently transferred into cryo-solution containing 32% (v/v) PEG 4000, 0.10 M $MgCl_2$, 6xH2O, 0.10 M Tris/HCl pH 8.8 and 22% glycerol. These crystals were tested in the in-house machine first and good diffraction quality crystals were measured at the National Synchrotron Radiation Research Center (NSRRC, Hsinchu, Taiwan) for data collection.

3.2.6 Data collection, processing and structure determination

Fluorescence scanning of selenomethionyl $F_{1–94}$I69M mutant crystal showed signal (Figure 3.2.18A-B) at the selenium absorption edge which allows to determine the correct phase for structure determination. Crystals diffracted to a maximum resolution of 2.3 Å (Figure 3.2.18C). The crystals belong to $C222_1$ space group, with unit-cell parameters *a* = 47.22 Å, b = 160.83 Å, c = 102.74 Å. Four molecules are present in one asymmetric unit. The crystal contains 46.28% of solvent and the calculated Matthews coefficient (V_m) (134) was 2.29 $Å^3$ Da^{-1}. Summary of data collection and processing statistics of F_{1-94} and its SeMetF_{1-94}I69M mutant are summarized in Table 3.2.2.

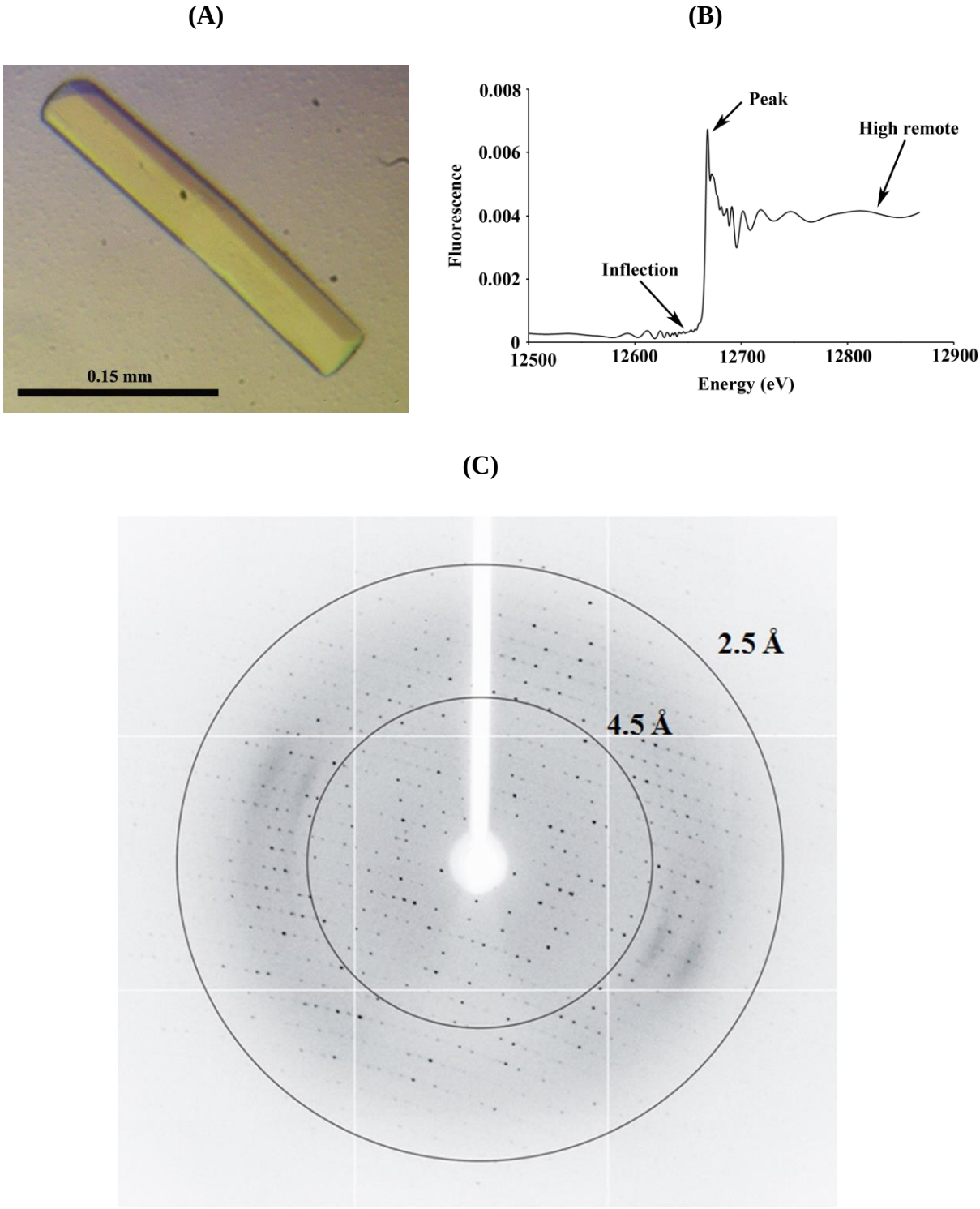


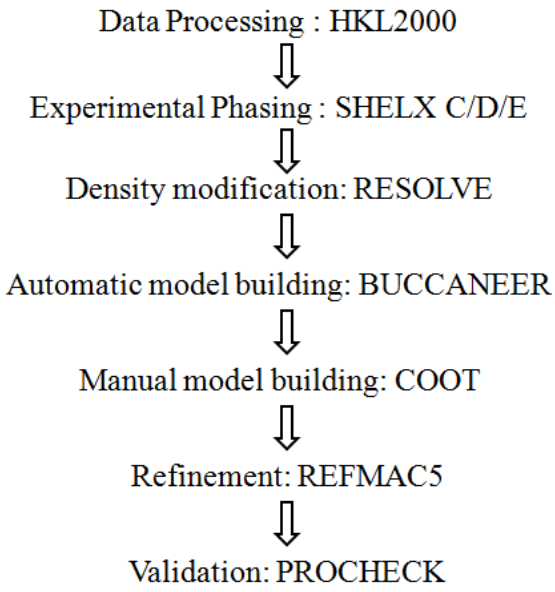
Figure 3.2.18. *(A)* $SeMetF_{1\text{-}94}I69M$ protein crystals in the condition of 30% (v/v) PEG 4000, 0.05 M $MgCl_2$, $6xH_2O$, 0.10 M Tris/HCl (pH 8.8) and 10% glycerol. Fluorescence scanning *(B)* and diffraction pattern *(C)* of $SeMetF_{1\text{-}94}I69M$ protein crystal (109). Circles showing resolution shell for 4.5 Å and 3 Å.

Table 3.2.2: Summary of data collection parameters and crystal characterization of SeMetF_{1-94}I69M mutant

	Peak	***Inflection***	***Remote***
Wavelength (Å)	0.978683	0.978836	0.963626
Space group	$C222_1$	$C222_1$	$C222_1$
Unit cell parameters (Å)			
a =	47.223	47.146	47.153
b =	160.833	160.312	160.334
c =	102.738	102.429	102.439
Molecules in asymmetric unit	4	4	4
Solvent content (%)	46.28	45.86	45.87
Resolution limits (Å)	50.0-2.45 (2.54-2.45)	50.0-2.33 (2.41-2.33)	50.0-2.29 (2.37-2.29)
No. of images	720	360	360
No. of reflections	33722	38421	40195
Unique reflections	14946	17096	17958
Multiplicity	12.7 (6.4)	6.5 (4.6)	6.4 (3.7)
Completeness (%)	99.4 (95.9)	99.5 (96.7)	98.8 (89.5)
R_{merge} (%)*	9.4 (58.6)	8.4 (43.4)	8.4 (45.9)
$\langle I//\sigma(I)\rangle$	23.4 (1.8)	18.8 (2.2)	17.1 (1.6)

*$R_{merge} = \Sigma\Sigma_i|I_h - I_{hi}|/\Sigma\Sigma_i I_h$, where I_h is the mean intensity for reflection h.

Phasing was performed by using three-wavelength MAD data sets. The structure was solved at 2.3 Å (R_{crys} = 0.15 and R_{free} = 0.21). Structure was fitted nicely into the electron density map which indicates high quality model of F_{1-94}I69M structure. Flow chart of structure determination, refinement and validation programs have been shown below.



The validation statistics for the final structure is summarized below.

Table 3.2.3: Summary of refinement statistic of SeMetF_{1-94}I69M mutant crystal structure

R factor (%)	15.2
R free (%)	21.2
Ramachandran statistics	
Residues in most favored (%) region	89.9
Residues in additionally allowed (%) region	10.1
Residues in generously (%) region	0.0
Residues in disallowed (%) region	0.0
R.M.S. Deviations	
Bond length (Å)	0.015
Bond angle (^{o})	1.7
Mean atomic B values	39.03
Number of water molecule	245

Four molecules are present in one asymmetric unit (Figure 3.2.19A). There was no clear electron density of the N-terminal region which could be due to the disordered His-tag region. The four β-strands (β1-β4) are oriented parallel to each other and are connected by four α-helices (α1-α4) (Figure 3.2.19B). The four β-strands (β1-β4) are formed by amino acid residues spanning from L7-D13, F37-Y40, I64-M69 and A90-E93, respectively as well as the four α-helices (α1-α4) with residues ranging from E14-L22, K47-E59, Q71-N76 and R78-D82. The F_{1-94} crystal structure contains seven loops, including a polar and large loop

between α1-β2 with amino acid sequence of $_{26}$GQITPETQEK$_{35}$ (Figure 3.2.20A). The electron density at the loop regions is very clear (Figure 3.2.20B). Surface electrostatic potential shows that one side of the protein is hydrophobic in nature and the opposite side is composed of both acidic and basic residues (Figure 3.2.21). The hydrophobic site is comprised of the residues I8, A9, V10, I11 and A12 located on β1; I64, A65, I66, L67 and L68 on β3; A90, I91, L92 and I94 on β4; G19, L20, L21 and L22 on α1. The electrostatic surface potential of the β-sheet is highly hydrophobic due to the presence of hydrophobic amino acids (β1:$_{7}$LIAVIAD$_{13}$, β2:$_{37}$FFVY$_{40}$, β3:$_{64}$IAILLM$_{69}$, β4:$_{90}$AILE$_{93}$), while the helices are mainly dominated by hydrophilic residues. The longer α2 helix contains mostly hydrophilic residues ($_{47}$KEEITDKFNHFTE$_{59}$) while the shorter helices, α3 ($_{71}$QHIAEN$_{76}$) and α4 ($_{78}$RARVD$_{82}$) contain hydrophilic and hydrophobic residues, respectively. In contrast, the first helix, α1 starts with a hydrophilic patch of five residues ($_{14}$EDTTT$_{18}$) followed by a hydrophobic patch containing four residues ($_{19}$GLLL$_{22}$). Sequence analysis of subunit F with other eukaryotic V-ATPases and related A-ATP synthases revealed two important loops present in the $F_{1\text{-}94}$ structure. One unique polar loop which is larger than other loop spanning between α1-β2 ($_{26}$GQITPETQEK$_{35}$) and a highly conserved loop with a $_{60}$ERDD$_{63}$ motif between α2-β3, which are present only in eukaryotic V-ATPases (Figure 3.2.20A).

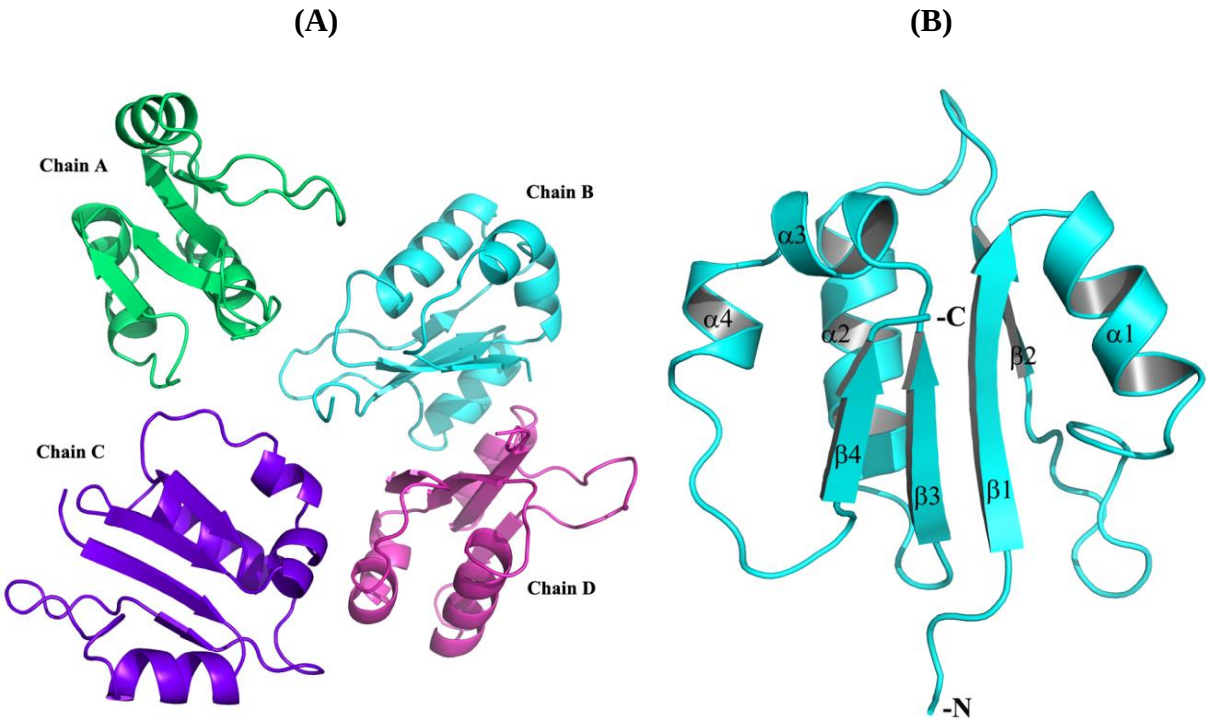
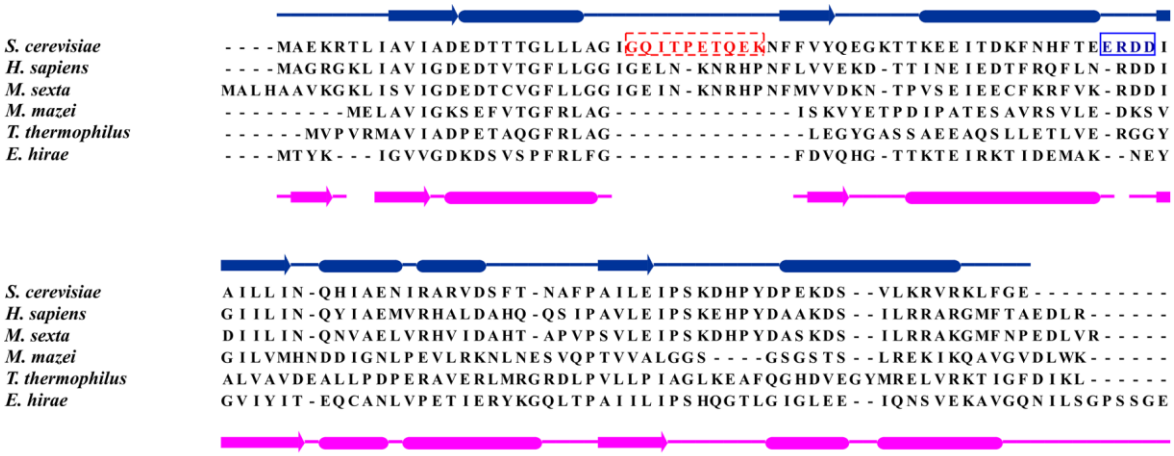


Figure 3.2.19: Crystal structure of *S. cerevisiae* subunit F, $F_{1\text{-}94}$. *(A)* Ribbon diagram of *S. cerevisiae* subunit F, $F_{1\text{-}94}$, showing four molecules in one asymmetric unit. *(B)* Ribbon structure of single molecule showing a Rossman fold labeled with alternating α-helices and β-strands.

(A)



(B)

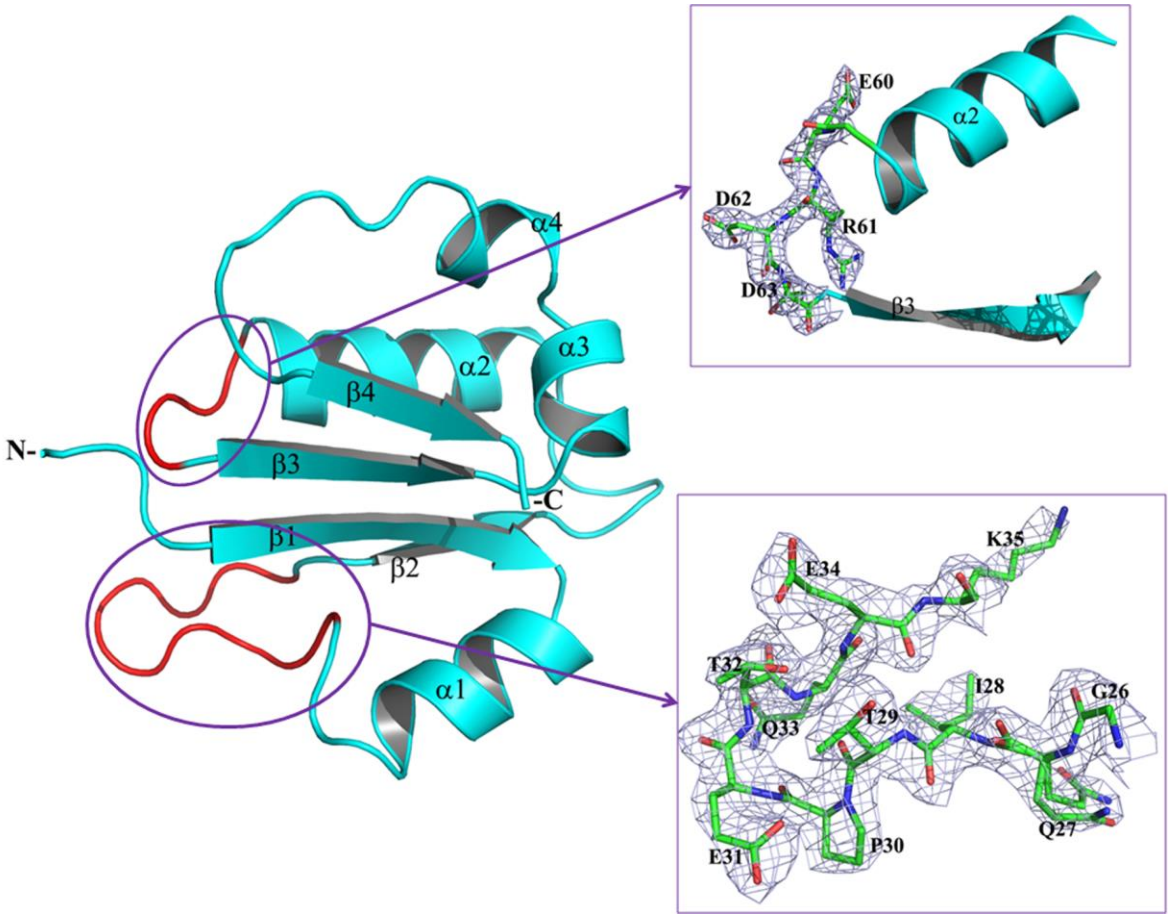


Figure 3.2.20: *(A)* Multiple sequence alignment of subunit F from eukaryotic V-ATPases and A-ATP synthases. The secondary structure elements of the eukaryotic subunit F from *S. cerevisiae* and the A-ATP synthase subunit F from *E. hirae* are shown in blue and magenta, respectively. *(B)* Two important loops are marked in red color (*left*). Electron density of these loops is shown (*right*).

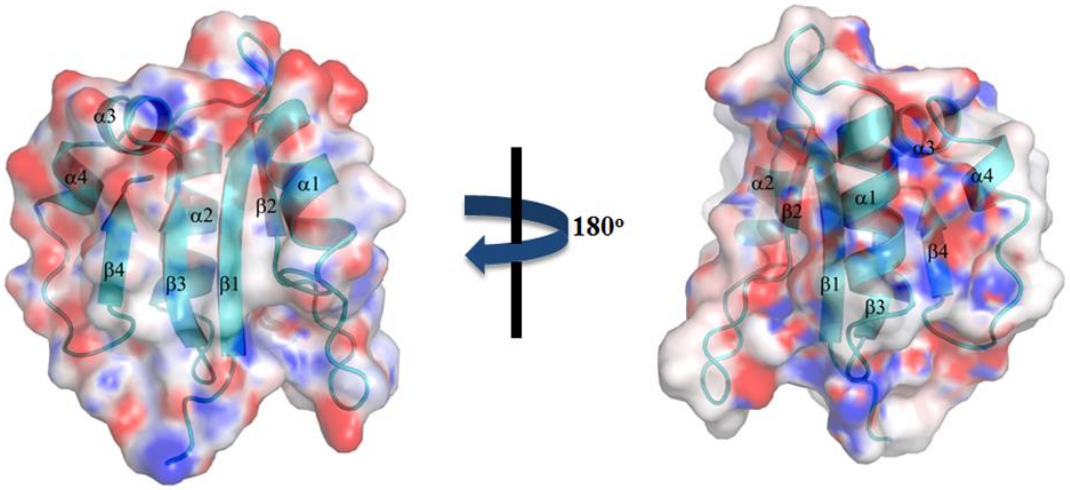
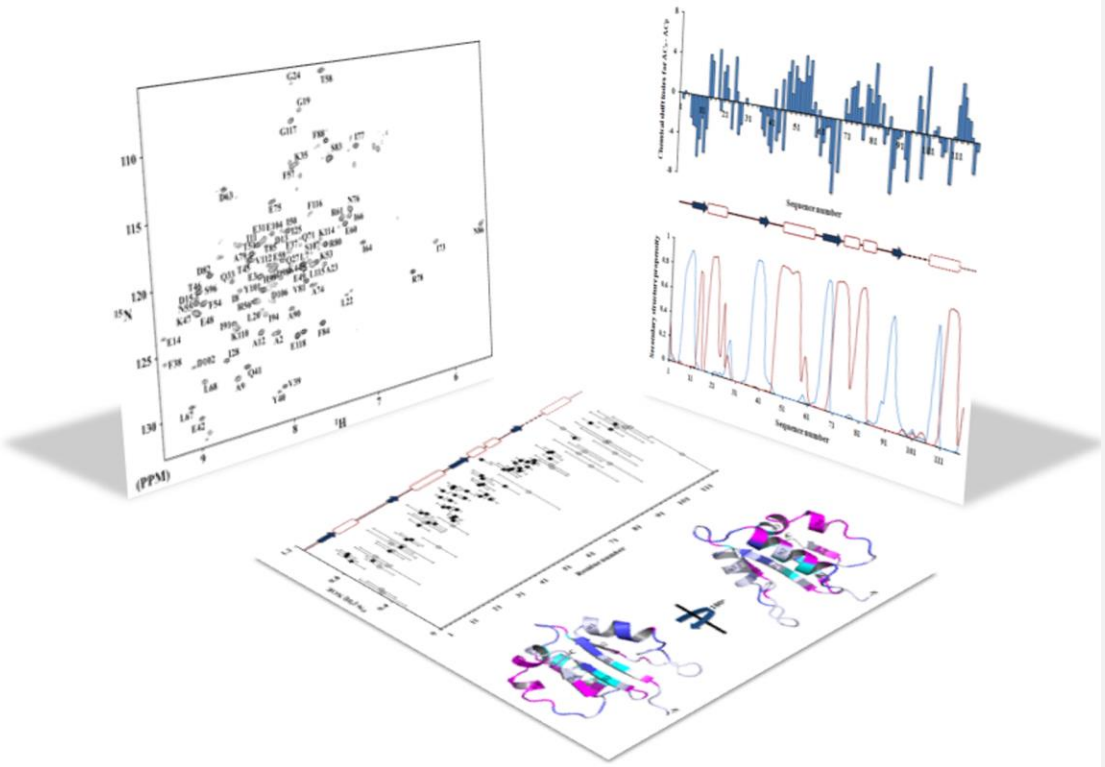


Figure 3.2.21: Surface electrostatic potential of $F_{1\text{-}94}$, with one side of protein showing hydrophobic (white) surface whereas opposite face is composed of basic (blue) and acidic (red) charges (modified from Basak *et al*; 2013) (140).

3.3 Structure and dynamics of entire *S. cerevisiae* subunit F analyzed by NMR spectroscopy



3.3.1 Introduction

Subunit F of V-ATPase plays a leading role in the coupling event as it is part of the central stalk. This subunit undergoes conformational changes and/or rearrangements during enzyme function and is proposed to interact with catalytic subunits A and B as well as the stalk subunits C, D, E and H (12). It has been shown that subunit H and F crosslinked in V_1 ATPase which does not occurring in assembled V_1V_o-ATPase (67).

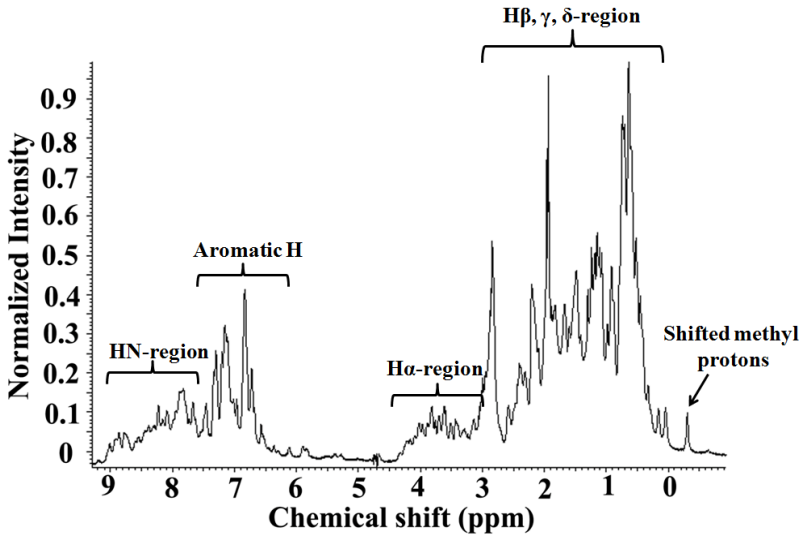


Figure 3.3.1: 1D NMR spectrum of subunit F. One dimensional NMR spectrum of subunit F recorded at 298 K in 25 mM phosphate buffer, pH 6.8, 200 mM NaCl and 5 mM EDTA on Bruker Avance 600 MHz machine.

In order to investigate the proposed interaction, one dimensional NMR spectrum of subunit F was recorded in 25 mM phosphate buffer (pH 6.8), which showed characteristic pattern of peaks and were dispersed in amide region (6 to 10 ppm), α-region (3.5 to 5 ppm), and uphill shifted methyl protons region (-0.5 to 1.0 ppm) (141) (Figure 3.3.1). The 2D ^{1}H-^{15}N HSQC experiments were done in 25 mM phosphate buffer at pH 6.8 by changing different parameters *e.g.* temperature, addition of mild detergent. The 2D ^{1}H-^{15}N HSQC spectra of subunit F at different optimization conditions are shown in figure 3.3.2A-B and figure 3.3.3A-B. Thereafter 3D NMR experiments were performed in the condition of 25 mM phosphate buffer, pH 6.8, 200 mM NaCl and 5 mM EDTA at 25^{o} C, but during acquisition, the protein precipitated. Further optimization was needed to stabilize subunit F in order to perform 3D NMR experiments.

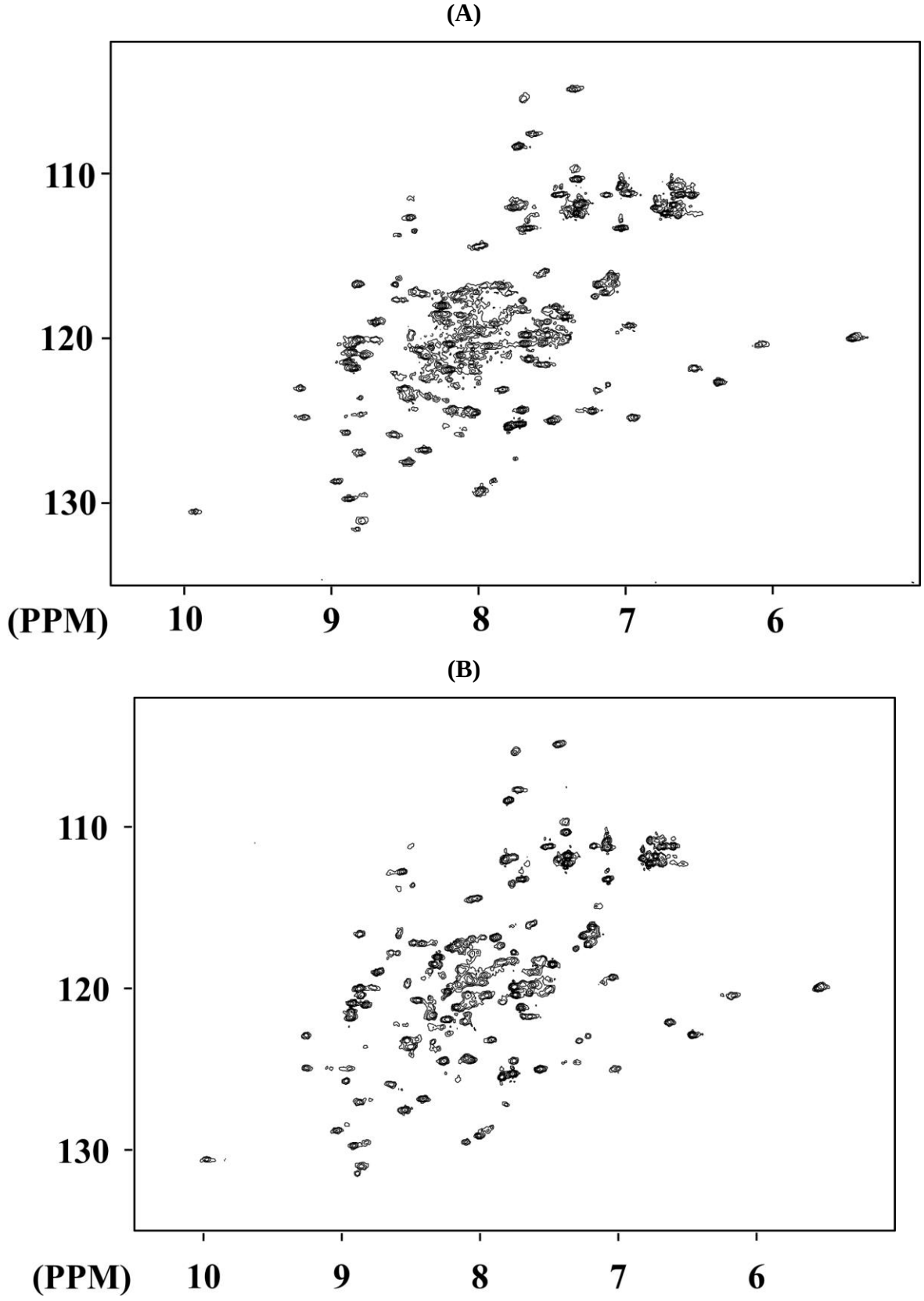


Figure 3.3.2: ^{1}H-^{15}N HSQC spectrum of subunit F acquired at different temperatures. *(A) 2D* HSQC NMR spectrum of subunit F recorded in 25 mM phosphate buffer, pH 6.8, 200 mM NaCl and 5 mM EDTA on Bruker Avance 600 MHz machine at 288 K *(A)* and 298 K *(B)*.

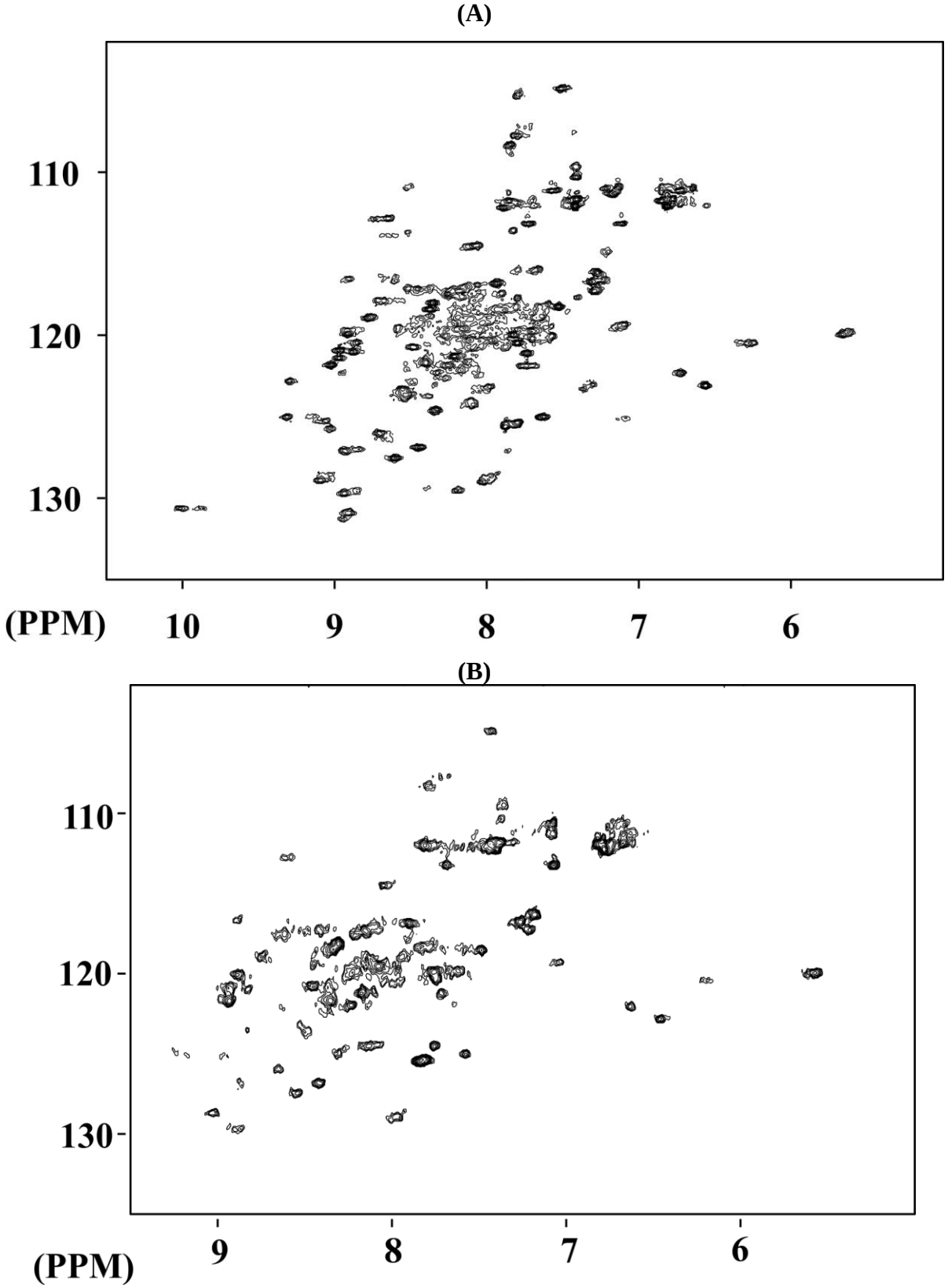


Figure 3.3.3: Optimization of parameters for acquisition of ^{1}H-^{15}N HSQC spectrum of subunit F. *(A) 2D* HSQC NMR spectrum of subunit F obtained in 25 mM phosphate buffer, pH 6.8, 200 mM NaCl and 5 mM EDTA at 303 K. *(B)*Representative *2D* ^{1}H-^{15}N HSQC NMR spectrum of subunit F in the presence of 2% dodecyl maltoside (DDM).

The protein buffer was changed to 50 mM Tris-HCl at pH 7.5, 200 mM NaCl, 5 mM EDTA. Then the 3D NMR experiments were conducted at 20° C on a 700 MHz NMR spectrometer where the protein was stable for entire acquisition period. At final optimized condition, subunit F produced a better spectrum which was used to assign amino acids (Figure 3.3.4).

3.3.2 Resonance assignments of subunit F of the *S. cerevisiae* V_1V_O ATPase

In order to determine the interaction of subunit F with other subunits and to study dynamics in solution, resonance assignment of this subunit was necessary. The sequential resonance assignments of the backbone amide groups are labeled in a 2D $^{1}H/^{15}N$-HSQC spectrum (Figure 3.3.4), indicating cross-peaks of backbone ^{1}H and ^{15}N of individual residues. The folding of subunit F was confirmed by good dispersion of amide peaks as well as up-field shifted methyls. Combination of triple-resonance CBCA(CO)NH, HNCACB and HNCO spectra were used to assign the backbone HN, ^{15}N, $^{13}C^{\alpha}$, $13C^{\beta}$ and 13C′ resonances (Figure 3.3.4). A total 84 residues were identified from the HSQC peaks and most of the unassigned residues are present in the flexible loops of entire domain (Figure 3.3.5). Flexible parts of the structure were not observed in the HSQC spectrum, which could be due to the fast amide solvent exchange or chemical exchange in these parts. SAXS data also indicated that the protein is monomeric in solution mentioned in section 3.1.4 (142) at the concentration used for NMR, thus it can be speculated that the broadness of the peaks in 2D HSQC is not due to oligomerization. Different types of line shapes of the peaks were observed in the HSQC spectrum. There are single intense peaks which are mostly from the N-terminal region. By comparison, the C-terminal part has less intense peaks indicating, that this region is more flexible in nature. There are some unassigned less intense peaks which are most probably coming from a minor conformational state. This difference in intensity of HSQC peaks is

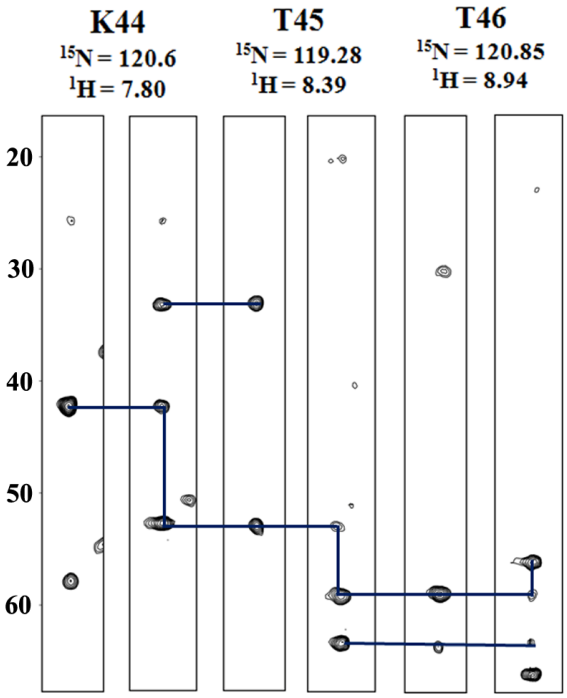


Figure 3.3.4: An example of sequential assignment of three residues (K44, T45 and T46) by CBCA(CO)NH and HNCACB experiments in case of subunit F of *S. cerevisiae* V-ATPase.

clear evidence that the protein undergoes conformational exchange on the microsecond to millisecond time scale. Due to incomplete sequential assignment, some sharp peaks could not be assigned.

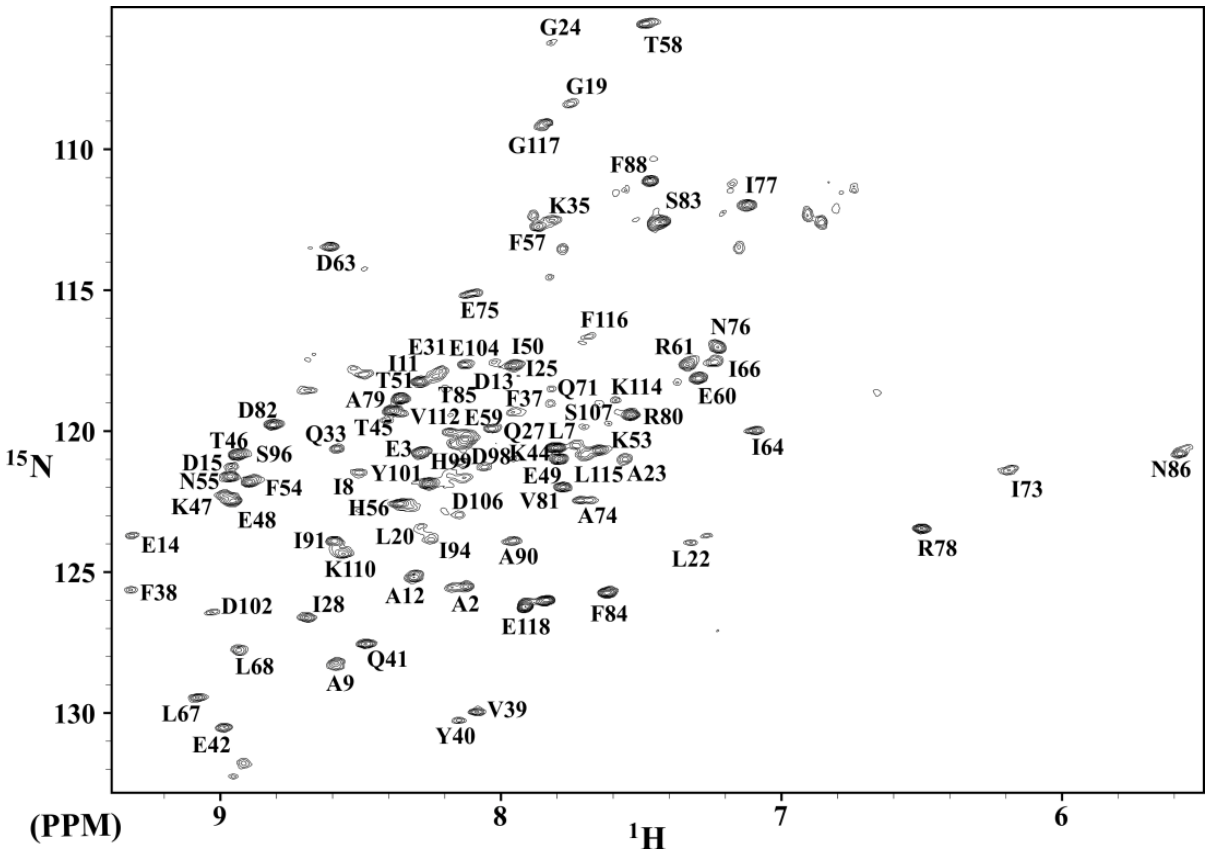


Figure 3.3.5: ^{15}N-^{1}H HSQC spectrum of the entire subunit F of V-ATPase from *S. cerevisiae*. The NMR data were acquired at 20° C on a 700 MHz NMR spectrometer in 50 mM Tris-HCl buffer at pH 7.5, 200 mM NaCl, 5 mM EDTA (140).

3.3.3 Characterization of full length subunit F by NMR spectroscopy

The NMR study on full length F subunit was performed together with Dr. J. Lim from our lab to characterize the protein in the solution. The sequential backbone assignment was carried out for 118-residue F subunit. 84 out of ~90 observable peaks (excluding proline and sidechains) were assigned (Figure 3.3.5). The secondary structure prediction based on Cα and Cβ chemical shifts was done which indicated, that there is an alternating arrangement of four β-strands and five α-helices (Figure 3.3.6A), which nicely correlates with the crystal structure of $F_{1\text{-}94}$ which is shown in chapter 3.2.6. The four β-strands are β1 (T6-D13), β2 (N36-Q41), β3 (R61-L68) and β4 (F88-I94), whereas the five α-helices are α1 (E14-L22), α2 (K47-T58), α3 (N70-N76), α4 (R78-F84) and α5 (K110-F116). These are consistent with the X-ray

crystal structure of $F_{1\text{-}94}$ and the NMR structure of α5-peptide ($F_{90\text{-}116}$) solved in the presence of trifluoroethanol (20) (Figure 3.3.6A). In addition, the propensity of helix and beta strand formation predicted by TALOS+ (34) matches the secondary structure prediction by NMR and also confirms the presence of C-terminal helix α5 in solution (Figure 3.3.6B).

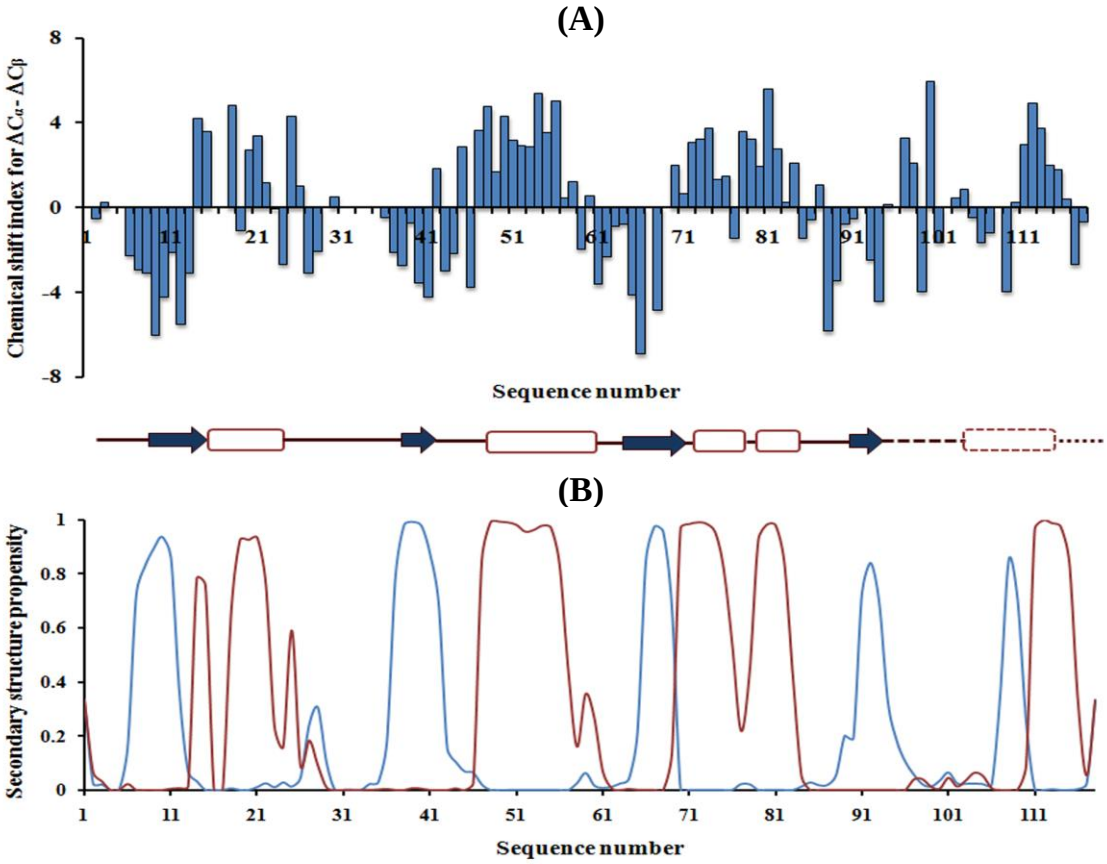


Figure 3.3.6: (*A)* Backbone characterization of full length subunit F. The secondary structure elements (β1 to β4), which correspond to the $F_{1\text{-}94}$ crystal structure and the NMR solution structure of $F_{90\text{-}116}$ peptide are shown below. Based on NMR assignment, contiguous positive and negative chemical shift index values correspond to α-helix and β-strand, respectively. *(B)* Secondary structure propensity values, predicted using TALOS+ (143), show the existence of a short α5 at the C-terminus. β- strands and α- helices are shown in brown and blue color, respectively (modified from Basak *et al.*, 2013) (140).

3.3.4 Dynamic studies of subunit F by NMR spectroscopy

To study the dynamics of C-terminal helical region as well as the major N-terminal region of subunit F of *S. cerevisiae* V-ATPase, ^{15}N-[^{1}H] heteronuclear NOE experiment was performed. This experiment provides residue-specific information about the conformational dynamics on pico- to nanosecond timescale. In general, high NOE values reflect rigidity in the tertiary structure. Based on the NOE values and a mean NOE value (error, σ) of 0.8 (0.12), strands β1 to β4, helices α2 to α4 helices are less flexible (NOE values > 0.8) than both N- and C-terminal helices α1 and α5 with similar average NOE values, 0.78 (0.14) and 0.76

(0.14), respectively (Figure 3.3.7A). Mapping onto the $F_{1\text{-}94}$ tertiary structure shows that the protein's buried core is more rigid than its outermost surface, with increasing flexibilities around helices α1 and also α5, probably conferred by the long loops adjacent to them (Figure 3.3.7B). Taken together, the high mean NOE value suggested a rigid F subunit with slightly less rigid helices α1 and α5 of the central stalk rotor.

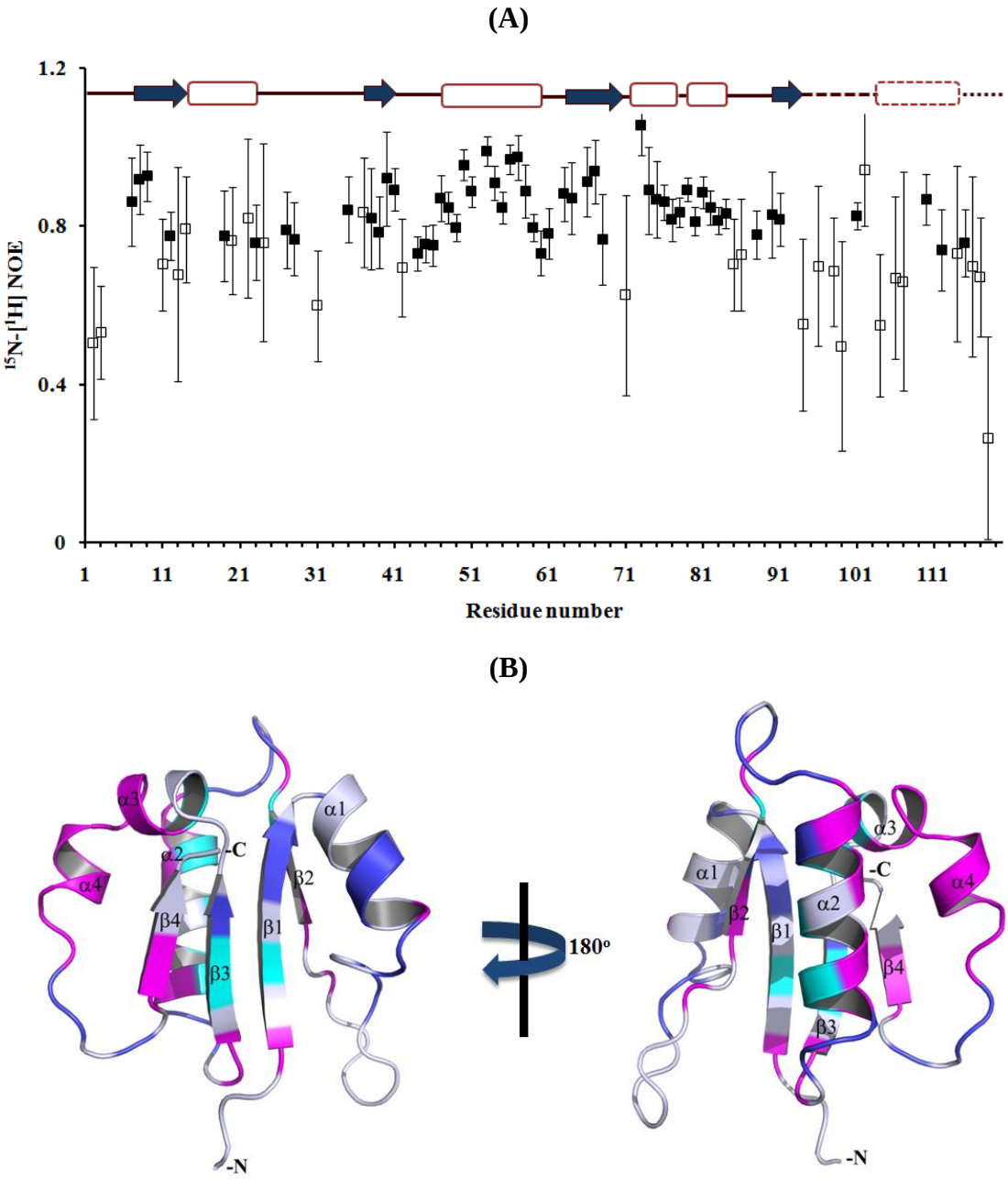
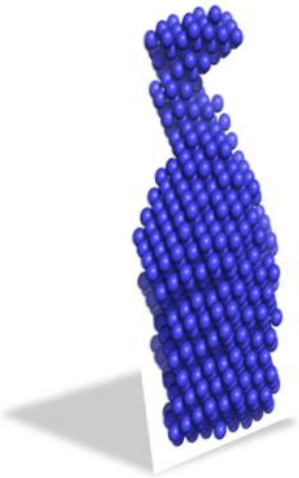
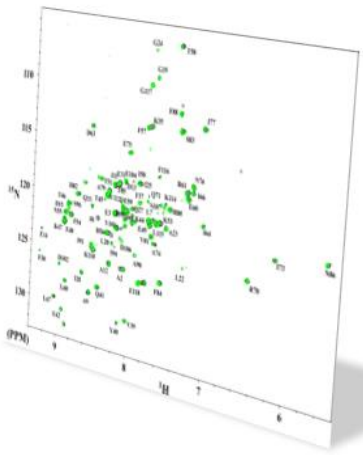


Figure 3.3.7: Conformational mobility of *S. cerevisiae* subunit F. *(A)* ^{15}N-[^{1}H] heteronuclear NOE with high values > 0.8 (rigid) and low values < 0.8 (flexible). NOE values with error > 15 % (white box) and < 15 % (black box) are also shown. *(B)* Using heteronuclear ^{15}N-[^{1}H] NOE values with error bar < 15 %, the mean value (0.80) and its error, (σ, 0.12) are calculated. *S. cerevisiae* $F_{1\text{-}94}$ is coloured blue (least rigid, > 0.68), magenta (> 0.80, mean) and cyan (most rigid, > 0.92, mean + 1σ), whereas NOE values with error > 15 % are coloured white (n.a.) (modified from Basak *et a.*, 2013) (140).

3.4 Interaction studies of subunit F with stalk subunits of V_1V_O ATPase from *S. cerevisiae* by NMR spectroscopy





3.4.1 Introduction

V-ATPases have a well known unique role of reversible dissociation that regulates the ATP hydrolysis in a definite manner. It was proven that subunit C plays a vital role during the reversible disassembly of V-ATPase which leads to structural alteration in the stalk part (83). The mechanism of this event is not well understood. Photo-activated cross-linking study showed that during reversible disassembly, there is a structural movement of subunit H which leads to a 10 Å cross-linked product with subunit F (67) by using maleimidobenzophenone (MBP) reagent. They reported that subunit F and the C-terminal part of H come closer during disassembly. In order to investigate the interaction, subunit C, H, $H_{8\text{-}354}$ and $H_{354\text{-}478}$ were constructed. Subunit *d* is part of stalk region which belongs to V_O domain and is the least characterized protein compared to other V-ATPase subunits. The function of this subunit is not yet characterized in eukaryotic system. Here, NMR experiment was conducted to analyse the probable interaction of stalk subunits based on the reported data at molecular level.

3.4.2 Cloning and production of subunit H, $H_{8\text{-}354}$, $H_{354\text{-}487}$ and C

The genes encoding subunit H**,** $H_{8\text{-}354}$, and $H_{354\text{-}487}$ were amplified using *S. cerevisiae* genomic DNA as a template for PCR amplification (Figure 3.4.1A-C). The amplified PCR products were digested with *NcoI* and *SacI* followed by the ligation with pET9d1-His_6 vector (86). Now the gene incorporated vector was transformed into *E. coli* cells (BL21 (DE3)) and subsequently grown on LB agar-plates which contain 30 μg/ml kanamycin as a antibiotic marker. Recombinant protein expression was induced by supplementing IPTG to a final concentration of 1 mM. The SDS-PAGE gel confirmed the induction of individual protein and is shown in figure 3.4.2A-D. Cloning and induction of subunit C was done according to the protocol mentioned previously (73).

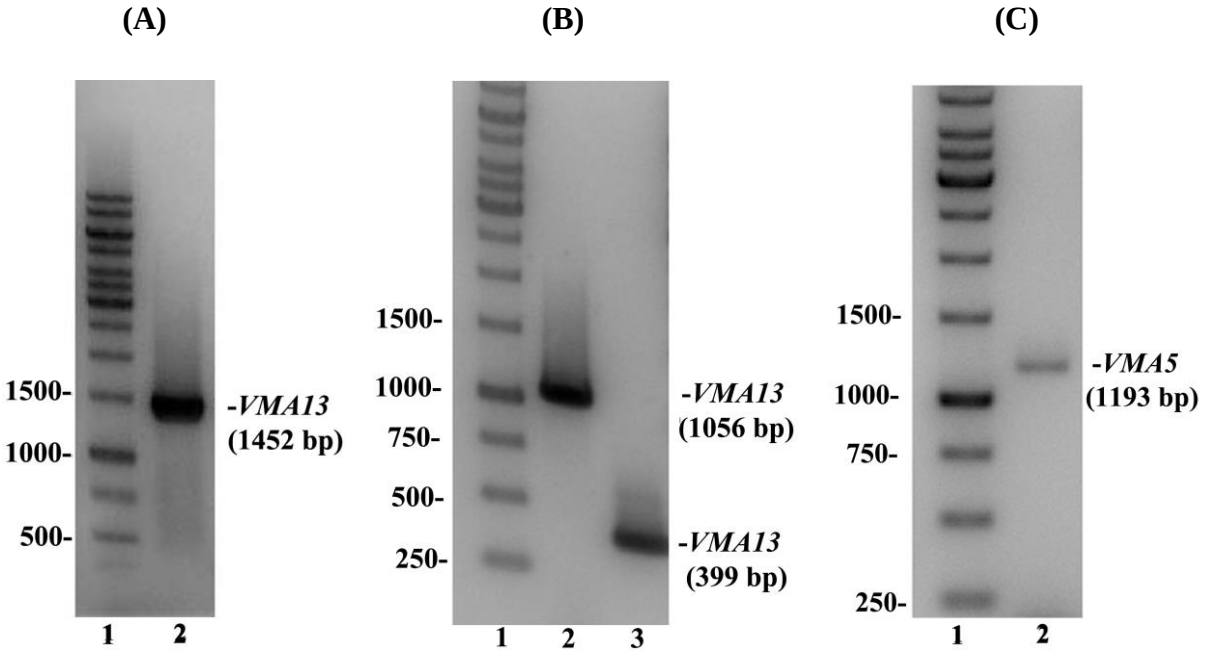


Figure 3.4.1: Amplification of gene by PCR cloning method. *(A)* PCR product of *VMA13* (1452 bp) which encodes subunit H. *(B)* PCR products of *VMA13* (1056 bp, *lane 1*) and *VMA13* (399 bp, *lane 2*) encode proteins $H_{8\text{-}354}$ and $H_{354\text{-}478}$, respectively. *(C)* PCR product of *VMA5* which encodes subunit C. DNA sample of 5 µl was loaded on an agarose gel (1%). *Lane 1* indicates the 500 bp marker and *lane 2* and *3* shows the amplified products.

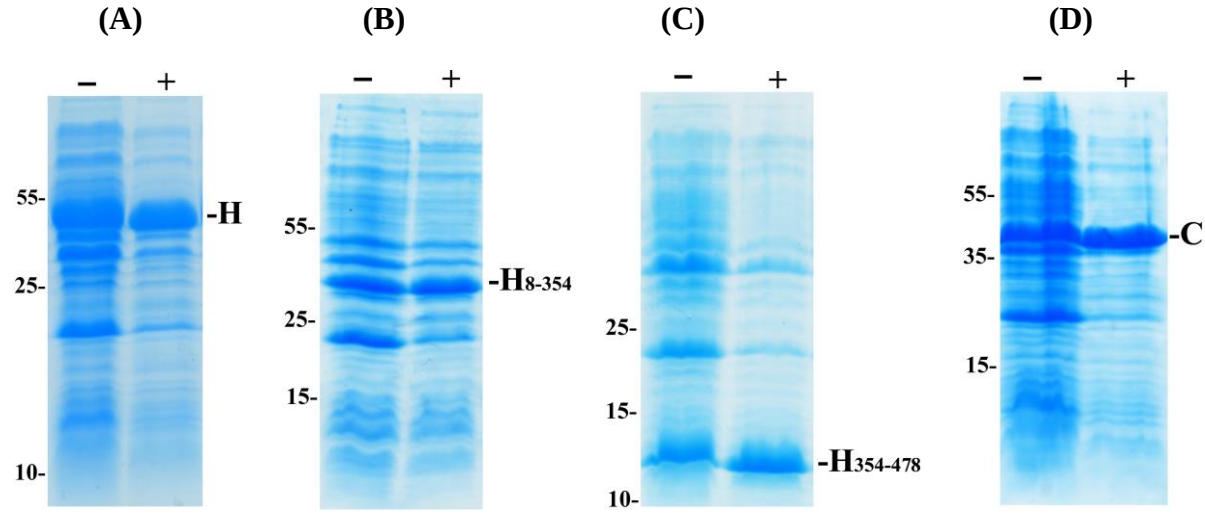


Figure 3.4.2: SDS-PAGE showing the expression of recombinant subunit H *(A)*, $H_{8\text{-}354}$ *(B)*, $H_{354\text{-}478}$ *(C)* and C *(D)* Proteins were induced IPTG to a final concentration of 1 mM. The + and – signs indicates the presence and absence of IPTG, respectively.

3.4.3 Purification of subunit H and $H_{354\text{-}478}$ from *S. cerevisiae* V-ATPase

Cells containing the His_6 tagged subunit H and $H_{354\text{-}487}$ were sonicated separately, followed by centrifugation to obtain the supernatant which was then incubate with Ni^{2+}-NTA beads. Bound protein was eluted by imidazole gradient (0-400 mM). Fractions between the

imidazole concentrations of 75 mM to 200 mM were pooled together and passed through the size exclusion column (Superdex75™ 10/300 GL), with buffer 50 mM Tris/HCl, pH 7.5, 150 mM NaCl and 1mM DTT. Protein fractions were pooled together, concentrated and 1 µl of the respective concentrated protein sample was loaded onto the SDS-PAGE gel to assess its purity (insert; on Figure 3.4.3A-B).

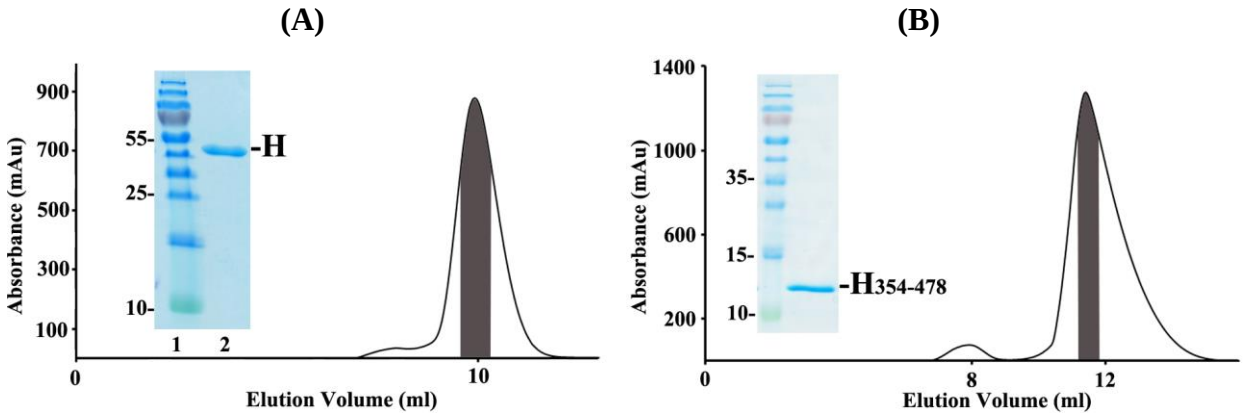


Figure 3.4.3: Superdex™ S75 purification chromatogram of subunit H *(A)* and $H_{354-478}$ *(B)* of V-ATPase from *S. cerevisiae*. Insert showing SDS-PAGE gel where *lane I* and *2* represent protein marker and final protein product from S75, respectively.

3.4.4 Purification of subunit C and *d* from *S. cerevisiae* V-ATPase

Purification of subunit C and *d* were done according to the protocol described previously (73,126).

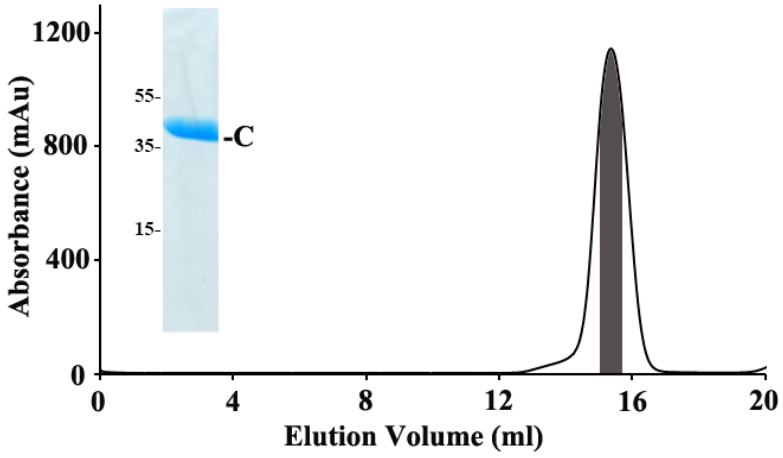


Figure 3.4.4: Superdex™ S200 HR 10/30 profile of subunit C of V-ATPase from *S. cerevisiae*. Insert showing SDS-PAGE gel of purified subunit C.

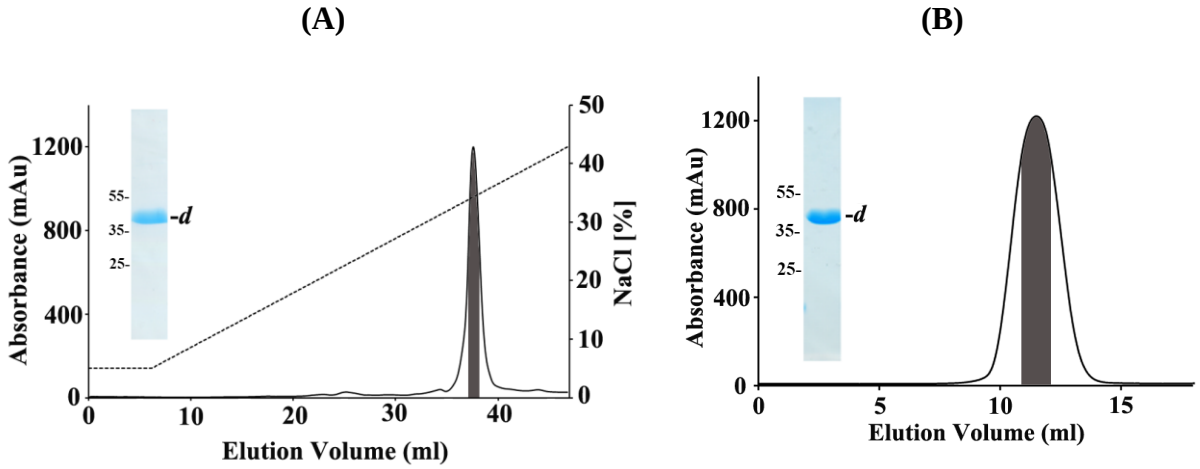


Figure 3.4.5: Purification chromatogram of subunit *d* of V-ATPase from *S. cerevisiae*. After purification by affinity chromatography, ptotein (subunit *d*) was loaded onto ion exchange RESOURCE™ Q column (6 ml) *(A)* subsequently onto the gel-filtration Superdex™S75 10/300 GL column *(B)* to get pure final product. In figure A the insert is showing SDS-PAGE gel after RESOURCE™ Q purification, whereas in figure B the insert represents a protein profile of subunit *d* after gel filtration chromatography.

3.4.5 NMR titration of F subunit against subunits H, $H_{354-478}$, C and *d*

To investigate the probable interactions in solution, NMR titration experiment was performed. As shown in figures 3.4.6A, peak shift or peak disappearance was not observed indicating the absence of interaction in solution even at an excess molar ratio of F:H of 1:5. The same phenomenon was observed when subunit F was titrated with $H_{354-478}$ (Figure 3.4.6B). ^{15}N-labeled subunit F was also titrated with subunit C and subunit *d* at molar ratio of F: C or *d* of 1:3, resulting in no observable peak shift or change (Figure 3.4.7A-B).

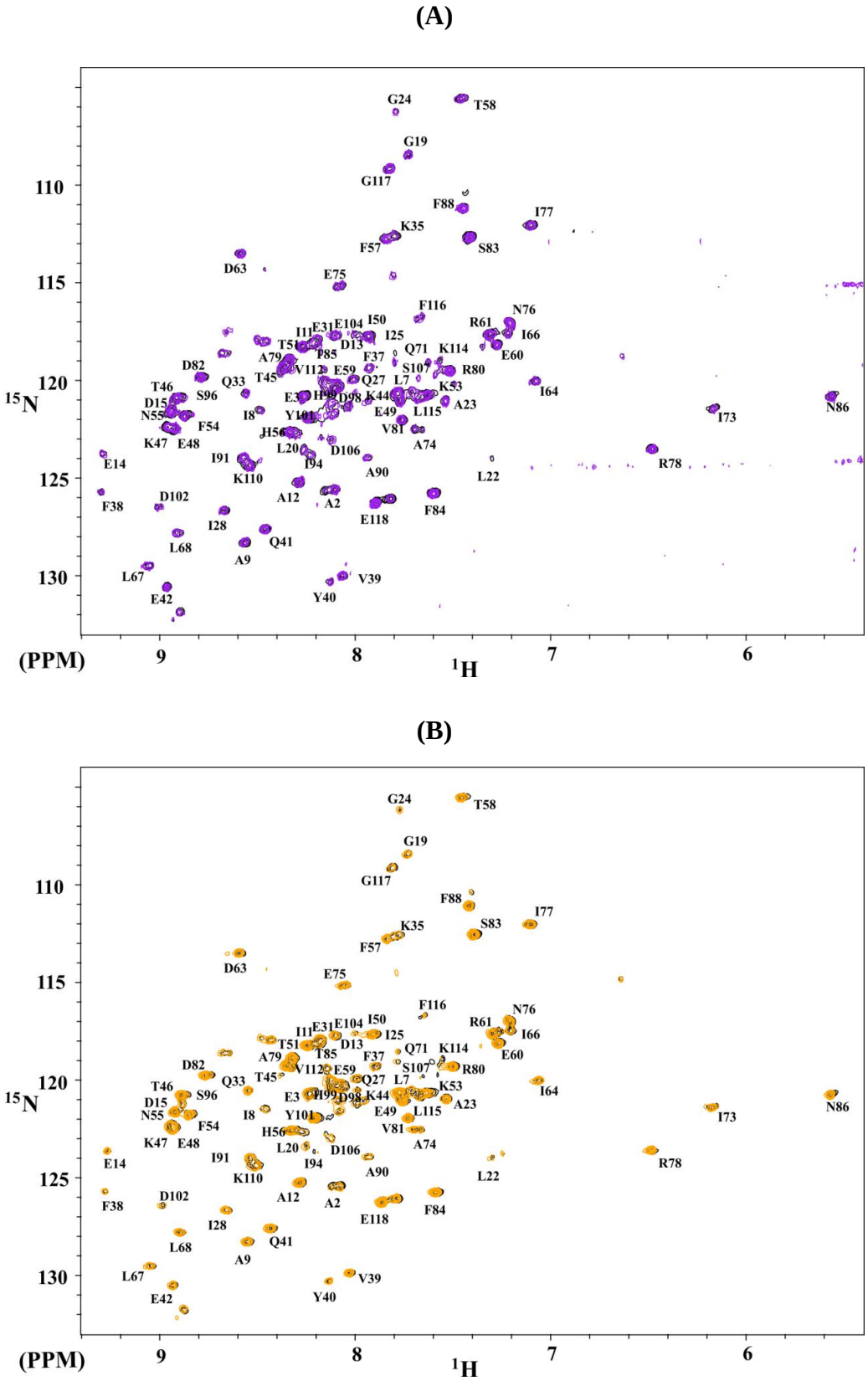


Figure 3.4.6: ^{15}N-^{1}H HSQC spectrum of subunit F titrated against subunit H *(A)* and $H_{354\text{-}478}$ *(B)* of V-ATPase from *S. cerevisiae*. The final molar ratio of subunit F (black): subunit H (purple)/$H_{354\text{-}478}$ (orange) is 1:5/3.

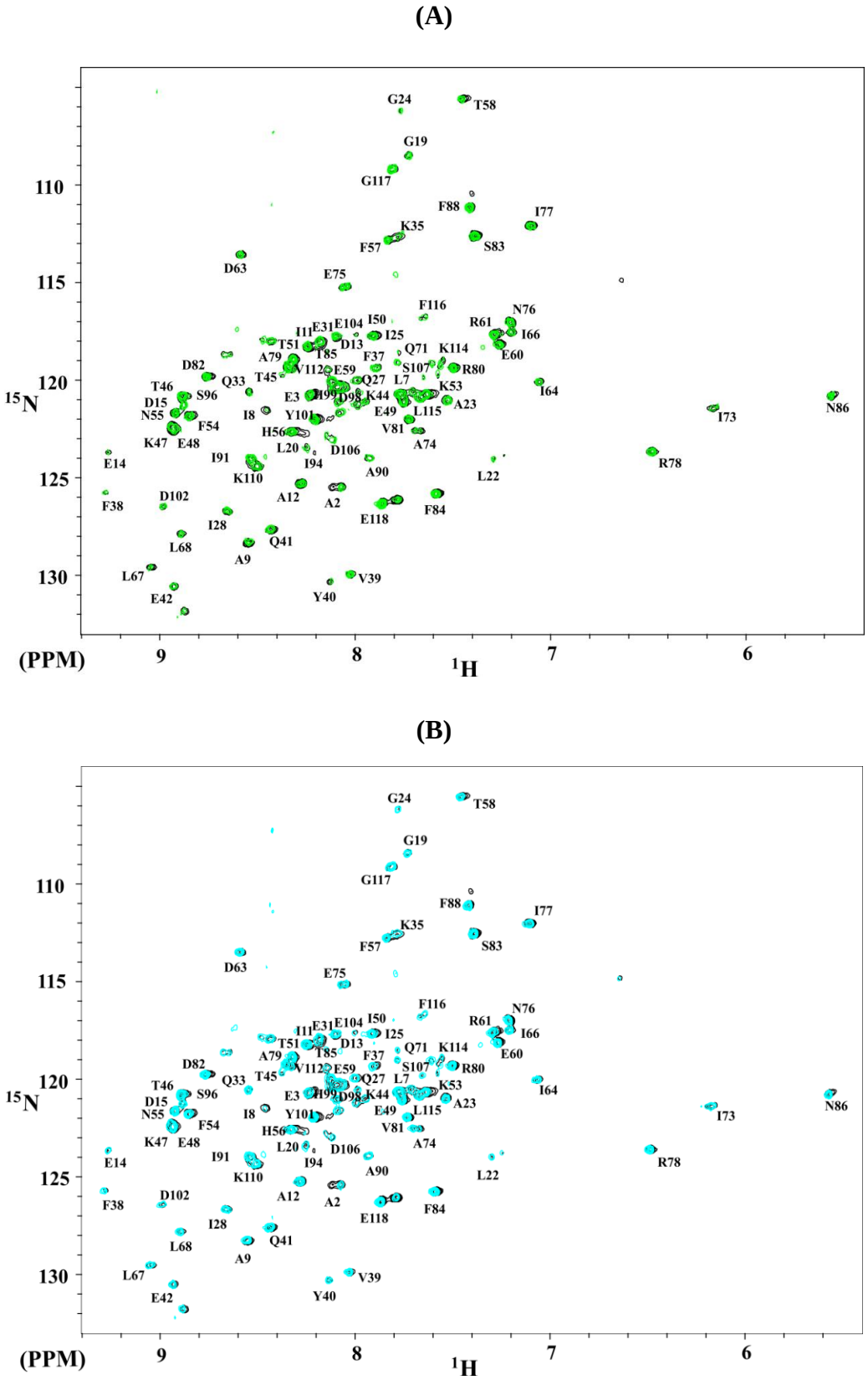
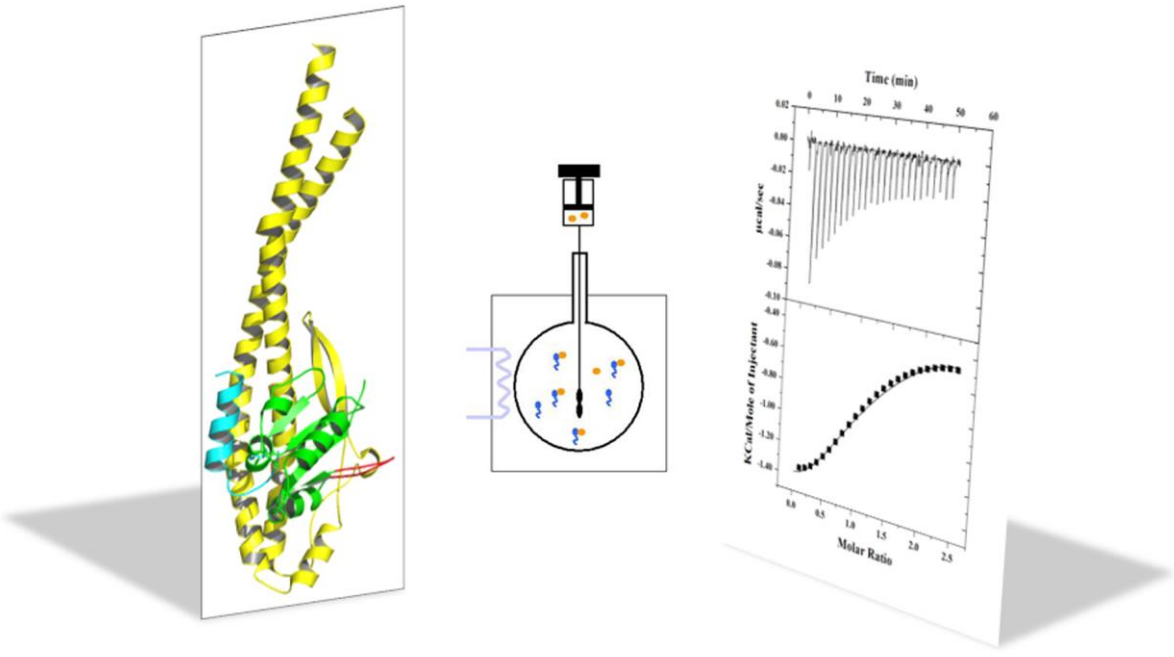


Figure 3.4.7: ^{15}N-^{1}H HSQC spectrum of subunit F titrated against subunit C and *d* of V-ATPase from *S. cerevisiae*. Molar ratio of subunit F (black): C (green)/*d* (cyan) is 1:3.

3.5 Interaction studies of subunit F and the DF-heterodimer with stalk subunits of the *S. cerevisiae* V_1V_O ATPase by ITC



3.5.1 Introduction

From the previous chapter it is clear that subunit F is not interacting with the stalk subunits *d*, C and H. Subunit D is also a part of central stalk, located together with subunit F. During reversible disassembly subunit D might undergo structural alteration together with subunit F. It has been reported that subunit D is not stable alone in A-type ATP synthases (144). In this chapter, I am focusing on the production of the DF heterodimer of *S. cerevisiase* V-ATPase and its interaction with stalk subunits by using ITC.

3.5.2 Cloning, production and purification of the DF complex

The *VMA7* and *VMA8* genes, encoding subunit D and subunit F, respectively, of V-ATPase from *S. cerevisiae* were amplified by using the genomic DNA from *S. cerevisiae* as the template for PCR amplification (Figure 3.5.1). Both genes were cloned into a pETDuet-1 expression vector (86). Both proteins were co-expressed under the condition specified in material and methods resulting in soluble 29 kDa and 14 kDa proteins, representing subunit D and subunit F of the *S. cerevisiae* V_1V_O ATPase, respectively. In the first purification step, affinity chromatography was used to bind His-tag proteins with Ni^{2+}- NTA beads, which was then eluted by an imidazole-gradient (25-200 mM) in buffer consisting of 50 mM TrisHCl (pH 8.5) and 200 mM NaCl. Both proteins were found at 75 to 200 mM imidazole concentrations and were pooled together and diluted with buffer to reach 50 mM salt concentration. Subsequently the fractions were applied to an ion exchange Resource Q column (6 ml) (Figure 3.5.2A), in order to isolate a pure DF-heterodimer. The main fraction was then passed through a gel filtration Superdex 200^{TM} HR 10/30 column (Amersham Pharmacia Biotech), to get pure and homogeneous protein (Figure 3.5.2B). SDS-PAGE gel was performed to analyze the purity of the isolated proteins (Figure 3.5.2A-B). SDS-PAGE gel was stained with Coomassie Brilliant Blue G250 which showed pure D and F subunits. Protein concentrations were determined by BCA (87) as

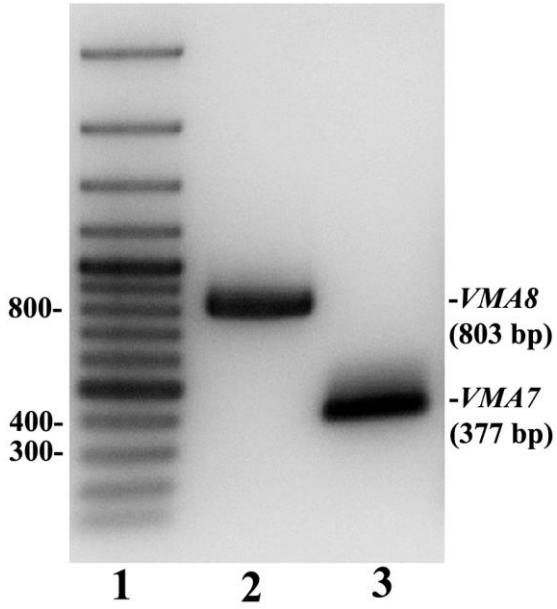


Figure 3.5.1: PCR amplification of *VMA7* and *VMA8* gene encoding subunit F and D, respectively. *Lane 1*, 2 *and 3* shows DNA ladder (100 bp, Fermentas), *VMA8* and and *VMA7*, respectively.

well as a nanodrop machine (BioSpec-nano, Shimadzu Scientific Instruments) to perform subsequent experiments.

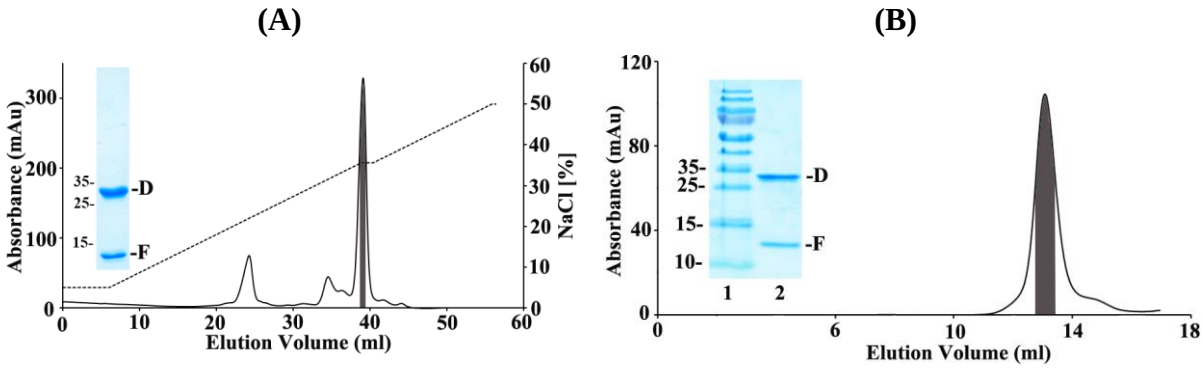


Figure 3.5.2: RESOURCE™ Q (A) and Superdex™ S200 *(B)* purification profile of DF complex of V-ATPase from *S. cerevisiae*. Insert in figure A shows the SDS-PAGE gel after RESOURCE™ Q purification. In figure B *lane I* and *2* represent a protein marker and profile of protein after size exclusion chromatography, respectively.

3.5.3 Isothermal titration calorimetry to study the interaction of subunit F and DF against subunits C, *d*, $H_{354\text{-}478}$

To assign the backbone amino acid of the complex structure it is still a challenging task due to the lack of information from the protein-protein interface. Isothermal titration calorimetry (ITC), a complementary method was used to investigate the probable interaction of DF complex as well as subunit F with subunit C, *d* and $H_{354\text{-}478}$. There is no observable interaction between subunit F and DF with subunit C and $H_{354\text{-}478}$ as shown in the ITC profile of subunit F and DF (Figure 3.5.3). The top panel of the ITC profile reveals a baseline after correction, whereas the bottom panel shows the heat release/mole profile of injected subunits. The titration of the DF complex with subunit *d* yielded an overall negative heat enthalpy, which indicates release of heat during reaction (Figure 3.5.4, *left*). In the case of subunit F, it did not show any heat changes upon addition of subunit *d* suggesting no interaction between them (Figure 3.5.4, *right*). Single site model equation was used to fit the binding isotherm. The interaction of DF complex and subunit *d* was occurring in equimolar ratio. The dissociation constant (K_d) calculated from equation was 52.9 µM, indicating a weak binding reaction between the DF-heterodimer and subunit *d*.

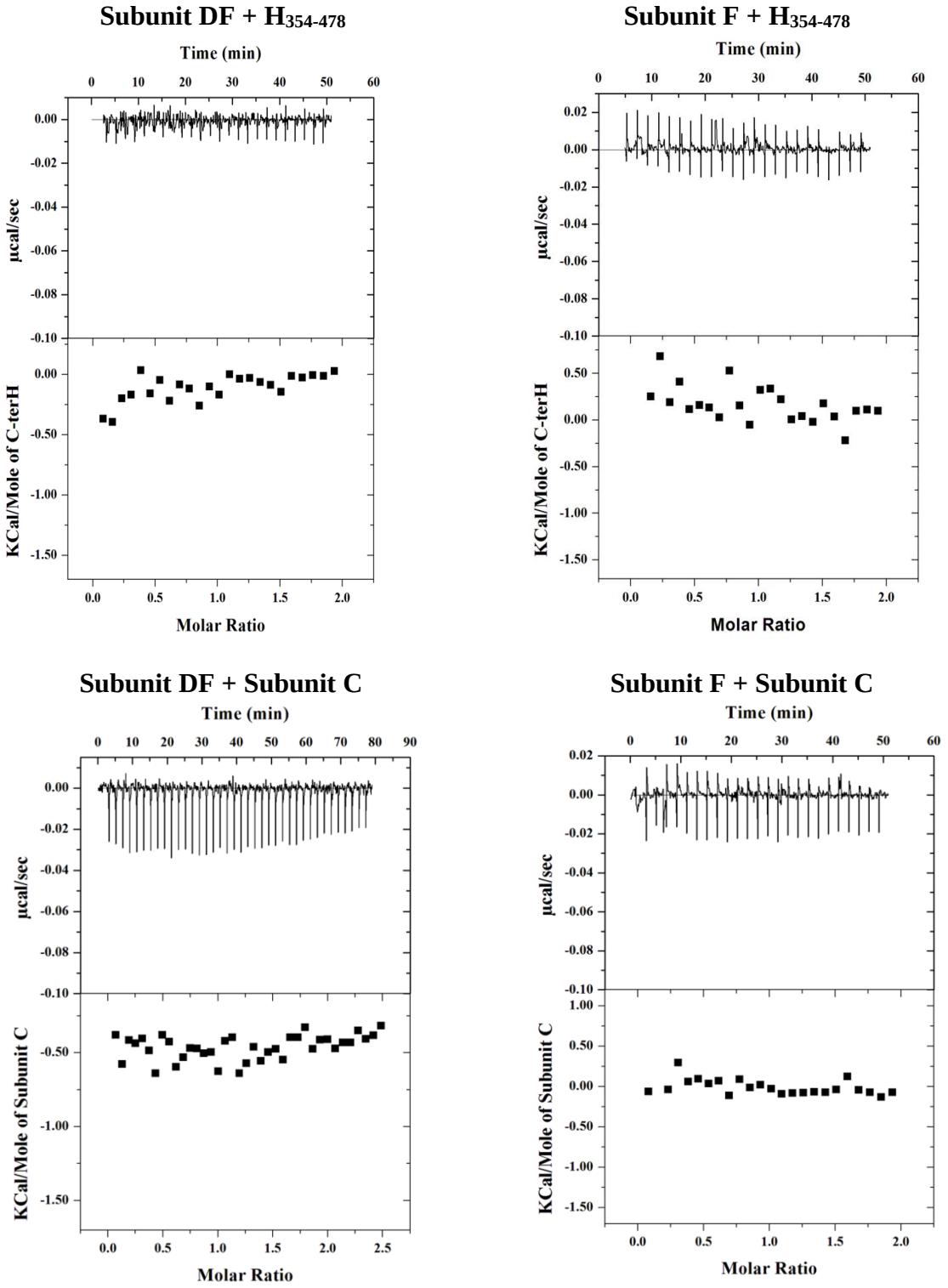


Figure 3.5.3: Binding affinity measurements for subunit F and the DF-heterodimer with subunit $H_{354-478}$ and C using ITC. Representative ITC profiles of DF-heterodimer (*left*) and subunit F (*right*) titration with subunit $H_{354-478}$ and C are shown for each experiment. The top and bottom panel shows the injection profile and heat release of each injection, respectively.

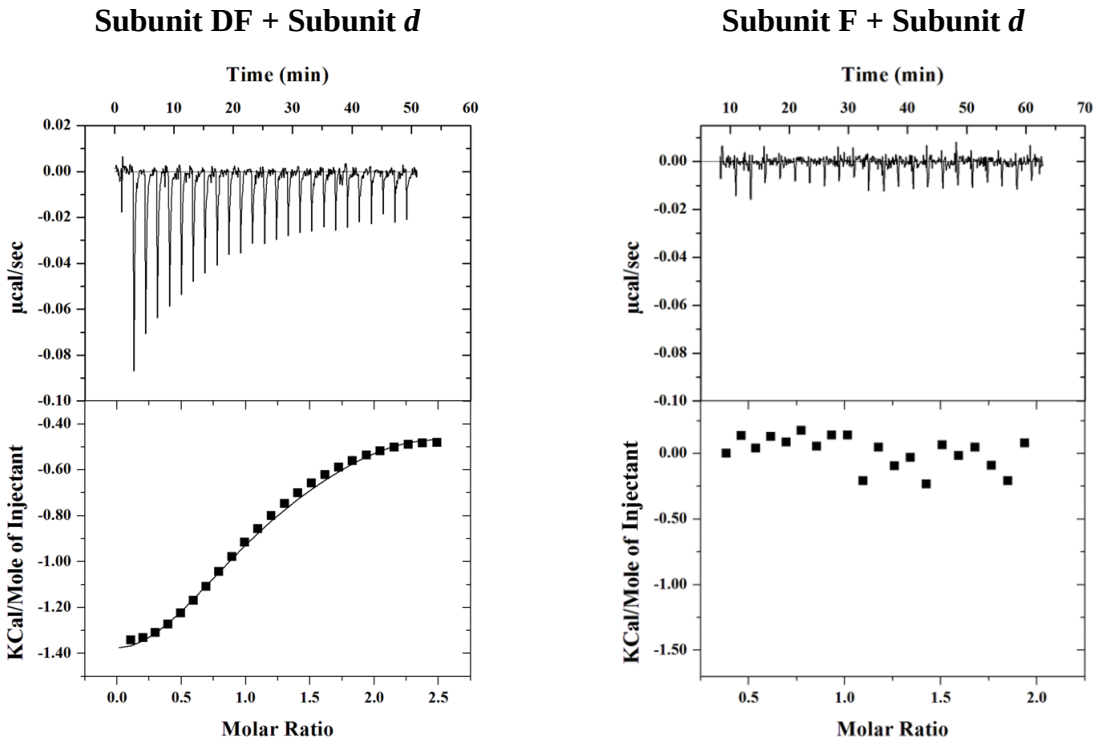


Figure 3.5.4: Binding experiments of subunit F and the DF-heterodimer with subunit *d* using ITC. Representative ITC profiles of DF-heterodimer (left) and subunit F (right) titration with subunit *d* is shown for each experiment. The top and bottom panel shows the injection profile and heat release of each injection, respectively (140).

3.6 Crystallization of the DF-heterodimer of V_1V_O ATPase from *S. cerevisiae*

3.6.1 Crystallization of the DF-heterodimer

Crystallization drops for subunit DF were setup in both sitting and hanging drops, using 5 mg/ml and 8 mg/ml protein concentrations. Initial crystals were observed in crystal screen 1 condition 11 of Hampton Research (0.10 M Sodium citrate tribasic dihydrate pH 5.6, 1.0 M $NH_4H_2PO_4$) at 18^o C. In hanging drops, numerous tiny crystals were observed with a dimension of 0.05 mm x 0.04 mm x 0.01 mm (Figure 3.6.1A). These crystals were small in size and do not diffract in the in-house Rigaku machine. After performing initial grid screen, crystals become bigger in size (0.13 mm x 0.10 mm x 0.04 mm) in the condition of 0.12 M sodium citrate tribasic dihydrate pH 5.6, 1.0 M $NH_4H_2PO_4$ (Figure 3.6.1B). Therefore, crystals were analyzed at the in-house Rigaku machine and diffracted to maximum resolution of 5 Å (Figure 3.6.1C). The diffraction data was indexed by HKL2000 suite program. The crystal belongs to trigonal space group P3 with unit cell parameters of $a = b = 167.73$ Å, $c = 127.14$ Å. The crystals were not stable for full dataset collection. To improve the stability of the DF crystals, further optimization is needed.

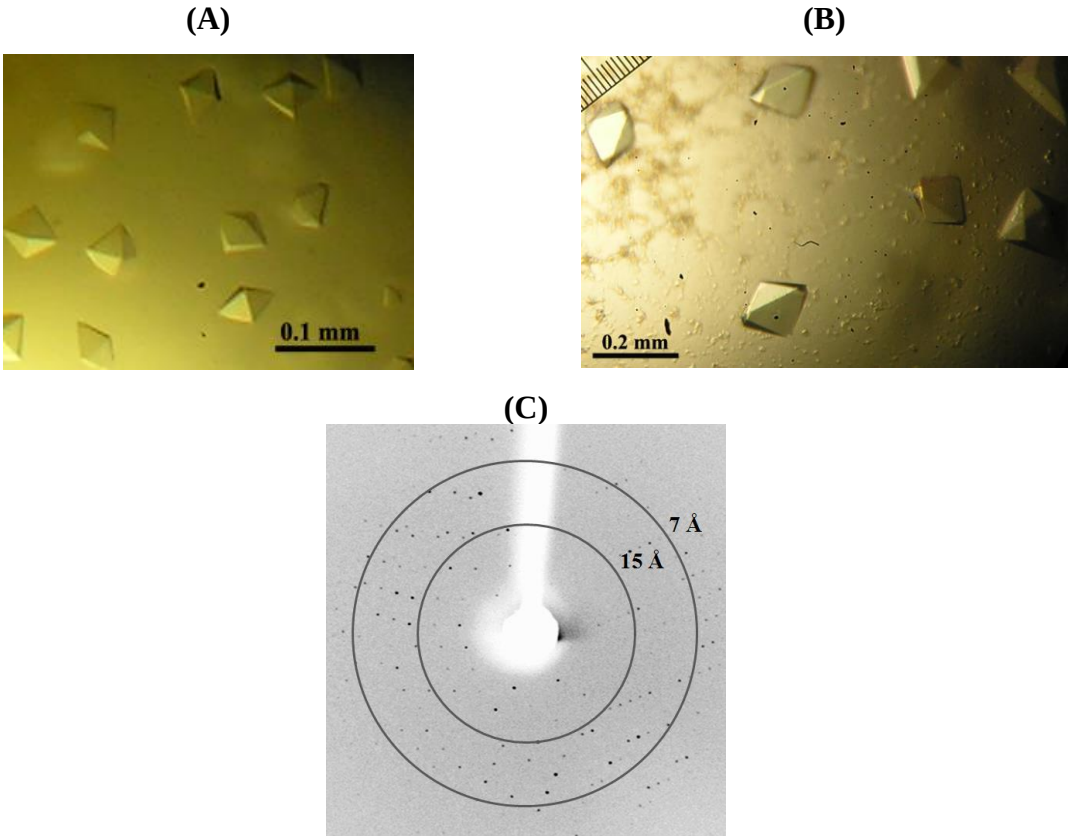


Figure 3.6.1: Crystals of the DF-heterodimer. *(A)* crystals obtained in the screen condition of 0.10 M Sodium citrate tribasic dihydrate pH 5.6, 1.0 M Ammonium phosphate monobasic. *(B)* Crystals of DF-heterodimer in the optimised condition (0.12 M Sodium citrate tribasic dihydrate pH 5.6, 1.0 M $NH_4H_2PO_4$). *(C)* The diffraction pattern of DF crystal after initial optimization. Circles show resolution shell for 15 Å and 7 Å, respectively.

Microseeding was performed to improve the quality of the crystals. Microseed were prepared in mother solution and directly added to the crystallization drop. The crystals were yielded which polarized under the microscope (Figure 3.6.2A). The crystals tested at in-house Rigaku machine showed few spots at the low resolution (Figure 3.6.2B).

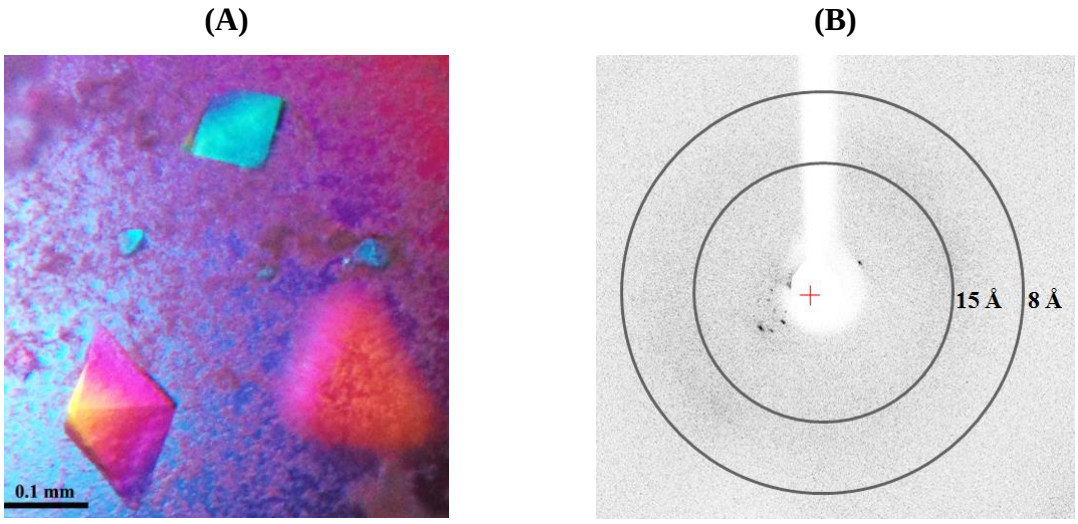


Figure 3.6.2: *(A)* Crystals of the DF-heterodimer appeared after microseeding in the condition of 0.12 M Sodium citrate tribasic dihydrate pH 5.6, 1.0 M $NH_4H_2PO_4$. *(B)* The diffraction pattern of DF crystal after microseeding.

Optimization was also done at the protein purification level. Step gradient has been introduced in the ion exchange Resource Q column (6 ml) to elute protein resulting the separation of peak2 from peak1 (Figure 3.6.3A). Protein fractions from peak2 were pooled together and applied onto the Superdex^{TM} S200. Only 20-30% of the apical peak fractions were pooled together to get highly pure protein (Figure 3.6.3B).

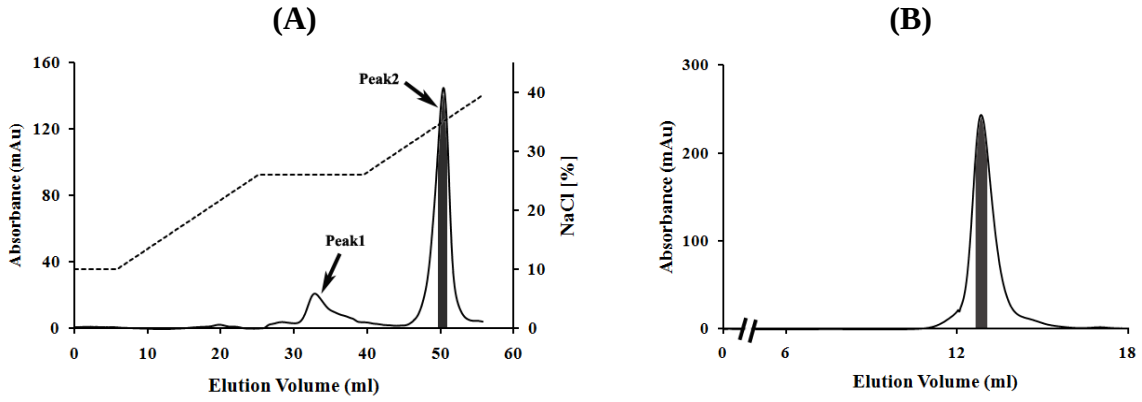


Figure 3.6.3: RESOURCE^{TM} Q (A) and Superdex^{TM} S200 *(B)* purification profile of DF-heterodimer of V-ATPase from *S. cerevisiae*.

4. Discussion

4.1 Eukaryotic V-ATPase from *S. cerevisiae*

V-ATPases are abundant in eukaryotes and function as a pH regulator in the intracellular compartment. Energy released after ATP is hydrolysis by V_1 domain, is used for proton translocation through the endomembrane maintaining pH in the vesicles. Dissociation of V_1 from the proton translocating V_O part is the unique feature of eukaryotic V_1V_o-ATPases, which is reversible and depending on several signalling pathways (145). During this period V-ATPases undergoes structural rearrangement or alteration and these changes are predicted to be more in stalk part (Figure 4.1A-C). It has been reported that subunit H has the ability to inhibit hydrolysis of ATP in the dissociated V_1 part. It was speculated that the reason might be due to the bridging of subunit H with central stalk subunit D and F. The inhibition of hydrolysis is not seen in the intact complex. To understand the connection between the central stalk with the ATP hydrolysis and the reversible dissociation, it is important to study their structure at the atomic level. In this thesis central stalk subunits are characterized using different complementary techniques to get deeper inside about the structural feature and its relationship with the ATP hydrolysis as well as reversible disassembly of the V-ATPase from *S. cerevisiae*.

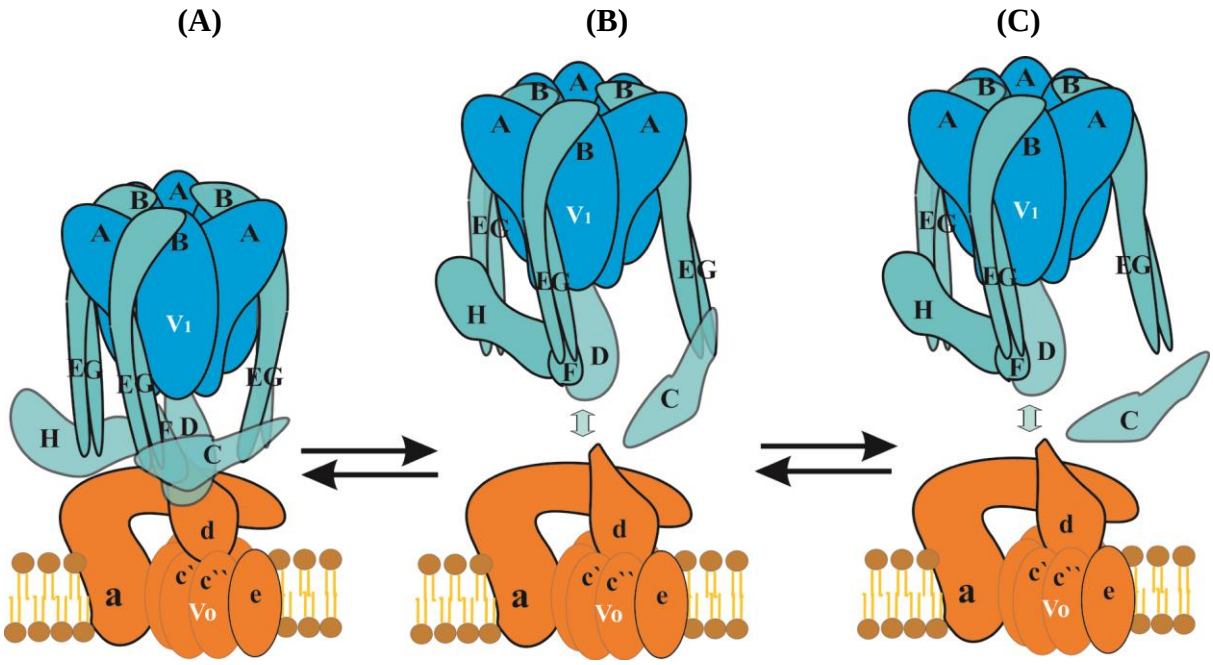


Figure 4.1: Topological model of V_1V_O ATPase. *(A)* Assembled V_1V_O ATPase. *(B)* Intermediate stage of dissociation. *(C)* Fully dissociated stage.

4.2 Solution shape of subunit F from *S. cerevisiae* V-ATPase

Small-angle X-ray scattering is a well establish method for the determination of low resolution three dimensional solution structures and is complementary to the high resolution experiments like X-ray crystallography and NMR spectroscopy (146). SAXS experiment was performed to get initial information about the domain feature of subunit F and $F_{1\text{-}94}$ which reveals two domains present in the subunit F. Subunit F has one egg-like major domain and one hook-like minor domain which are linked by a linker segment (Figure 3.1.5). The C-terminal truncated forms of subunit F, $F_{1\text{-}94}$ was generated where 24 amino acids form the C-terminal part are removed. This protein shows missing hook-like region (Figure 3.1.10), reflected by decreased D_{max} value (Figure 3.1.9B). The missing hook-like segment is composed of the C-terminal peptide $_{95}$PSKDHPYDPEKDSVLKRVRKLFGE$_{118}$. The peptide structure of *S. cerevisiae* $F_{90\text{-}116}$ was solved by NMR and fits nicely into the hook-like segment of subunit F (Figure 4.2A). Both low and high resolution structure of subunit F from *Methanococcus mazei* Gö1 (F_{Mm}) A-ATPsynthase has been solved in our lab which clearly shows that N- and C-terminal domains are separated in solution (Figure 4.2B). Compared to this protein, subunit F from *S. cerevisiae* V-ATPase (F_{Sc}) shows two well defined domains separated by linker segment inspite of having lesser sequence identity (15%).

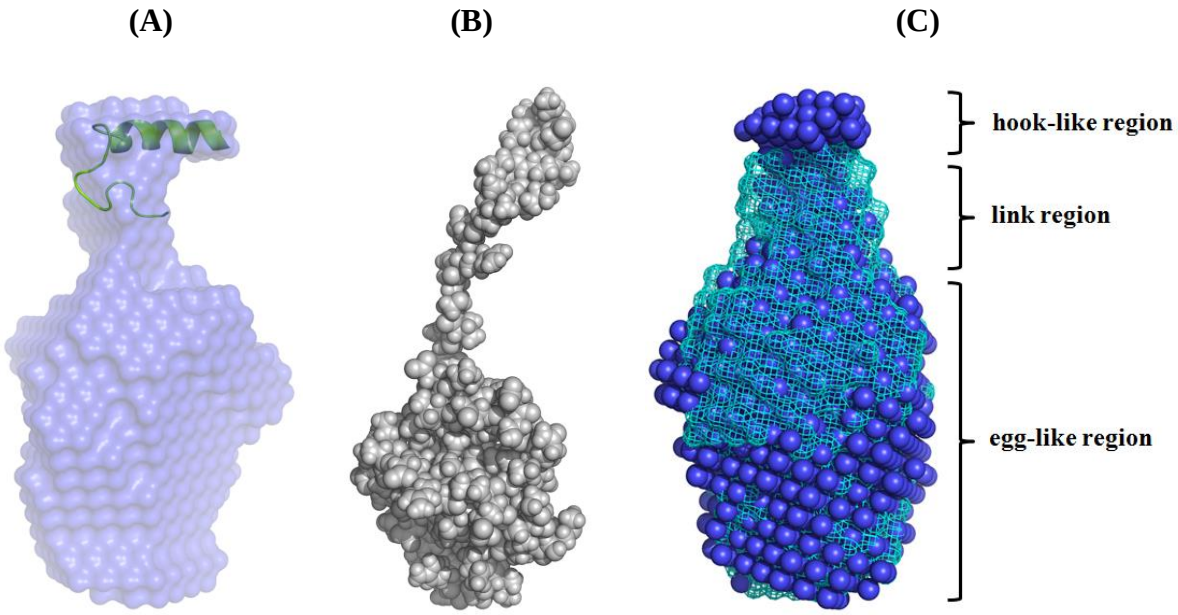


Figure 4.2: Comparison of F subunit from V_1V_O ATPase and A_1A_O ATP synthase. (*A*) Superimposition of NMR structure of $F_{90\text{-}116}$ peptide onto SAXS structure of subunit F. (*B*) Sphere representation of the subunit F from *Methanococcus mazei* Gö1 A_1A_O ATP synthase (85). (C) Superimposition of SAXS structures of subunit F and $F_{1\text{-}94}$ in blue and cyan, respectively.

Subunit F from *Thermus thermophilus* (F_{Tt}) A-ATP synthase also shows a similar structural feature compared to the low resolution solution structure of F_{Sc} (147). All available subunit F structures from different prokaryotic systems revealed a mixed content of alternative pattern of α-helices and β-strands. Circular dichroism experimental data of F_{Sc} also show mixed population of helices and strands (Figure 3.1.3). The very C-terminal α-helical peptide of prokaryotic and eukaryotic subunit F shows similarity in both cases (148). By contrast, both the low and high resolution structure of subunit F_{Mm} of the *M. mazei* Gö1 A_1A_O ATP synthase solved by SAXS (142) and NMR (85) shows major N-terminal globular domain, which is smaller compared to the F subunit of *S. cerevisiae* V-ATPase. Residues 79-101 at the very C-terminus of F_{Mm} form an extended stretch which is not present in F_{Sc} (Figure 4.2B-C). Because of the extended C-terminal region, 101 amino acid subunit F_{Mm} has a D_{max} of 7.8 nm (142), which is as long as the 118 amino acid residues of subunit F_{Sc} (7.7 nm). The direct involvement of the C-terminal helix of F_{Mm} in the coupling process of the proton pump as well as the catalytic process in the A_3B_3-hexamer of the A_1A_O ATP synthase was reported (85,142,148). Now the C-terminal region of the solution structure of F_{Sc} and F_{Mm} showed a hook-like and extended region, respectively, which are diverse in nature. This diversity in the structure might indicate a diverse mechanism in coupling event in the V-ATPase and A_1A_O ATP synthase. The low resolution structure of hydrated V_1 ATPase from *Manduca sexta* (149), determined by solution X-ray scattering experiment, shows that the central stalk has a length of 11 nm. The cryo-EM structure of the subunit C depleted V_1(-C) complex revealed an about 7 nm long central stalk (41). The difference in the length is due to the absence of subunit C in the V_1 complex. The D_{max} of F_{Sc} is 7.7 nm which would nicely occupy the length of the central stalk. Therefore, the diversity in the stalk part may reflect the necessity of diverse structural features of subunits which are involved in the coupling process.

4.3 Structural and mechanistic features of $F_{1\text{-}94}$ of the eukaryotic V-ATPase

The crystallographic structure of subunit F was solved at 2.3 Å using MAD technique. Subunit $F_{1\text{-}94}$ has an egg-like elliptical shape with a dimension of 30 x 16 x 38 Å (Figure 3.2.19B). The overall structure of $F_{1\text{-}94}$ belongs to an alpha-beta class of proteins with Rossmann fold topology where four-parallel β-strands are intermittently surrounded by four α-helices forming an egg-shaped structure. Backbone ΔC_α-ΔC_β based on random coil values and ^{13}C and ^{1}H chemical shifts using TALOS+ (143) of the entire subunit F of the *S. cerevisiae* V-ATPase is consistent with the crystallographic $F_{1\text{-}94}$ structure as well as the

recently determined C-terminal peptide structure of F_{90-116} (87). The present high resolution structural data of subunit F of a eukaryotic V-ATPase could be compared with the previously solved NMR solution- and X-ray crystallographic structure of subunit F of the related A-ATP synthase (83, 142, 145). The related F subunit shows sequence similarity less than 23%. The F_{1-94} structure with the C-terminal F_{90-116} from *S.cerevisiae* V-ATPase was superimposed with related subunit F of the *T. thermophilus* A-ATP synthase (147) (Figure 4.3A) which showed an overall r.m.s. deviation of 4.38 Å, revealing the diversity in the arrangement of α3 and ß3 present in the N-terminal globular domain as well as the major difference in ß4-sheet-helix-loop-helix in the C-terminal domain (140).

The extended conformation of the C-terminal portion of the subunit F from *T. thermophilus* and *M. mazei* Gö1 A_1A_O ATP synthase, is proposed to make contacts with catalytic subunit B (142,147,148). Superimposition of the F_{1-94} crystal structure from the *S.cerevisiae* V-ATPase (140), including the C-terminal NMR structure of F_{90-116} (87) with the A_1A_O ATP synthase subunit F from *M. mazei* Gö1 (F_{Mm} (85)) and *E. hirae* (F_{Eh} (144)) showed an overall r.m.s. deviation of 2.51 Å and 2.16 Å, respectively (Figure 4.3B). Superimposition revealed the presence of two loops between α1-β2 ($_{26}$GQITPETQEK$_{35}$) and α2-β3 ($_{60}$ERDDI$_{64}$) which are unique in eukaryotic subunit F (Figure 4.3B) (140).

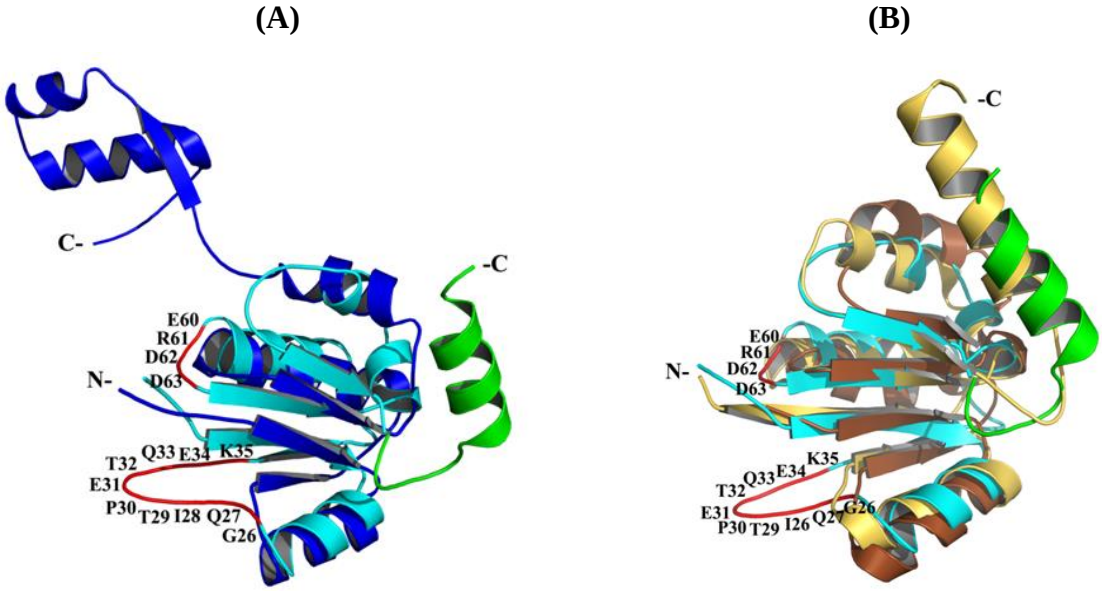


Figure 4.3: Superimposition of the eukaryotic subunit F from *S. cerevisiae* with subunit F of the related A-ATP synthase. *(A)* Overlap of *S. cerevisiae* F_{1-94} and α5-$F_{103-113}$ peptide on the *T. thermophilus* subunit F structure (147). *(B)* Ribbon representations of *S. cerevisiae* with subunit F, F_{1-94} (cyan), with the A-ATP synthase subunit F from *E. hirae* (golden (144)) and *M. mazei* Gö1 (wheat (85)). The α5-$F_{103-113}$ peptide (green) is fitted in the *S. cerevisiae* F_{1-94}. The $_{26}$GQITPETQEK$_{35}$-loop of the *S. cerevisiae* F subunit is colored in red (140).

Multiple sequence alignment of subunit F form different organism confirm the uniqueness of the loop part which contains polar residues ($_{26}$GQITPETQEK$_{35}$) and is found only in eukaryotic F subunits (Figure 3.2.20A). Interestingly, another loop contains a $_{60}$ERDDI$_{64}$-sequence motif and is conserved only among the eukaryotic F subunits (Figure 3.2.20A). The crystal structure of $F_{1\text{-}94}$ was docked into the assembled A_3B_3DF complex and finally superimposed onto the EM map of *S. cerevisiae* V-ATPase (Figure 4.7A), which reveals the localization of the $_{26}$GQITPETQEK$_{35}$-loop in the complex. The loop is exposed to the outside and nicely occupies into the density of the EM structure (Figure 4.7A) (46).

The modeled subunit D and high resolution crystal structure of $F_{1\text{-}94}$ of the *S. cerevisiae* V-ATPase were docked together using recently solved crystal structure of the subunit DF-heterodimer of the *E. hirae* A-ATP synthase as a template (144). Similar hydrophobic patches were found in the $F_{1\text{-}94}$ crystal structure as well as modeled subunit D from the *S. cerevisiae* V-ATPase (Figure 4.4). The presence of hydrophobic patches is important for a stable complex formation which allows the proteins to function as a heterodimer.

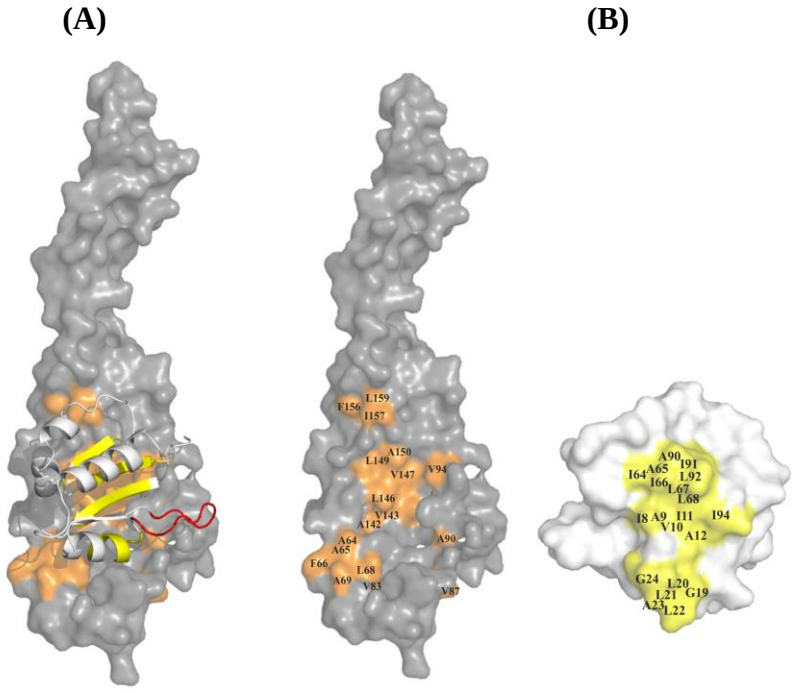


Figure 4.4: Model of $F_{1\text{-}94}$ in complex with subunit D from *S. cerevisiae*. *(A)* The crystal structure of *S. cerevisiae* $F_{1\text{-}94}$ was docked with the homology model of the *S. cerevisiae* subunit D based on the subunit DF-heterodimer of the *E. hirae* A-ATP synthase (144). *(B)* The surface representation of modelled subunit D (*left*) and $F_{1\text{-}94}$ (*right*) of the *S. cerevisiae* V-ATPase, where hydrophobic residues between the interface of subunit D and F, $F_{1\text{-}94}$ are labelled.

Further analysis revealed that the $_{91}$RXKXR$_{95}$ sequence motif in the β-hairpin region of the *S. cerevisiae* subunit D positioned closed to the *S. cerevisiae* subunit F loops, which includes the two unique loops ($_{24}$GIGQXXPE$_{31}$ and $_{60}$ERDDI$_{64}$) (140) (Figure 4.5).

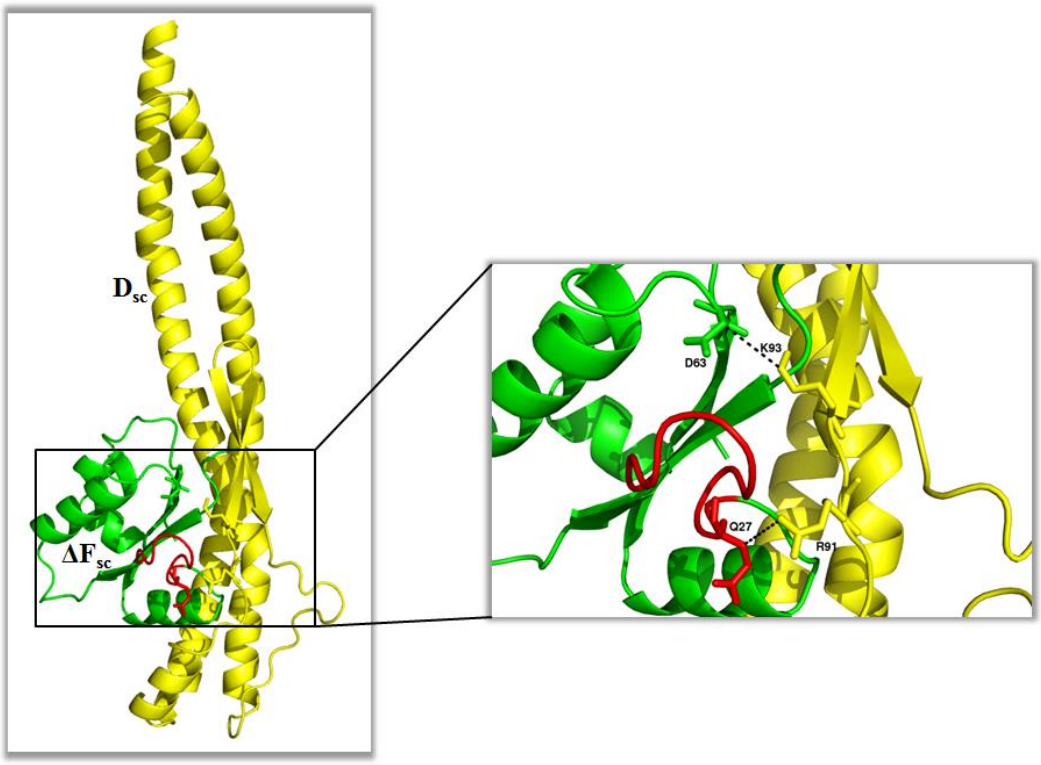


Figure 4.5: Charge interaction of F$_{1\text{-}94}$ with the *S. cerevisiae* subunit D homology model. Based on the DF-heterodimer from of the *E. hirae* A-ATP synthase (144), proximal charged residues in the eukaryotic *S. cerevisiae* subunit DF-heterodimer are shown as stick.

The β-hairpin of D subunit in the *E. hirae* A-ATP synthase was reported to stimulate ATPase activity two fold in the entire enzyme complex (144). Therefore, the close proximity of loops and β-hairpin of subunit F and D, respectively, might have influence in the regulation of ATPase activity of V-ATPases. Two basic residues K93 and R91 in the β-hairpin of D subunit are in close proximity to the acidic residue D63 of the $_{60}$ERDDI$_{64}$-loop and Q27 of the $_{26}$GQITPETQEK$_{35}$-loop of the *S. cerevisiae* F subunit. A multiple sequence alignment of subunit D from eukaryotic V-ATPases showed that K93 residue is conserved only among eukaryotic V-ATPases and not among F subunits of A-ATP synthases (Figure 4.6). The close vicinity of these two acidic-basic residues might allows the interaction between the central stalk subunits by formation of salt bridge which might be therefore regulate the movement of the β-hairpin in subunit D, leading to further alteration in the N- and C-terminal helices of

subunit D, which are in close proximity to the catalytic interfaces of the subunits A and B in V_1.

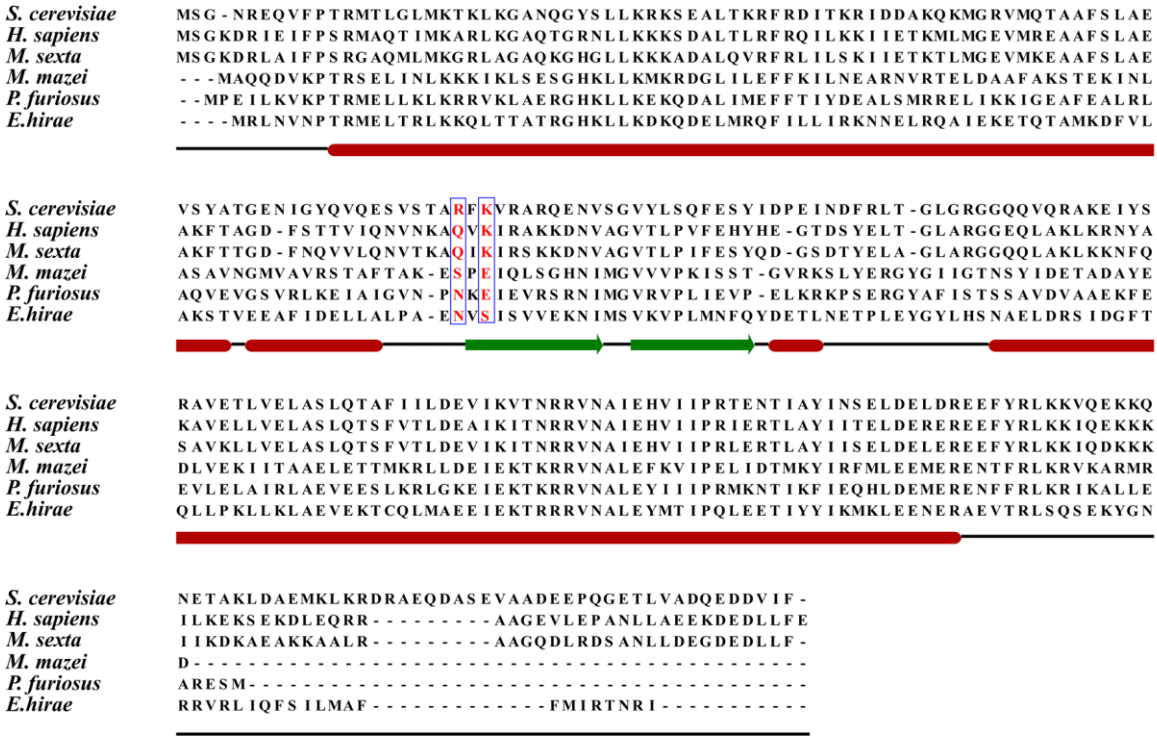


Figure 4.6: Multiple sequence alignment of subunit D from V-ATPases and A-ATP synthases from different organism. The secondary structure contents of *E. hirae* subunit D are shown below.

It is very important for the subunits of the enzyme to have sufficient flexibility, in order to regulate their activity during the dynamic equilibrium of reversible dissociation of V-ATPases. The ^{15}N-[^{1}H] heteronuclear NOE experiment of the entire subunit F of the *S. cerevisiae* V-ATPase revealed the formation of a rigid core by ß-strands, ß1 to ß4, and α2 to α4 (Figure 3.3.6A-B) whereas the N- and C-terminal helices α1 and α5 together with their adjacent loops $_{26}$GQITPETQEK$_{35}$ and $_{94}$IPSKDHPYD$_{102}$, respectively, are relatively more flexible in solution. Furthermore, $_{60}$ERDDI$_{64}$-loop is also considered to be flexible. This flexible loop together with the $_{26}$GQITPETQEK$_{35}$-loop of subunit F might have the ability to modulate the interaction with the residues of the β-hairpin in subunit D, leading to the changes in ATPase activity which is discussed above.

Docking of subunit F of *S. cerevisiae* V-ATPase onto the EM structure has been done using Situs (80) and EInemo (81) programs which show the final overall correlation coefficient 0.8357. The study revealed that the $_{26}$GQITPETQEK$_{35}$-loop of subunit F is facing towards

subunit H with a closest distance of 23 Å (Figure 4.7A). A previous study showed that the residue S381 of subunit H is facing to the central stalk subunit F and involved in the cross-linking with subunit F when the V_1 domain is disassembled (67). The S381 residue of subunit H was mutated to cysteine which yielded 10 Å cross link product with subunit F through 4-(N-Maleimido)benzophenone (MBP) crosslinking reagent induced by UV irradiation. The same crosslink could not be formed in the assembled V_1V_O ATPase. Based on these results it has been hypothesized that during reversible dissociation the C-terminal part of subunit H comes close to the central stalk and makes a bridge with both stator and rotor domains, causing blockage of rotation and leading to the decrease in ATP hydrolysis. This phenomenon is not occurring in the intact V-ATPase. A docking study showed that the distance between S381 of subunit H and the distal part of the $_{26}$GQITPETQEK$_{35}$-loop of subunit F is 32 Å, which is not close enough to generate a crosslink via the 10 Å long MBP cross linker. In a recent study it has been demonstrated that the linker part of subunit H allows the conformational change in the C-terminal domain of subunit H (46). Therefore, it is possible that during reversible dissociation of two domains of the V_1V_O ATPase the flexible C-terminal domain of subunit H might move slightly closer to its nearest neighbour, the exposed $_{26}$GQITPETQEK$_{35}$-loop of subunit F, where it causes conformational changes, leading to an inhibitory effect of ATPase activity in the V_1 ATPase (Figure 4.7B). The crosstalk among the subunit DF-heterodimer and the F-H interaction through the flexible $_{26}$GQITPETQEK$_{35}$- and $_{60}$ERDDI$_{64}$-loop might be involved in the regulation of ATPase activity during the disassembly and assembly of V_1V_O ATPase. However this hypothesis does not rule out the other possible influence by other subunits. The presented NMR and ITC data showed that there was no interaction between subunit F and DF with subunit H and C-terminal H. There might be additional subunit/s involved in the process in the whole V-ATPase or V_1 complex which is not clear in the isolated subunits study. It has been demonstrated that one EG heterodimer comes close contact with N-terminal domain of subunit H (46). Therefore, during disassembly of the enzyme, the EG heterodimer might hold subunit H by interacting with its N-terminal domain, followed by a movement of the C-terminal domain towards subunit F through the flexible linker region (Figure 4.7B).

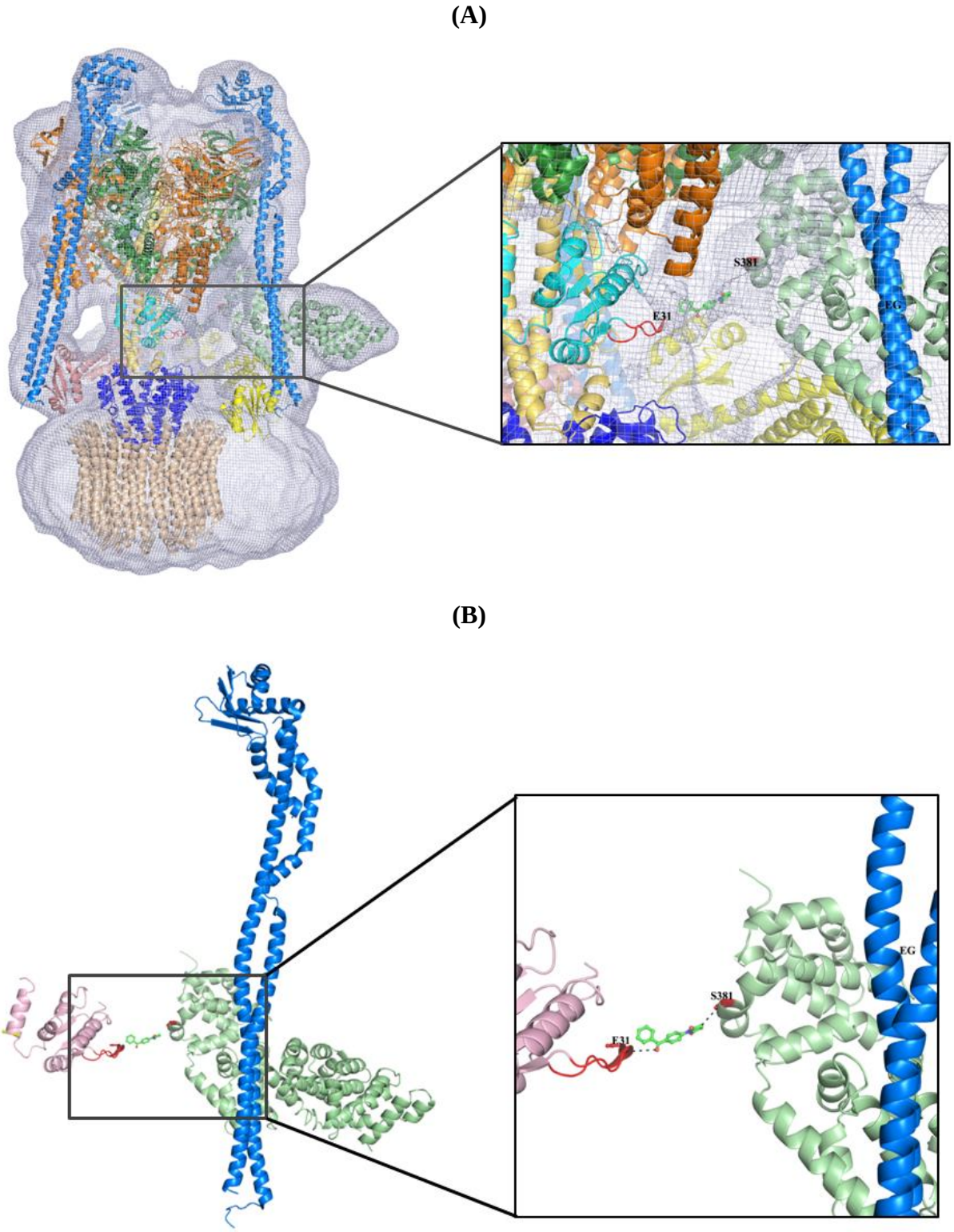


Figure 4.7: *(A)* Docking of existing high resolution structures of isolated subunits from V-ATPase and related A-ATP synthases onto the EM map of *S. cerevisiae* V-ATPase (46). Unique loops of subunit F and S381 of subunit H are highlighted in red color. *(B)* Close proximity of subunit F and H which might be involved in cross linking by MBP (stick representation) (140). EG heterodimer might be involved in interaction with N-terminal domain of subunit H.

During the process of reversible disassembly V_1 domain is reported to be dissociated from V_O domain resulting in detachment of the rotor subunits D and F from the subunit *d*. The ITC-data presented here showed that both D and F subunits are necessary for the interaction with subunit *d*. The determined dissociation constant (K_d) of DF-heterodimer and *d* binding is 52.9 µM, which reflects a weaker binding between these two partners which maybe stronger in the V_1V_O complex (140). The weak interaction between them might be essential for the dissociation of V_1 and V_O. Compared to this, the dissociation constant of related homologues subunits in non-dissociating A-ATP synthase from *E. hirae* was calculated to be 82 nM (144). From heteronuclear NOE experiment it is clear that helix α1of subunit F is flexible and is situated in the cavity of bottom part of subunit D (Figure 4.8). The interface of DF-heterodimer and subunit *d* should undergo more conformational alteration during the process of dissociation and reassembly of the V_1 and V_O sectors. Now the higher flexibility of α1 in subunit F (Figure 3.3.7A-B) might have influence on the alteration of bottom part of subunit D during the dissociation of the DF-heterodimer form the subunit *d*. These conformational changes might also have a role in the movement of subunit F closer to H, via the neighbouring $_{26}$GQITPETQEK$_{35}$-loop. Taken together, the eukaryotic subunit F plays a structural as well as mechanistic central role in the processes of reversible disassembly and ATP hydrolysis regulation of the V_1V_O ATPase by transmitting of at least the subunit movements of subunit *d* and H of the V_O and V_1 sector, respectively.

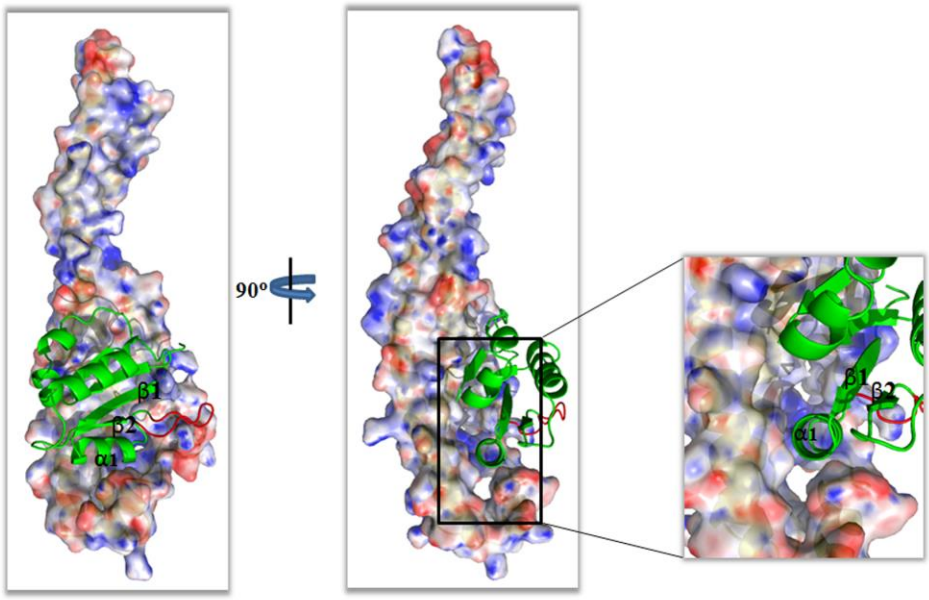


Figure 4.8: Fitting of helix α1 of $F_{1\text{-}94}$ inside the cavity of homology modelled subunit D from *S. cerevisiae*. The polar and hydrophobic match between helix α1 and the cavity of $D_{sc.}$ may be induced by the less rigid α1.

5. Conclusion

Eukaryotic V-ATPases are very important enzyme complexes involved in various physiological as well as pathophysiological events in higher organism including human beings. The mechanism of action of the enzyme is still not clear due to the lack of high resolution structure of the entire complex. In the present thesis, I have concentrated on the most dynamic region of the complex, which is called the central stalk composed of subunit D and F. The SAXS studies revealed that subunit F of the *S. cerevisiae* V-ATPase exists as a well defined two domain protein in solution. The major domain has an egg-like shape, formed by N-terminal amino acids. The C-terminus consists of a hook-like shape, which is α-helical in nature. The described crystal structure of $F_{1\text{-}94}$ revealed an alternating pattern of β-strands and α-helices with a dimension of 30 x 16 x 38 Å. The high resolution structure yielded two unique loops $_{26}$GQITPETQEK$_{35}$ and $_{60}$ERDDI$_{64}$, which are only present in eukayotic V-ATPases subunit F. NMR experiments confirmed the presence of secondary content, which are in line with the crystal structure and also confirmed the last α5 helix in the subunit F. Dynamic studies of subunit F from *S. cerevisiae* V-ATPase showed that the flexibility of this subunit is different from the subunit F from A-ATP synthases (138). Comparatively, the C-terminal helix of subunit F from *S. cerevisiae* V-ATPase is more rigid than the subunit F from *M. mazei* A-ATP synthase. This observation is justified from the low resolution structure the C-terminal part of subunit F from *S. cerevisiae* V-ATPase which is positioned orthogonally to the major domain. Based on the recently solved EM structure of V-ATPase (46), it is observed that subunit F and H can come closer during the process of reversible disassembly. The ^{15}N-[^{1}H] heteronuclear NOE experiments showed that the α1-helix is more flexible than the rest of the structural part which is adjacent to the unique $_{26}$GQITPETQEK$_{35}$. Flexibility of the particular helix might allow the movement of $_{26}$GQITPETQEK$_{35}$ loop towards the C-terminal region of subunit H, leading to the silencing of ATPase activity of the whole complex. Due to the presence of flexible linker part of subunit H, the C-terminal region might also move towards the unique $_{26}$GQITPETQEK$_{35}$ loop. The absence of interaction of subunit F with subunit H, $H_{354\text{-}478}$, C and *d* by both NMR and ITC indicates the dynamic and rather complicated process going on in the intact complex. Therefore, the DF-heterodimer was produced. The ITC experiment shows that the DF-heterodimer can bind to subunit *d* with a weak binding dissociation constant (K_d = 52.9 µM), which might be even stronger in whole complex. Weak binding between these two proteins might be necessary for the reversible dissociation to occur. The flexible α1-helix is situated at the cavity of the bottom part of subunit D found after homology modeling and docking experiment. Therefore, higher

flexibility of the α1-helix of subunit F might not only transmit the alterations in subunit *d* from the DF-heterodimer during dissociation but also allowing the movement of the unique loop $_{26}$GQITPETQEK$_{35}$ close towards the subunit H. Therefore, a high resolution structure of the DF-heterodimer is necessary to get deeper insight into this phenomenon.

6. References

1. Saroussi S. and Nelson N. (2009). *Vacuolar H^+-ATPase-an enzyme for all seasons*. Pflüg. Arch. Eur. J. Phy. **457**: 581-587.

2. Beyenbach K. W. and Wieczorek H. (2006). *The V-type H^+ ATPase: molecular structure and function, physiological roles and regulation*. J. Exp. Biol. **209**: 577-589.

3. Mellman I., Fuchs R. and Helenius A. (1986). *Acidification of the Endocytic and Exocytic Pathways*. Annu. Rev. Biochem. **55**: 663-700.

4. Nelson N. and Harvey W. R. (1999). *Vacuolar and Plasma Membrane Proton-Adenosinetriphosphatases*. Physiol. Rev. **79**: 361-385.

5. Nishi T. and Forgac M. (2002). *The vacuolar (H^+)-ATPases - nature's most versatile proton pumps*. Nat. Rev. Mol. Cell Biol. **3**: 94-103.

6. Morel N. (2003). *Neurotransmitter release: the dark side of the vacuolar-H+ATPase*. Biol. Cell **95**: 453-457.

7. Michaelson D. M. and Angel I. (1980). *Determination of DpH in cholinergic synaptic vesicles: its effect on storage and release of acetylcholine*. Life Sci. **27**: 39-44.

8. Marshansky V. (2007). *The V-ATPase a2-subunit as a putative endosomal pH-sensor*. Biochem. Soc. Trans. **35**: 1092-1099.

9. Recchi C. and Chavrier P. (2006). *V-ATPase: a potential pH sensor*. Nat. Cell Biol. **8**: 107-109.

10. Hurtado-Lorenzo A., Skinner M., Annan J. E., Futai M., Sun-Wada G.-H., Bourgoin S., Casanova J., Wildeman A., Bechoua S., Ausiello D. A., Brown D. and Marshansky V. (2006). *V-ATPase interacts with ARNO and Arf6 in early endosomes and regulates the protein degradative pathway*. Nat. Cell Biol. **8**: 124-136.

11. Brown D. and Breton S. (2000). *H^+V-ATPase-dependent luminal acidification in the kidney collecting duct and the epididymis/vas deferens: vesicle recycling and transcytotic pathways*. J. Exp. Biol. **203**: 137-145.

12. Grüber G. and Marshansky V. (2008). *New insights into structure-function relationships between archeal ATP synthase (A_1A_O) and vacuolar type ATPase (V_1V_O)*. BioEssays **30**: 1-14.

13. Breton S. and Brown D. (2007). *New insights into the regulation of V-ATPase-dependent proton secretion*. AJP - Renal Physiol. **292**: F1-10.

14. Zhang Z., Zheng Y., Mazon H., Milgrom E., Kitagawa N., Kish-Trier E., Heck A. J. R., Kane P. M. and Wilkens S. (2008). *Structure of the Yeast Vacuolar ATPase*. J. Biol. Chem. **283**: 35983-35995.

15. Forgac M. (2007). *Vacuolar ATPases: Rotary proton pumps in physiology and pathophysiology*. Nat. Rev. Mol. Cell Biol. **8**: 917-929.

16. Lepier A., Azuma M., Harvey W. R. and Wieczorek H. (1994). *K+/H+ antiport in the tobacco hornworm midgut: the K(+)-transporting component of the K+ pump*. J. Exp. Biol. **196**: 361-373.

17. Wieczorek H., Gruber G., Harvey W. R., Huss M. and Merzendorfer H. (1999). *The plasma membrane H^+-V-ATPase from tobacco hornworm midgut*. J. Bioenerg. Biomembr. **31**: 67-74.

18. Zhuang Z., Linser P. J. and Harvey W. R. (1999). *Antibody to H(+) V-ATPase subunit E colocalizes with portasomes in alkaline larval midgut of a freshwater mosquito (Aedes aegypti)*. J. Exp. Biol. **202**: 2449-2460.

19. Forgac M., Cantley L., Wiedenmann B., Altstiel L. and Branton D. (1983). *Clathrin-coated vesicles contain an ATP-dependent proton pump*. Proc. Natl. Acad. Sci. **80**: 1300-1303.

20. Moriyama Y., Maeda M. and Futai M. (1992). *The role of V-ATPase in neuronal and endocrine systems*. J. Exp. Biol. **172**: 171-178.

21. Morel N. (2003). *Neurotransmitter disease: The dark side of the vacuolar-H^+ATPase*. Biol. Cell **95**: 453-457.

22. Kane P. M. (2000). *Regulation of V-ATPases by reversible disassembly*. FEBS Lett. **469**: 137-141.

23. Kane P. M. (1995). *Disassembly and reassembly of the yeast vacuolar H^+-ATPase in vivo*. J. Biol. Chem. **270**: 17025-17032.

24. Adams D. S., Robinson K. R., Fukumoto T., Yuan S., Albertson R. C., Yelick P., Kuo L., McSweeney M. and Levin M. (2006). *Early, H^+-V-ATPase-dependent proton flux is necessary for consistent left-right patterning of non-mammalian vertebrates*. Development **133**: 1657-1671.

25. Alves dos Santos M. and Smidt M. (2011). *En1 and Wnt signaling in midbrain dopaminergic neuronal development*. Neural Develop. **6**: 23.

26. Cruciat C.-M., Ohkawara B., Acebron S. P., Karaulanov E., Reinhard C., Ingelfinger D., Boutros M. and Niehrs C. *Requirement of Prorenin Receptor and Vacuolar H^+-ATPase-Mediated Acidification for Wnt Signaling*. Science **327**: 459-463.

27. Hinton A., Bond S. and Forgac M. (2009). *V-ATPase functions in normal and disease processes*. Pflüg. Arch. Eur. J. Phy. **457**: 589-598.

28. Sennoune S. R., Luo D. and Martinez-Zaguilan R. (2004). *Plasmalemmal vacuolar-type H^+-ATPase in cancer biology*. Cell Biochem. Biophys. **40**: 185-206.

29. Sennoune S. R. and Martinez-Zaguilan R. (2007). *Plasmalemmal vacuolar H^+-ATPases in angiogenesis, diabetes and cancer*. J. Bioenerg. Biomembr. **39**: 427-433.

30. McHenry P., Wang W. L. W., Devitt E., Kluesner N., Davisson V. J., McKee E., Schweitzer D., Helquist P. and Tenniswood M. *Iejimalides A and B inhibit lysosomal vacuolar H^+-ATPase (V-ATPase) activity and induce S-phase arrest and apoptosis in MCF-7 cells*. J. Cell. Biochem. **109**: 634-642.

31. Sennoune S. R., Bakunts K., Martinez G. M., Chua-Tuan J. L., Kebir Y., Attaya M. N. and Martinez-Zaguilan R. (2004). *Vacuolar H^+-ATPase in human breast cancer cells with distinct metastatic potential: distribution and functional activity*. AJP - Cell Physiol. **286**: C1443-1452.

32. Lagadic-Gossmann D., Huc L. and Lecureur V. (2004). *Alterations of intracellular pH homeostasis in apoptosis: origins and roles*. Cell Death Differ. **11**: 953-961.

33. Gottlieb R. A., Giesing H. A., Zhu J. Y., Engler R. L. and Babior B. M. (1995). *Cell acidification in apoptosis: granulocyte colony-stimulating factor delays programmed cell death in neutrophils by up-regulating the vacuolar H(+)-ATPase*. Proc. Natl. Acad. Sci. **92**: 5965-5968.

34. Tanigaki K., Sasaki S. and Ohkuma S. (2003). *In bafilomycin A_1-resistant cells, bafilomycin A_1 raised lysosomal pH and both prodigiosins and concanamycin A inhibited growth through apoptosis*. FEBS Lett. **537**: 79-84.

35. Daniel C., Bell C., Burton C., Harguindey S., Reshkin S. J. and Rauch C. (2013). *The role of proton dynamics in the development and maintenance of multidrug resistance in cancer*. Biochim. Biophys. Acta **1832**: 606-617.

36. Nishihara T., Akifusa S., Koseki T., Kato S., Muro M. and Hanada N. (1995). *Specific inhibitors of vacuolar type H^+-ATPases induce apoptotic cell ceath*. Biochem. Biophys. Res. Commun. **212**: 255-262.

37. Xu J., Feng H. T., Wang C., Yip K. H., Pavlos N., Papadimitriou J. M., Wood D. and Zheng M. H. (2003). *Effects of Bafilomycin A1: An inhibitor of vacuolar H (+)-ATPases on endocytosis and apoptosis in RAW cells and RAW cell-derived osteoclasts*. J. Cell. Biochem. **88**: 1256-1264.

38. Yoshimoto Y. and Imoto M. (2002). *Induction of EGF-dependent apoptosis by vacuolar-type H^+-ATPase inhibitors in A431 cells overexpressing the EGF receptor*. Exp. Cell Res. **279**: 118-127.

39. Sumner J.-P., Dow J. A. T., Earley F. G. P., Klein U., Jäger D. and Wieczorek H. (1995). *Regulation of plasma membrane V-ATPase activity by dissociation of peripheral subunits*. J. Biol. Chem. **270**: 5649-5653.

40. Trombetta E. S., Ebersold M., Garrett W., Pypaert M. and Mellman I. (2003). *Activation of Lysosomal Function During Dendritic Cell Maturation*. Science **299**: 1400-1403.

41. Radermacher M., Ruiz T., Wieczorek H. and Grüber G. (2001). *The Structure of the V_1-ATPase Determined by Three-Dimensional Electron Microscopy of Single Particles*. J. Struct. Biol. **135**: 26-37.
42. Grüber G. (2003). *Introduction: a close look at the vacuolar ATPase*. J. Bioenerg. Biomembr. **35**: 277-280.

43. Lolkema J. S., Chaban Y. and Boekema E. J. (2003). *Subunit Composition, Structure, and Distribution of Bacterial V-Type ATPases*. J. Bioenerg. Biomembr. **35**: 323-335-335.

44. Kluge C., Lahr J., Hanitzsch M., Bolte S., Golldack D. and Dietz K. J. (2003). *New Insight into the Structure and Regulation of the Plant Vacuolar H^+-ATPase*. J. Bioenerg. Biomembr. **35**: 377-388-388.

45. Grüber G., Radermacher M., Ruiz T., Godovac-Zimmermann J., Canas B., Kleine-Kohlbrecher D., Huss M., Harvey W. R. and Wieczorek H. (2000). *Three-Dimensional Structure and Subunit Topology of the V_1 ATPase from Manduca sexta Midgut* Biochemistry **39**: 8609-8616.

46. Benlekbir S., Bueler S. A. and Rubinstein J. L. (2012). *Structure of the vacuolar-type ATPase from Saccharomyces cerevisiae at 11-Å resolution*. Nat. Struct. Mol. Biol. **19**: 1356-1362.

47. Sambade M. and Kane P. M. (2004). *The yeast vacuolar proton-translocating ATPase contains a subunit homologous to the Manduca sexta and bovine e subunits that Is essential for function*. J. Biol. Chem. **279**: 17361-17365.

48. Compton M. A., Graham L. A. and Stevens T. H. (2006). *Vma9p (Subunit e) is an integral membrane V_O subunit of the yeast V-ATPase*. J. Biol. Chem. **281**: 15312-15319.

49. Supek F., Supekova L., Mandiyan S., Pan Y. C. E., Nelson H. and Nelson N. (1994). *A novel accessory subunit for vacuolar H^+-ATPase from chromaffin granules*. J. Biol. Chem. **269**: 24102-24106.

50. Jansen E. J. R., Holthuis J. C. M., McGrouther C., Burbach J. P. H. and Martens G. J. M. (1998). *Intracellular trafficking of the vacuolar H^+-ATPase accessory subunit Ac45*. J. Cell Sci. **111**: 2999-3006.

51. Ludwig J., Kerscher S., Brandt U., Pfeiffer K., Getlawi F., Apps D. K. and Schägger H. (1998). *Identification and characterization of a novel 9.2-kDa membrane sector-associated protein of vacuolar proton-ATPase from chromaffin granules*. J. Biol. Chem. **273**: 10939-10947.

52. Demirci F. Y. K., White N. J., Rigatti B. W., Lewis K. F. and Gorin M. B. (2001). *Identification, genomic structure, and screening of the vacuolar proton-ATPase membrane sector-associated protein M8-9 gene within the COD1 critical region (Xp11.4)*. Mol. Vis. **7**: 234-239.

53. Merzendorfer H., Huss M., Schmid R., Harvey W. R. and Wieczorek H. (1999). *A Novel Insect V-ATPase Subunit M9.7 Is Glycosylated Extensively*. J. Biol. Chem. **274**: 17372-17378.

54. Nishi T., Kawasaki-Nishi S. and Forgac M. (2003). *Expression and Function of the Mouse V-ATPase d Subunit Isoforms*. J. Biol. Chem. **278**: 46396-46402.

55. Sun-Wada G.-H., Murata Y., Namba M., Yamamoto A., Wada Y. and Futai M. (2003). *Mouse Proton Pump ATPase C Subunit Isoforms (C2-a and C2-b) Specifically Expressed in Kidney and Lung*. J. Biol. Chem. **278**: 44843-44851.

56. Murata Y., Sun-Wada G. H., Yoshimizu T., Yamamoto A., Wada Y. and Futai M. (2002). *Differential Localization of the Vacuolar H^+ Pump with G Subunit Isoforms (G1 and G2) in Mouse Neurons*. J. Biol. Chem. **277**: 36296-36303.

57. Karet F. (2005). *Physiological and Metabolic Implications of V-ATPase Isoforms in the Kidney*. J. Bioenerg. Biomembr. **37**: 425-429-429.

58. Nishi T. (2005). *Subunit isoforms controlling the functions of the V-ATPase*. Seikagaku **77**: 354-358.

59. Strompen G., Dettmer J., Stierhof Y. D., Schumacher K., Jürgens G. and Mayer U. (2005). *Arabidopsis vacuolar H^+-ATPase subunit E isoform 1 is required for Golgi organization and vacuole function in embryogenesis*. Plant J. **41**: 125-132.

60. Armbrüster A., Christina H., Anne H., Karin S., Michael B. and Grüber, G. (2005). *Evidence for major structural changes in subunit C of the vacuolar ATPase due to nucleotide binding*. FEBS Lett. **579**: 1961-1967.

61. Hong-Hermesdorf A., Brux A., Grüber A., Grüber G. and Schumacher K. (2006). *A WNK kinase binds and phosphorylates V-ATPase subunit C*. FEBS Lett. **580**: 932-939.

62. Graham L. A., Hill K. J. and Stevens T. H. (1995). *VMA8 encodes a 32-kDa VO subunit of the Saccharomyces cerevisiae vacuolar H+-ATPase required for function and assembly of the enzyme complex*. J. Biol. Chem. **270**: 15037-15044.

63. Owegi M. A., Carenbauer A. L., Wick N. M., Brown J. F., Terhune K. L., Bilbo S. A., Weaver R. S., Shircliff R., Newcomb N. and Parra-Belky K. J. (2005). *Mutational analysis of the stator subunit E of the yeast V-ATPase*. J. Biol. Chem. **280**: 18393-18402.

64. Charsky C. M. H., Schumann N. J. and Kane P. M. (2000). *Mutational analysis of subunit G (Vma10p) of the yeast vacuolar H^+-ATPase*. J. Biol. Chem. **275**: 37232-37239.

65. Parra K. J., Keenan K. L. and Kane P. M. (2000). *The H subunit (Vma13p) of the yeast V-ATPase inhibits the ATPase activity of cytosolic V1 complexes*. J. Biol. Chem. **275**: 21761-21767.

66. Sagermann M., Stevens T. H. and Matthews B. W. (2001). *Crystal structure of the regulatory subunit H of the V-type ATPase of Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. **98**: 7134-7139.

67. Jefferies K. C. and Forgac M. (2008). *Subunit H of the vacuolar (H^+) ATPase inhibits ATP hydrolysis by the free V1 domain by interaction with the rotary subunit F*. J. Biol. Chem. **283**: 4512-4519.

68. Nelson H., Mandiyan S. and Nelson N. (1994). *The Saccharomyces cerevisiae VMA7 gene encodes a 14-kDa subunit of the vacuolar H(+)-ATPase catalytic sector*. J. Biol. Chem. **269**: 24150-24155.

69. Omri D., Felix F. and Nathan N. (2004). *Crystal structure of yeast V-ATPase subunit C reveals its stator function* EMBO Rep. **5**: 1148-1152.

70. Oot Rebecca A., Huang L.-S., Berry Edward A. and Wilkens S. (2012). *Crystal Structure of the Yeast Vacuolar ATPase Heterotrimeric EGChead Peripheral Stalk Complex*. Structure **20**: 1881-1892.

71. Rishikesan S., Gayen S., Thaker Y. R., Vivekanandan S., Manimekalai M. S. S., Yau Y. H., Shochat S. G. and Grüber G. (2009). *Assembly of subunit d (Vma6p) and G (Vma10p) and the NMR solution structure of subunit G (G_{1-59}) of the Saccharomyces cerevisiae V_1V_O ATPase*. Biochim. Biophys. Acta - Bioenergetics **1787**: 242-251.

72. Rishikesan S., Manimekalai M. S. S. and Grüber G. (2010). *The NMR solution structure of subunit G (G61-101) of the eukaryotic V_1V_O ATPase from Saccharomyces cerevisiae*. Biochim. Biophys. Acta - Biomembranes **1798**: 1961-1968.

73. Armbrüster A., Dmitri I. S., Coscun U., Sandra J., Susanne M. B. and Grüber G. (2004). *Structural analysis of the stalk subunit Vma5p of the yeast V-ATPase in solution*. FEBS Lett. **570**: 119-125.

74. Curtis K. K., Francis S. A., Oluwatosin Y. and Kane P. M. (2002). *Mutational Analysis of the Subunit C (Vma5p) of the Yeast Vacuolar H^+-ATPase*. J. Biol. Chem. **277**: 8979-8988.

75. Landolt-Marticorena C., Williams K. M., Correa J., Chen W. and Manolson M. F. (2000). *Evidence that the NH2 terminus of Vph1p, an integral subunit of the V0 sector of the yeast V-ATPase, interacts directly with the Vma1p and Vma13p subunits of the V1 Sector*. J. Biol. Chem. **275**: 15449-15457.

76. Jones R. P. O., Durose L. J., Findlay J. B. C. and Harrison M. A. (2005). *Defined sites of interaction between subunits E (Vma4p), C (Vma5p), and G (Vma10p) within the stator structure of the Vacuolar H(+)-ATPase*. Biochemistry **44**: 3933-3941.

77. Grüber G., Radermacher M., Ruiz T., Godovac-Zimmermann J., Canas B., Kleine-Kohlbrecher D., Huss M., Harvey W. R. and Wieczorek H. (2000). *Three-dimensional structure and subunit topology of the V_1 ATPase from Manduca sexta midgut*. Biochemistry **39**: 8609-8616.

78. Rishikesan S., Thaker Y. R., Priya R., Gayen S., Manimekalai M. S. S., Hunke C. and Grüber G. (2008). *Spectroscopical identification of residues of subunit G of the yeast V-ATPase in its connection with subunit E*. Mol. Membr. Biol. **25**: 400 - 410.

79. Rizzo V. F., Coskun U., Radermacher M., Ruiz T., Armbrüster A. and Grüber G. (2003). *Resolution of the V_1 ATPase from Manduca sexta into subcomplexes and visualization of an ATPase-active A_3B_3EG complex by electron microscopy*. J. Biol. Chem. **278**: 270-275.

80. Rusu M., Birmanns S. and Wriggers W. (2008). *Biomolecular pleiomorphism probed by spatial interpolation of coarse models.* Bioinformatics **24**: 2460-2466.

81. Suhre K. and Sanejouand Y. H. (2004). *ElNémo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement.* Nucleic Acids Res. **32**: 610-614.

82. Boyer P. D. (1997). *The ATP synthase—a splendid molecular machine*. Annu. Rev. Biochem. **66**: 717-749.

83. Oot R. A. and Wilkens S. (2012). *Subunit interactions at the V_1-V_O interface in yeast vacuolar ATPase*. J. Biol. Chem. **287**: 13396-13406.

84. Coskun Ü., Rizzo V. F., Koch M. H. J. and Grüber G. (2004). *Ligand-dependent structural changes in the V_1 ATPase from Manduca sexta*. J. Bioenerg. Biomembr. **36**: 249-256.

85. Gayen S., Vivekanandan S., Biuković G., Grüber G. and Yoon H. S. (2007). *NMR Solution Structure of Subunit F of the Methanogenic A_1A_O Adenosine Triphosphate Synthase and Its Interaction with the Nucleotide-Binding Subunit B*. Biochemistry **46**: 11684-11694.

86. Grüber G., Godovac-Zimmermann J., Link T. A., Coskun Ü., Rizzo V. F., Betz C. and Bailer S. M. (2002). *Expression, purification, and characterization of subunit E, an essential subunit of the vacuolar ATPase*. Biochem. Biophys. Res. Commun. **298**: 383-391.

87. Basak S., Gayen S., Thaker Y. R., Manimekalai M. S. S., Roessle M., Hunke C. and Grüber G. (2011). *Solution structure of subunit F (Vma7p) of the eukaryotic V_1V_O ATPase from Saccharomyces cerevisiae derived from SAXS and NMR spectroscopy*. Biochim. Biophys. Acta - Biomembranes **1808**: 360-368.

88. Laemmli U. K. (1970). *Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4*. Nature **227**: 680-685.

89. Manavalan P. and Johnson W. C. (1987). *Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra*. Anal. Biochem. **167**: 76-85.

90. Sreerama N. and Woody R. W. (1993). *A Self-Consistent Method for the Analysis of Protein Secondary Structure from Circular Dichroism*. Anal. Biochem. **209**: 32-44.

91. Provencher S. W. (1982). *A constrained regularization method for inverting data represented by linear algebraic or integral equations*. Comput. Phys. Commun. **27**: 213-227.

92. Andrade M. A., Chacon P., Merelo J. J. and Moran F. (1993). *Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network*. Protein Eng., Des. Sel. **6**: 383-390.

93. Deléage G. and Geourjon C. (1993) An interactive graphic program for calculating the secondary structure content of proteins from circular dichroism spectrum. Oxford University Press

94. Böhm G., Muhr R. and Jaenicke R. (1992). *Quantitative analysis of protein far UV circular dichroism spectra by neural networks*. Protein Eng. **5**: 191-195.

95. Roessle M. W., Klaering R., Ristau U., Robrahn B., Jahn D., Gehrmann T., Konarev P., Round A., Fiedler S., Hermes C. and Svergun D. (2007). *Upgrade of the small-angle X-ray scattering beamline X33 at the European Molecular Biology Laboratory, Hamburg*. J. Appl. Crystallogr. **40**: s190-s194.

96. Round A. R., Franke D., Moritz S., Huchler R., Fritsche M., Malthan D., Klaering R., Svergun D. I. and Roessle M. (2008). *Automated sample-changing robot for solution scattering experiments at the EMBL Hamburg SAXS station X33*. J. Appl. Crystallogr. **41**: 913-917.

97. Konarev P. V., Volkov V. V., Sokolova A. V., Koch M. H. J. and Svergun D. I. (2003). *PRIMUS: a Windows PC-based system for small-angle scattering data analysis*. J. Appl. Crystallogr. **36**: 1277-1282.

98. Guinier A. and Fournet G. (1955). *Small-angle Scattering of X-rays*. Wiley, New York

99. Svergun D. (1992). *Determination of the regularization parameter in indirect-transform methods using perceptual criteria*. J. Appl. Crystallogr. **25**: 495-503.

100. Svergun D. I. (1999). *Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing*. Biophys. J. **76**: 2879-2886.

101. Svergun D. I., Petoukhov M. V. and Koch M. H. J. (2001). *Determination of Domain Structure of Proteins from X-Ray Solution Scattering*. Biophys. J **80**: 2946-2953.

102. Ho S. N., Hunt H. D., Horton R. M., Pullen J. K. and Pease L. R. (1989). *Site-directed mutagenesis by overlap extension using the polymerase chain reaction*. Gene **77**: 51-59.

103. Thomas M. R. (1994). *Simple, effective cleanup of DNA ligation reactions prior to electro-transformation of E. coli*. Biotechniques **16**: 988-990.

104. Goddard T. and Kneller D. (1997). *SPARKY 3, University of California, San Francisco, CA.*

105. Wüthrich K. (1986) *NMR of Proteins and Nuclei acids*, Wiley, Interscience, New York

106. Cornilescu G., Delaglio F. and Bax A. (1999). *Protein backbone angle restraints from searching a database for chemical shift and sequence homology*. J. Biomol. NMR **13**: 289-302.

107. Güntert P., Mumenthaler C. and Wüthrich K. (1997). *Torsion angle dynamics for NMR structure calculation with the new program D*. J. Mol. Biol. **273**: 283-298.

108. Schäfer I. B., Bailer S. M., Düser M. G., Börsch M., Bernal R. A., Stock D. and Grüber G. (2006). *Crystal Structure of the Archaeal A_1A_O ATP Synthase Subunit B from Methanosarcina mazei Gö1: Implications of Nucleotide-binding Differences in the Major A_1A_O Subunits A and B.* J. Mol. Biol. **358**: 725-740.

109. Basak S., Malathy Sony Subramanian M. and Asha Manikkoth B. (2012). *Crystallization and preliminary X-ray crystallographic analysis of subunit F, $F_{1\text{-}94}$, an essential coupling subunit of the eukaryotic V_1V_O ATPase from Saccharomyces cerevisiae.* Acta Crystallogr. F**68**: 1055-1059.

110. Zbyszek O. and Wladek M. (1997). *Processing of X-ray diffraction data collected in oscillation mode*. Methods Enzymol. **276**: 307-326.

111. Sheldrick G. M., Hauptman H. A., Weeks C. M., Miller R. and Uson I. (2001) *International Tables for Macromolecular Crystallography*, Kluwer Academic Publishers

112. Schneider T. R. and Sheldrick G. M. (2002). *Substructure solution with SHELXD*. Acta Crystallogr. D **58**: 1772-1779.

113. Sheldrick G. M. (2002). *Macromolecular phasing with SHELXE*. Z. Kristallogr. **217**: 644-650.

114. Terwilliger T. (2000). *Maximum-likelihood density modification*. Acta Crystallogr. D **56**: 965-972.

115. Cowtan K. (2006). *The Buccaneer software for automated model building. 1. Tracing protein chains*. Acta Crystallogr. D **62**: 1002-1011.

116. Emsley P., Lohkamp B., Scott W. G. and Cowtan K. (2010). *Features and development of Coot*. Acta Crystallogr. D **66**: 486-501.

117. Murshudov G. N., Skubak P., Lebedev A. A., Pannu N. S., Steiner R. A., Nicholls R. A., Winn M. D., Long F. and Vagin A. A. (2011). *REFMAC5 for the refinement of macromolecular crystal structures*. Acta Crystallogr. D **67**: 355-367.

118. Laskowski R. A., MacArthur M. W., Moss D. S. and Thornton J. M. (1993). *PROCHECK: a program to check the stereochemical quality of protein structures*. J. Appl. Crystallogr. **26**: 283-291.

119. DeLano W. (2002) The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA

120. Krissinel E. and Henrick K. (2004). *Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions*. Acta Crystallogr. D **60**: 2256-2268.

121. Delaglio F., Grzesiek S., Vuister G., Zhu G., Pfeifer J. and Bax A. (1995). *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*. J. Biomol. NMR **6**: 277-293.

122. Cavanagh J., Fairbrother W. J., III A. G. P. and Skelton N. J. (1995). *Protein NMR Spectroscopy: Principles and Practice*

123. Braunschweiler L. and Ernst R. R. (1983). *Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy*. J. Magn. Reson. **53**: 521-528.

124. Wüthrich K. (1990). *Protein structure determination in solution by NMR spectroscopy*. J. Biol. Chem. **265**: 22059-22062.

125. Armbrüster A., Svergun D. I., Coskun Ü., Juliano S., Bailer S. M. and Grüber G. (2004). *Structural analysis of the stalk subunit Vma5p of the yeast V-ATPase in solution*. FEBS Lett. **570**: 119-125.

126. Thaker Y. R., Roessle M. and Grüber G. (2007). *The boxing glove shape of subunit d of the yeast V-ATPase in solution and the importance of disulfide formation for folding of this protein*. J. Bioenerg. Biomembr. **39**: 275-289.

127. Takeuchi K. and Wagner G. (2006). *NMR studies of protein interactions*. Curr. Opin. Struct. Biol. **16**: 109-117.

128. Vaynberg J. and Qin J. (2006). *Weak protein-protein interactions as probed by NMR spectroscopy*. Trends Biotechnol. **24**: 22-27.

129. Kneller D. and Goddard T. (1997). *SPARKY 3.105 edit.* San Francisco, CA, University of California.

130. Pierce M. M., Raman C. S. and Nall B. T. (1999). *Isothermal Titration Calorimetry of Protein-Protein Interactions*. Methods **19**: 213–221.

131. Graham L. A., Flannery A. R. and Stevens T. H. (2003). *Structure and Assembly of the Yeast V-ATPase*. J. Bioenerg. Biomembr. **35**: 301-312.

132. Parsons L. S. and Wilkens S. (2012). *Probing subunit-subunit interactions in the yeast vacuolar ATPase by peptide arrays*. PLoS ONE **7**: e46960.

133. Coskun Ü., Rizzo V. F., Koch M. H. J. and Grüber G. (2004) *Ligand-dependent structural changes in the V_1 ATPase from Manduca sexta*. J. Bioenerg. Biomembr. **36**: 249-256.

134. Matthews B. W. (1968). *Solvent content of protein crystals*. J. Mol. Biol. **33**: 491-497.

135. Kigawa T., Yamaguchi-Nunokawa E., Kodama K., Matsuda T., Yabuki T., Matsuda N., Ishitani R., Nureki O. and Yokoyama S. (2002). *Selenomethionine incorporation into a protein by cell-free synthesis*. J. Struct. Funct. Genomics **2**: 29-35-35.

136. Barton W. A., Tzvetkova-Robev D., Erdjument-Bromage H., Tempst P. and Nikolov D. B. (2006). *Highly efficient selenomethionine labeling of recombinant proteins produced in mammalian cells*. Protein Sci. **15**: 2008-2013.

137. Gassner N. C., Baase W. A., Lindstrom J. D., Lu J., Dahlquist F. W. and Matthews B. W. (1999). *Methionine and Alanine Substitutions Show That the Formation of Wild-Type-like Structure in the Carboxy-Terminal Domain of T4 Lysozyme Is a Rate-Limiting Step in Folding* Biochemistry **38**: 14451-14460.

138. Senda M., Muto S., Horikoshi M. and Senda T. (2008). *Effect of leucine-to-methionine substitutions on the diffraction quality of histone chaperone SET/TAF-I[beta]/INHAT crystals*. Acta Crystallogr. F **64**: 960-965.

139. Hege T. and Baumann U. (2001). *The conserved methionine residue of the metzincins: a site-directed mutagenesis study*. J. Mol. Biol. **314**: 181-186.

140. Basak S., Lim J., Manimekalai M. S. S., Balakrishna A. M. and Grüber G. (2013). *Crystal- and NMR structures give insights into the role and dynamics of subunit F of the eukaryotic V-ATPase from Saccharomyces cerevisiae*. J. Biol. Chem. **288**: 11930-11939.

141. Page R., Peti W., Wilson I. A., Stevens R. C. and Wüthrich K. (2005). *NMR screening and crystal quality of bacterially expressed prokaryotic and eukaryotic proteins in a structural genomics pipeline*. Proc. Natl. Acad. Sci. **102**: 1901-1905.

142. Schäfer I., Rössle M., Biuković G., Müller V. and Grüber G. (2006). *Structural and functional analysis of the coupling subunit F in solution and topological arrangement of the stalk domains of the methanogenic A_1A_O ATP synthase*. J. Bioenerg. Biomembr. **38**: 83-92.

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

144. Saijo S., Arai S., Hossain K. M. M., Yamato I., Suzuki K., Kakinuma Y., Ishizuka-Katsura Y., Ohsawa N., Terada T., Shirouzu M., Yokoyama S., Iwata S. and Murata T. (2011). *Crystal structure of the central axis DF complex of the prokaryotic V-ATPase*. Proc. Natl. Acad. Sci. **108**: 19955-19960.

145. Kane P. M. (2012). *Targeting Reversible Disassembly as a Mechanism of Controlling V-ATPase Activity*. Curr. Protein Pept. Sci. **13**: 117-123.

146. Mertens H. D. T. and Svergun D. I. (2010) *Structural characterization of proteins and complexes using small-angle X-ray solution scattering*. J. Struct. Biol. **172**: 128-141.

147. Makyio H., Iino R., Ikeda C., Imamura H., Tamakoshi M., Iwata M., Stock D., Bernal R. A., Carpenter E. P., Yoshida M., Yokoyama K. and Iwata S. (2005). *Structure of a central stalk subunit F of prokaryotic V-type ATPase/synthase from Thermus thermophilus*. EMBO J. **24**: 3974-3983.

148. Raghunathan D., Gayen S., Grüber G. and Verma C. S. *Crosstalk along the Stalk: Dynamics of the Interaction of Subunits B and F in the A_1A_O ATP Synthase of Methanosarcina mazei Gö1*. Biochemistry **49**: 4181-4190.

149. Svergun D. I., Konrad S., Huss M., Koch M. H. J., Wieczorek H., Altendorf K., Volkov V. V. and Grüber G. (1998). *Quaternary Structure of V_1 and F_1 ATPase: Significance of Structural Homologies and Diversities.* Biochemistry **37**: 17659-17663.

Author's publications during the Ph.D period August 2009 to July 2013

1. **Basak, S.**, Gayen, S., Thaker, Y. R., Manimekalai, M. S. S., Roessle, M., Hunke, C., and Grüber, G. (2011) *Solution structure of subunit F (VMA7p) of the eukaryotic V_1V_O ATPase from Saccharomyces cerevisiae derived from SAXS and NMR spectroscopy.* Biochim. Biophys. Acta-Biomembranes **1808**, 360-368

2. **Basak, S.**, Gayen, S., Ramalingam, J., Grüber, A., Preiser, R. P., and Grüber, G. (2011) *NMR solution structure of $NBD94_{483-502}$ of the nucleotide binding domain of the Plasmodium yoelii reticulocyte binding protein Py235.* FEMS Microbiol. Letters, **318**, 152-158

3. **Basak, S.**, Balakrishna, A. M., Manimekalai, M. S. S., and Grüber, G. (2012) *Crystallization and preliminary X-ray crystallographic analysis of subunit F, F_{1-94}, an essential coupling subunit of the eukaryotic V_1V_O ATPase from Saccharomyces cerevisiae.* Acta Cryst. **F68**, 1055-1059

4. Biuković, G., **Basak, S.**, Manimekalai, M. S. S., Rishikesan, S., Roessle, M., Dick, T., Rao, S., Hunke, C., and Grüber, G. (2013) *Variations of subunit ε of the Mycobacterium tuberculosis F_1F_O ATP synthase and a novel model for mechanism of action of the TB drug TMC207.* Antimicrob. Agents Chemother. **57**, 168-176

5. **Basak, S.**, Lim, J., Manimekalai, M. S. S., Balakrishna, A. M., and Grüber, G. (2013) *Crystal- and NMR structures give insights into the role and dynamics of subunit F of the eukaryotic V-ATPase from Saccharomyces cerevisiae.* J. Biol. Chem. **288**, 11930-11939

Conference attendance

1. "Satellite meeting of the 15th International photosynthesis congress", Singapore (18th to 20th August 2010).

Posters

1. The 6th International Conference on Structural Biology and Functional Genomics, Singapore (6th to 8th December 2010).

   Title: "Solution structure of subunit F of the eukaryotic V_1V_O ATPase from *Saccharomyces cerevisiae*."
2. The 17th European Bioenergetics Conference, Freiburg, Germany (15th to 20th September 2012).

   Title: "First crystallographic structure of the eukaryotic V_1V_O ATPase subunit F from *Saccharomyces cerevisiae*."

Abstracts

1. **Basak, S.**, Balakrishna, A. M., Manimekalai, M. S. S. and Grüber, G. (2012) *First crystallographic structure of the eukaryotic V_1V_O ATPase subunit F from Saccharomyces cerevisiae.*

Awards

1. Received "Travel grant awarded by the 17th European Bioenergetics Conference organizing committee" Freiburg, Germany (15th to 20th September 2012).

Appendix

Chemical shift assignments of subunit F from *S. cerevisiae* V-ATPase

The chemical shift assignment are listed in the NMR-STAR format (Columns are as follows: Atom chemical shift ID, Atom chemical shift compound index ID, Atom chemical shift compound ID, Atom chemical shift atom ID, Atom chemical shift atom type, Atom chemical shift atom isotope number, Atom chemical shift value, Atom chemical shift value error, Atom chemical shift ambiguity code.

1	2	ALA	H	H	1	8.18	0.020	1
2	2	ALA	C	C	13	173.32	0.400	1
3	2	ALA	CA	C	13	49.73	0.400	1
4	2	ALA	CB	C	13	16.11	0.400	1
5	2	ALA	N	N	15	125.55	0.400	1
6	3	GLU	H	H	1	8.29	0.020	1
7	3	GLU	C	C	13	174.97	0.400	1
8	3	GLU	CA	C	13	54.29	0.400	1
9	3	GLU	CB	C	13	26.72	0.400	1
10	3	GLU	N	N	15	120.77	0.400	1
11	7	LEU	H	H	1	7.73	0.020	1
12	7	LEU	C	C	13	171.93	0.400	1
13	7	LEU	CA	C	13	51.86	0.400	1
14	7	LEU	CB	C	13	41.47	0.400	1
15	7	LEU	N	N	15	120.49	0.400	1
16	8	ILE	H	H	1	8.497	0.020	1
17	8	ILE	C	C	13	174.42	0.400	1
18	8	ILE	CA	C	13	-	0.400	1
19	8	ILE	CB	C	13	-	0.400	1
20	8	ILE	N	N	15	121.46	0.400	1
21	9	ALA	H	H	1	8.59	0.020	1
22	9	ALA	C	C	13	171.26	0.400	1
23	9	ALA	CA	C	13	46.50	0.400	1
24	9	ALA	CB	C	13	18.37	0.400	1
25	9	ALA	N	N	15	128.16	0.400	1
26	11	ILE	H	H	1	8.242	0.020	1
27	11	ILE	C	C	13	174.99	0.400	1
28	11	ILE	CA	C	13	-	0.400	1
29	11	ILE	CB	C	13	-	0.400	1
30	11	ILE	N	N	15	117.87	0.400	1
31	12	ALA	H	H	1	8.30	0.020	1
32	12	ALA	C	C	13	170.28	0.400	1
33	12	ALA	CA	C	13	47.71	0.400	1
34	12	ALA	CB	C	13	19.08	0.400	1
35	12	ALA	N	N	15	125.08	0.400	1
36	13	ASP	H	H	1	8.01	0.020	1
37	13	ASP	C	C	13	173.76	0.400	1
38	13	ASP	CA	C	13	49.62	0.400	1
39	13	ASP	CB	C	13	38.91	0.400	1
40	13	ASP	N	N	15	117.63	0.400	1
41	14	GLU	H	H	1	9.31	0.020	1
42	14	GLU	C	C	13	174.32	0.400	1
43	14	GLU	CA	C	13	58.05	0.400	1
44	14	GLU	CB	C	13	27.09	0.400	1
45	14	GLU	N	N	15	123.61	0.400	1
46	15	ASP	H	H	1	8.96	0.020	1
47	15	ASP	C	C	13	176.07	0.400	1
48	15	ASP	CA	C	13	54.88	0.400	1
49	15	ASP	CB	C	13	37.49	0.400	1
50	15	ASP	N	N	15	121.52	0.400	1
51	19	GLY	H	H	1	7.74	0.020	1
52	19	GLY	C	C	13	173.79	0.400	1
53	19	GLY	CA	C	13	44.24	0.400	1
54	19	GLY	N	N	15	108.32	0.400	1
55	20	LEU	H	H	1	8.27	0.020	1
56	20	LEU	C	C	13	173.84	0.400	1
57	20	LEU	CA	C	13	55.02	0.400	1
58	20	LEU	CB	C	13	38.97	0.400	1
59	20	LEU	N	N	15	123.52	0.400	1
60	22	LEU	H	H	1	*7.27*	0.020	1
61	22	LEU	C	C	13	*173.06*	0.400	1
62	22	LEU	CA	C	13	*53.69*	0.400	1
63	22	LEU	CB	C	13	*39.20*	0.400	1
64	22	LEU	N	N	15	*123.70*	0.400	1
65	23	ALA	H	H	1	7.56	0.020	1
66	23	ALA	C	C	13	175.70	0.400	1
67	23	ALA	CA	C	13	50.22	0.400	1
68	23	ALA	CB	C	13	16.11	0.400	1
69	23	ALA	N	N	15	121.00	0.400	1
70	24	GLY	H	H	1	7.81	0.020	1
71	24	GLY	C	C	13	174.86	0.400	1
72	24	GLY	CA	C	13	42.68	0.400	1
73	24	GLY	N	N	15	106.06	0.400	1
74	25	ILE	H	H	1	7.94	0.020	1
75	25	ILE	C	C	13	171.23	0.400	1
76	25	ILE	CA	C	13	62.21	0.400	1
77	25	ILE	CB	C	13	34.85	0.400	1
78	25	ILE	N	N	15	118.04	0.400	1
79	27	GLN	H	H	1	8.03	0.020	1
80	27	GLN	C	C	13	172.73	0.400	1
81	27	GLN	CA	C	13	57.88	0.400	1
82	27	GLN	CB	C	13	33.57	0.400	1
83	27	GLN	N	N	15	119.92	0.400	1
84	28	ILE	H	H	1	8.69	0.020	1
85	28	ILE	C	C	13	171.95	0.400	1
86	28	ILE	CA	C	13	57.93	0.400	1
87	28	ILE	CB	C	13	37.07	0.400	1
88	28	ILE	N	N	15	126.62	0.400	1
89	31	GLU	H	H	1	8.21	0.020	1
90	31	GLU	C	C	13	174.99	0.400	1
91	31	GLU	CA	C	13	-	0.400	1
92	31	GLU	CB	C	13	-	0.400	1
93	31	GLU	N	N	15	117.86	0.400	1
94	33	GLN	H	H	1	8.583	0.020	1
95	33	GLN	C	C	13	173.80	0.400	1
96	33	GLN	CA	C	13	-	0.400	1
97	33	GLN	CB	C	13	-	0.400	1
98	33	GLN	N	N	15	120.67	0.400	1
99	35	LYS	H	H	1	7.82	0.020	1
100	35	LYS	C	C	13	172.55	0.400	1
101	35	LYS	CA	C	13	52.98	0.400	1
102	35	LYS	CB	C	13	-	0.400	1
103	35	LYS	N	N	15	112.53	0.400	1
104	37	PHE	H	H	1	7.96	0.020	1
105	37	PHE	C	C	13	169.69	0.400	1
106	37	PHE	CA	C	13	52.80	0.400	1
107	37	PHE	CB	C	13	36.78	0.400	1
108	37	PHE	N	N	15	119.33	0.400	1
109	38	PHE	H	H	1	9.32	0.020	1
110	38	PHE	C	C	13	172.55	0.400	1
111	38	PHE	CA	C	13	54.27	0.400	1
112	38	PHE	CB	C	13	38.87	0.400	1
113	38	PHE	N	N	15	125.64	0.400	1
114	39	VAL	H	H	1	8.09	0.020	1
115	39	VAL	C	C	13	170.51	0.400	1
116	39	VAL	CA	C	13	57.89	0.400	1
117	39	VAL	CB	C	13	28.85	0.400	1
118	39	VAL	N	N	15	129.96	0.400	1
119	40	TYR	H	H	1	8.13	0.020	1
120	40	TYR	C	C	13	171.68	0.400	1
121	40	TYR	CA	C	13	54.18	0.400	1
122	40	TYR	CB	C	13	38.94	0.400	1
123	40	TYR	N	N	15	130.21	0.400	1
124	41	GLN	H	H	1	8.48	0.020	1
125	41	GLN	C	C	13	170.94	0.400	1

126	41	GLN	CA	C	13	50.93	0.400	1	201	58	THR	CA	C	13	59.99	0.400	1
127	41	GLN	CB	C	13	27.75	0.400	1	202	58	THR	CB	C	13	66.29	0.400	1
128	41	GLN	N	N	15	127.55	0.400	1	203	58	THR	N	N	15	105.45	0.400	1
129	42	GLU	H	H	1	8.98	0.020	1	204	59	GLU	H	H	1	8.12	0.020	1
130	42	GLU	C	C	13	171.38	0.400	1	205	59	GLU	C	C	13	174.49	0.400	1
131	42	GLU	CA	C	13	55.32	0.400	1	206	59	GLU	CA	C	13	54.23	0.400	1
132	42	GLU	CB	C	13	26.12	0.400	1	207	59	GLU	CB	C	13	28.85	0.400	1
133	42	GLU	N	N	15	130.47	0.400	1	208	59	GLU	N	N	15	120.28	0.400	1
134	44	LYS	H	H	1	7.80	0.020	1	209	60	GLU	H	H	1	7.30	0.020	1
135	44	LYS	C	C	13	170.22	0.400	1	210	60	GLU	C	C	13	174.58	0.400	1
136	44	LYS	CA	C	13	54.16	0.400	1	211	60	GLU	CA	C	13	54.71	0.400	1
137	44	LYS	CB	C	13	32.16	0.400	1	212	60	GLU	CB	C	13	26.81	0.400	1
138	44	LYS	N	N	15	120.60	0.400	1	213	60	GLU	N	N	15	118.14	0.400	1
139	45	THR	H	H	1	8.39	0.020	1	214	61	ARG	H	H	1	7.33	0.020	1
140	45	THR	C	C	13	173.52	0.400	1	215	61	ARG	C	C	13	176.12	0.400	1
141	45	THR	CA	C	13	60.92	0.400	1	216	61	ARG	CA	C	13	51.17	0.400	1
142	45	THR	CB	C	13	65.58	0.400	1	217	61	ARG	CB	C	13	28.72	0.400	1
143	45	THR	N	N	15	119.28	0.400	1	218	61	ARG	N	N	15	117.64	0.400	1
144	46	THR	H	H	1	8.94	0.020	1	219	63	ASP	H	H	1	*8.61*	0.020	1
145	46	THR	C	C	13	172.37	0.400	1	220	63	ASP	C	C	13	*173.18*	0.400	1
146	46	THR	CA	C	13	57.70	0.400	1	221	63	ASP	CA	C	13	*49.99*	0.400	1
147	46	THR	CB	C	13	68.96	0.400	1	222	63	ASP	CB	C	13	*37.05*	0.400	1
148	46	THR	N	N	15	120.85	0.400	1	223	63	ASP	N	N	15	*113.46*	0.400	1
149	47	LYS	H	H	1	8.98	0.020	1	224	64	ILE	H	H	1	*7.10*	0.020	1
150	47	LYS	C	C	13	172.83	0.400	1	225	64	ILE	C	C	13	*172.85*	0.400	1
151	47	LYS	CA	C	13	57.46	0.400	1	226	64	ILE	CA	C	13	*56.05*	0.400	1
152	47	LYS	CB	C	13	29.63	0.400	1	227	64	ILE	CB	C	13	*33.80*	0.400	1
153	47	LYS	N	N	15	122.25	0.400	1	228	64	ILE	N	N	15	*119.98*	0.400	1
154	48	GLU	H	H	1	8.96	0.020	1	229	66	ILE	H	H	1	7.24	0.020	1
155	48	GLU	C	C	13	176.80	0.400	1	230	66	ILE	C	C	13	175.06	0.400	1
156	48	GLU	CA	C	13	57.69	0.400	1	231	66	ILE	CA	C	13	56.26	0.400	1
157	48	GLU	CB	C	13	25.60	0.400	1	232	66	ILE	CB	C	13	40.16	0.400	1
158	48	GLU	N	N	15	122.50	0.400	1	233	66	ILE	N	N	15	117.57	0.400	1
159	49	GLU	H	H	1	7.79	0.020	1	234	67	LEU	H	H	1	9.08	0.020	1
160	49	GLU	C	C	13	175.48	0.400	1	235	67	LEU	C	C	13	168.50	0.400	1
161	49	GLU	CA	C	13	56.25	0.400	1	236	67	LEU	CA	C	13	49.48	0.400	1
162	49	GLU	CB	C	13	27.21	0.400	1	237	67	LEU	CB	C	13	41.20	0.400	1
163	49	GLU	N	N	15	120.99	0.400	1	238	67	LEU	N	N	15	129.45	0.400	1
164	50	ILE	H	H	1	7.96	0.020	1	239	68	LEU	H	H	1	*8.94*	0.020	1
165	50	ILE	C	C	13	177.24	0.400	1	240	68	LEU	C	C	13	*170.92*	0.400	1
166	50	ILE	CA	C	13	62.21	0.400	1	241	68	LEU	CA	C	13	*50.11*	0.400	1
167	50	ILE	CB	C	13	34.85	0.400	1	242	68	LEU	CB	C	13	*41.06*	0.400	1
168	50	ILE	N	N	15	117.67	0.400	1	243	68	LEU	N	N	15	*127.76*	0.400	1
169	51	THR	H	H	1	8.28	0.020	1	244	71	GLN	H	H	1	7.82	0.020	1
170	51	THR	C	C	13	174.78	0.400	1	245	71	GLN	C	C	13	176.24	0.400	1
171	51	THR	CA	C	13	65.42	0.400	1	246	71	GLN	CA	C	13	57.20	0.400	1
172	51	THR	CB	C	13	-	0.400	1	247	71	GLN	CB	C	13	29.13	0.400	1
173	51	THR	N	N	15	118.12	0.400	1	248	71	GLN	N	N	15	118.52	0.400	1
174	53	LYS	H	H	1	7.65	0.020	1	249	73	ILE	H	H	1	6.19	0.020	1
175	53	LYS	C	C	13	176.57	0.400	1	250	73	ILE	C	C	13	173.73	0.400	1
176	53	LYS	CA	C	13	53.94	0.400	1	251	73	ILE	CA	C	13	61.11	0.400	1
177	53	LYS	CB	C	13	26.87	0.400	1	252	73	ILE	CB	C	13	34.84	0.400	1
178	53	LYS	N	N	15	120.68	0.400	1	253	73	ILE	N	N	15	121.40	0.400	1
179	54	PHE	H	H	1	8.90	0.020	1	254	74	ALA	H	H	1	7.72	0.020	1
180	54	PHE	C	C	13	174.70	0.400	1	255	74	ALA	C	C	13	174.62	0.400	1
181	54	PHE	CA	C	13	60.01	0.400	1	256	74	ALA	CA	C	13	51.61	0.400	1
182	54	PHE	CB	C	13	36.45	0.400	1	257	74	ALA	CB	C	13	13.74	0.400	1
183	54	PHE	N	N	15	121.83	0.400	1	258	74	ALA	N	N	15	122.46	0.400	1
184	55	ASN	H	H	1	8.96	0.020	1	259	75	GLU	H	H	1	8.10	0.020	1
185	55	ASN	C	C	13	-	0.400	1	260	75	GLU	C	C	13	176.26	0.400	1
186	55	ASN	CA	C	13	53.32	0.400	1	261	75	GLU	CA	C	13	55.15	0.400	1
187	55	ASN	CB	C	13	34.89	0.400	1	262	75	GLU	CB	C	13	26.50	0.400	1
188	55	ASN	N	N	15	121.60	0.400	1	263	75	GLU	N	N	15	115.11	0.400	1
189	56	HIS	H	H	1	8.36	0.020	1	264	76	ASN	H	H	1	7.23	0.020	1
190	56	HIS	C	C	13	174.63	0.400	1	265	76	ASN	C	C	13	176.12	0.400	1
191	56	HIS	CA	C	13	56.62	0.400	1	266	76	ASN	CA	C	13	52.68	0.400	1
192	56	HIS	CB	C	13	25.30	0.400	1	267	76	ASN	CB	C	13	36.31	0.400	1
193	56	HIS	N	N	15	122.59	0.400	1	268	76	ASN	N	N	15	117.03	0.400	1
194	57	PHE	H	H	1	7.87	0.020	1	269	77	ILE	H	H	1	7.12	0.020	1
195	57	PHE	C	C	13	173.65	0.400	1	270	77	ILE	C	C	13	173.50	0.400	1
196	57	PHE	CA	C	13	55.22	0.400	1	271	77	ILE	CA	C	13	56.43	0.400	1
197	57	PHE	CB	C	13	36.61	0.400	1	272	77	ILE	CB	C	13	34.87	0.400	1
198	57	PHE	N	N	15	112.73	0.400	1	273	77	ILE	N	N	15	111.97	0.400	1
199	58	THR	H	H	1	7.48	0.020	1	274	78	ARG	H	H	1	6.50	0.020	1
200	58	THR	C	C	13	173.09	0.400	1	275	78	ARG	C	C	13	173.35	0.400	1

276	78	ARG	CA	C	13	55.15	0.400	1
277	78	ARG	CB	C	13	25.45	0.400	1
278	78	ARG	N	N	15	123.46	0.400	1
279	79	ALA	H	H	1	8.36	0.020	1
280	79	ALA	C	C	13	174.22	0.400	1
281	79	ALA	CA	C	13	51.89	0.400	1
282	79	ALA	CB	C	13	14.49	0.400	1
283	79	ALA	N	N	15	118.89	0.400	1
284	80	ARG	H	H	1	7.54	0.020	1
285	80	ARG	C	C	13	177.94	0.400	1
286	80	ARG	CA	C	13	54.01	0.400	1
287	80	ARG	CB	C	13	25.96	0.400	1
288	80	ARG	N	N	15	119.44	0.400	1
289	81	VAL	H	H	1	7.78	0.020	1
290	81	VAL	C	C	13	175.95	0.400	1
291	81	VAL	CA	C	13	63.72	0.400	1
292	81	VAL	CB	C	13	28.34	0.400	1
293	81	VAL	N	N	15	121.98	0.400	1
294	82	ASP	H	H	1	8.81	0.020	1
295	82	ASP	C	C	13	176.50	0.400	1
296	82	ASP	CA	C	13	53.67	0.400	1
297	82	ASP	CB	C	13	37.09	0.400	1
298	82	ASP	N	N	15	119.78	0.400	1
299	83	SER	H	H	1	7.43	0.020	1
300	83	SER	C	C	13	175.15	0.400	1
301	83	SER	CA	C	13	55.96	0.400	1
302	83	SER	CB	C	13	60.79	0.400	1
303	83	SER	N	N	15	112.60	0.400	1
304	84	PHE	H	H	1	7.62	0.020	1
305	84	PHE	C	C	13	170.58	0.400	1
306	84	PHE	CA	C	13	56.14	0.400	1
307	84	PHE	CB	C	13	35.89	0.400	1
308	84	PHE	N	N	15	125.74	0.400	1
309	85	THR	H	H	1	8.227	0.020	1
310	85	THR	C	C	13	172.96	0.400	1
311	85	THR	CA	C	13	-	0.400	1
312	85	THR	CB	C	13	-	0.400	1
313	85	THR	N	N	15	117.90	0.400	1
314	86	ASN	H	H	1	5.58	0.020	1
315	86	ASN	C	C	13	171.78	0.400	1
316	86	ASN	CA	C	13	49.65	0.400	1
317	86	ASN	CB	C	13	35.36	0.400	1
318	86	ASN	N	N	15	120.81	0.400	1
319	88	PHE	H	H	1	7.47	0.020	1
320	88	PHE	C	C	13	173.49	0.400	1
321	88	PHE	CA	C	13	51.19	0.400	1
322	88	PHE	CB	C	13	38.88	0.400	1
323	88	PHE	N	N	15	111.14	0.400	1
324	90	ALA	H	H	1	7.96	0.020	1
325	90	ALA	C	C	13	172.61	0.400	1
326	90	ALA	CA	C	13	49.32	0.400	1
327	90	ALA	CB	C	13	15.96	0.400	1
328	90	ALA	N	N	15	123.89	0.400	1
329	91	ILE	H	H	1	8.60	0.020	1
330	91	ILE	C	C	13	173.39	0.400	1
331	91	ILE	CA	C	13	57.10	0.400	1
332	91	ILE	CB	C	13	34.61	0.400	1
333	91	ILE	N	N	15	123.90	0.400	1
334	94	ILE	H	H	1	8.25	0.020	1
335	94	ILE	C	C	13	171.85	0.400	1
336	94	ILE	CA	C	13	57.11	0.400	1
337	94	ILE	CB	C	13	38.51	0.400	1
338	94	ILE	N	N	15	123.85	0.400	1
339	96	SER	H	H	1	8.91	0.020	1
340	96	SER	C	C	13	174.44	0.400	1
341	96	SER	CA	C	13	-	0.400	1
342	96	SER	CB	C	13	65.527	0.400	1
343	96	SER	N	N	15	120.82	0.400	1
344	98	ASP	H	H	1	8.15	0.020	1
345	98	ASP	C	C	13	173.71	0.400	1
346	98	ASP	CA	C	13	54.23	0.400	1
347	98	ASP	CB	C	13	38.31	0.400	1
348	98	ASP	N	N	15	120.46	0.400	1
349	99	HIS	H	H	1	8.17	0.020	1
350	99	HIS	C	C	13	171.82	0.400	1
351	99	HIS	CA	C	13	51.30	0.400	1
352	99	HIS	CB	C	13	28.98	0.400	1
353	99	HIS	N	N	15	120.56	0.400	1
354	101	TYR	H	H	1	8.26	0.020	1
355	101	TYR	C	C	13	172.57	0.400	1
356	101	TYR	CA	C	13	54.39	0.400	1
357	101	TYR	CB	C	13	37.22	0.400	1
358	101	TYR	N	N	15	121.87	0.400	1
359	102	ASP	H	H	1	9.03	0.020	1
360	102	ASP	C	C	13	173.59	0.400	1
361	102	ASP	CA	C	13	-	0.400	1
362	102	ASP	CB	C	13	-	0.400	1
363	102	ASP	N	N	15	126.42	0.400	1
364	104	GLU	H	H	1	8.13	0.020	1
365	104	GLU	C	C	13	175.30	0.400	1
366	104	GLU	CA	C	13	54.61	0.400	1
367	104	GLU	CB	C	13	26.41	0.400	1
368	104	GLU	N	N	15	117.65	0.400	1
369	106	ASP	H	H	1	8.20	0.020	1
370	106	ASP	C	C	13	173.14	0.400	1
371	106	ASP	CA	C	13	50.95	0.400	1
372	106	ASP	CB	C	13	38.81	0.400	1
373	106	ASP	N	N	15	122.85	0.400	1
374	107	SER	H	H	1	7.703	0.020	1
375	107	SER	C	C	13	62.421	0.400	1
376	107	SER	CA	C	13	56.16	0.400	1
377	107	SER	CB	C	13	-	0.400	1
378	107	SER	N	N	15	119.84	0.400	1
379	110	LYS	H	H	1	*8.57*	0.020	1
380	110	LYS	C	C	13	*171.65*	0.400	1
381	110	LYS	CA	C	13	*51.579*	0.400	1
382	110	LYS	CB	C	13	*27.15*	0.400	1
383	110	LYS	N	N	15	*124.35*	0.400	1
384	111	ARG	H	H	1	-	0.020	1
385	112	VAL	H	H	1	8.18	0.020	1
386	112	VAL	C	C	13	176.17	0.400	1
387	112	VAL	CA	C	13	63.44	0.400	1
388	112	VAL	CB	C	13	28.72	0.400	1
389	112	VAL	N	N	15	120.01	0.400	1
390	114	LYS	H	H	1	7.59	0.020	1
391	114	LYS	C	C	13	176.27	0.400	1
392	114	LYS	CA	C	13	55.23	0.400	1
393	114	LYS	CB	C	13	29.06	0.400	1
394	114	LYS	N	N	15	118.90	0.400	1
395	115	LEU	H	H	1	7.70	0.020	1
396	115	LEU	C	C	13	176.45	0.400	1
397	115	LEU	CA	C	13	54.02	0.400	1
398	115	LEU	CB	C	13	38.89	0.400	1
399	115	LEU	N	N	15	120.82	0.400	1
400	116	PHE	H	H	1	7.69	0.020	1
401	116	PHE	C	C	13	175.47	0.400	1
402	116	PHE	CA	C	13	55.16	0.400	1
403	116	PHE	CB	C	13	36.60	0.400	1
404	116	PHE	N	N	15	116.65	0.400	1
405	117	GLY	H	H	1	7.83	0.020	1
406	117	GLY	C	C	13	173.36	0.400	1
407	117	GLY	CA	C	13	42.66	0.400	1
408	117	GLY	N	N	15	109.04	0.400	1
409	118	GLU	H	H	1	7.91	0.020	1
410	118	GLU	C	C	13	170.50	0.400	1
411	118	GLU	CA	C	13	54.81	0.400	1
412	118	GLU	CB	C	13	28.17	0.400	1
413	118	GLU	N	N	15	126.23	0.400	1