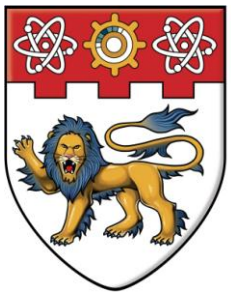


NANYANG TECHNOLOGICAL UNIVERSITY

THE STALKS OF THE MOTOR A_1A_O ATP SYNTHASE COME INTO THE FOCUS: STRUCTURAL INSIGHT INTO SUBUNITS E, F AND H

SHOVANLAL GAYEN

SCHOOL OF BIOLOGICAL SCIENCES

2010

The stalks of the motor A_1A_O ATP synthase come into the focus: structural insight into subunits E, F and H

Shovanlal Gayen

SCHOOL OF BIOLOGICAL SCIENCES

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

2010

Acknowledgements

It is with my profound gratitude and thankfulness I wish to acknowledge the help of my supervisor A/Prof. Dr. Gerhard Grüber in my research life. 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 wish to express also my deep appreciations and indebtedness to A/Prof. Gerhard Grüber for introducing me to the wonderful and interesting field of ATP synthase.

I would especially like to thank A/Prof. Yoon Ho Sup for his helpful discussions and suggestions and also for providing me the facilities initially in his laboratory during the solution structure determination of subunit F. I must also give my heartfelt thanks to A/Prof. Konstantin Pervushin for his help, giving new ideas and suggestions during structural characterization of the protein H_{1-47} by NMR spectroscopy. I am indebted to them for their time and patience, in explaining concepts to me.

I would also like to take this opportunity to thank Dr. Subramanian Vivekanandan to train me initially the field of biomolecular NMR during solution structure determination of subunit F. I am very grateful to him for his valuable knowledge, suggestions on NMR throughout the entire tenure of my Ph.D. Many special thanks to Dr. Goran Biuković for helping me to learn the basic techniques of Molecular Biology. This has been a wonderful experience to work with them.

I am very thankful to Dr. Devanathan Raghunathan and Adjunct Assoc. Prof. Chandra S. verma (Bioinformatics Institute (A*STAR), Singapore) for the molecular dynamics study on subunit F. I am also very thankful to Dr. Asha Balakrishna for cloning and purification of E_{1-100}, $E_{101-206}$ and E_{1-52} constructs.

I also give my heartfelt thanks to Dr. Sathan Thennarasu (University of Michigan, USA) for providing me the support when I needed. A vital note of appreciations is due to my other laboratory colleagues Dr. Cornelia Hunke, Dr. Malathy Sony Subramanian Manimekalai, Dr. Ardina Grüber, Dr. Toh Yew Kwang, Dr. Kamariah Neelagandan, Dr. Youg Raj Thaker, Mr. Anil Kumar, Mr. Vikeramjit Singh Tadwal, Mr. Sankaranarayanan Rishikesan, Mrs Ragunathan Priya, Mr. Sandip Basak, Mr. Phat Vinh Dip and Mrs V.P. Hema Priya for their cooperation, adjustment and encouragement during this long period of research. I am very thankful to my supervisor A/Prof. Gerhard Grüber for critical comments and reading my thesis several times and my colleagues, who helped me in thesis correction.

I am grateful to the School of Biological Sciences, Nanyang Technological University, Singapore for providing a Ph.D studentship without which this work would not have been possible. The research was also supported by A*STAR, Biomedical Research Council (06/1/22/19/467).

I dedicate this thesis to my father and mother whom I owe all that I have become today. I thank also my brother 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

Contents

List of figures

Introduction

Materials and methods

Results

Discussion

List of tables

Abbreviations

CPMG	Carr-Purcell-Meiboom-Gill
CSA	Chemical shift anisotropy
hetNOE	Heteronuclear Overhauser effect
HSQC	Heteronuclear single quantum correlation
BMRB	Biological magnetic resonance bank
CSI	Chemical shift index
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
NMR	Nuclear magnetic resonance
PDB	Protein data bank
Ppm	Parts per million
RMSD	Root mean square deviation
R_1	Longitudinal or spin-lattice relaxation rate constant
R_2	Transverse or spin-spin relaxation rate constant
TOCSY	Total correlation spectroscopy
ATP	Adenosine triphosphate
ADP	Adenosine diphosphate
BSA	Bovine Serum Albumin
CD	Circular Dichroism
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
FPLC	Fast protein liquid chromatography
IPTG	Isopropyl-β-D-thiogalactoside
NTA	Nitrilotriacetic acid
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
SAXS	Small Angle X-ray Scattering
SDS	Sodium dodecyl sulphate
Tris	Tris-(hydroxymethyl) aminomethane
A-ATP synthase	Archaeal ATP synthase

Abstract

A_1A_O adenosine triphosphate (ATP) synthases from archaea represent the second class of ATP synthases due to their unusual structural and mechanistic features, which may be due to archaea's unique phylogenetic position in the tree of life. The enzyme uses both H^+ and Na^+ ion gradients across the membrane to synthesize ATP. A_1A_O ATP synthases comprises subunits in the stoichiometry of A_3:B_3:C:D:E_x:F:H_2:*a*:c_x. The water soluble A_1 domain (A_3:B_3) has a pseudohexagonal arrangement of major subunits A and B and it is responsible for ATP synthesis or hydrolysis. The energy provided for or released during the process of ATP hydrolysis or synthesis is transmitted via the stalk parts to the membrane-bound A_O domain to facilitate ion conduction. The stalk parts have an important role for the regulation of this class of enzymes and are separated into a central part (C:D:F) and peripheral parts (E_x:H_2:a_N). In order to get a detailed picture, structural characterization has been done for the central stalk subunit F and the peripheral stalk subunits H and E. The solution structure of the central stalk subunit F was determined by NMR spectroscopy, revealing a distinct two-domain structure, with the N-terminal 78 residues and a flexible C-terminal part formed by the residues 79-101. The two domains are loosely associated, providing new insights into energy coupling between A_O and A_1.

The subunit H is a component of one of the two peripheral stalks connecting the upper water soluble A_1 ATPase and the membrane bound subcomplex A_O. Small angle X-ray scattering (SAXS) was used to determine the first low-resolution structure of this molecule in solution. The protein is dimeric and has a boomerang-like shape, composed of two arms of 12.0 and 6.8 nm in length. NMR analysis together with cross-linking and mutagenesis experiments has shown that subunit H is divided into a C-terminal coiled-coil domain and an N-terminal domain formed by adjacent helices. The solution structure of the peptide comprising the last twenty amino acids of subunit H, $H_{85\text{-}104}$, and cross linking analysis at the C-terminus of entire H subunit indicate that the C-terminus is important for the structure formation of subunit H. To understand how the N-terminal domain of H is formed, the solution NMR structure of $H_{1\text{-}47}$ was solved, revealing an α-helix between residues 15–42 and a flexible N-terminal stretch. The α-helix includes a kink that would bring the two helices of the C-terminus into the coiled-coil arrangement.

Subunit E is another essential peripheral stalk subunit in A_1A_O ATP synthase. To get into the details, the N-terminal structure of subunit E, the truncated construct $E_{1–52}$ of *M. jannaschii* A_1A_O ATP synthase was solved by NMR spectroscopy. The structure indicates a possible helix-helix

interaction with subunit H. NMR titration experiments were performed between full length / truncated domains of subunits E, H and F. The data provide for the first time, structural insights of subunits E, H and F and their arrangement within the A_1A_O ATP synthase.

1. Introduction

1.1 Overview on archaea

Archaea are considered to be among the oldest organisms on the planet. They form a separate evolutionary kingdom located relatively near to the root of the evolutionary tree (SCHÄFER *et al.*, 1999; DEPPENMEIER, 2002). Figure 1.1 shows that archaea are located in a distinct evolutionary branch from eukarya and bacteria. It was speculated that they may have a separate energy conserving mechanism from eukarya or bacteria and may represent an efficient form of energy conserving mechanism. In addition, they can live under extremely harsh conditions, at temperatures above 100 °C, tolerance of extreme pH values (pH 1-12) and salt concentrations (up to 5M) (MÜLLER *et al.*, 2005). These facts have been increasing the interest about archaea and their special kind of cellular adaptation mechanism that can stabilize their

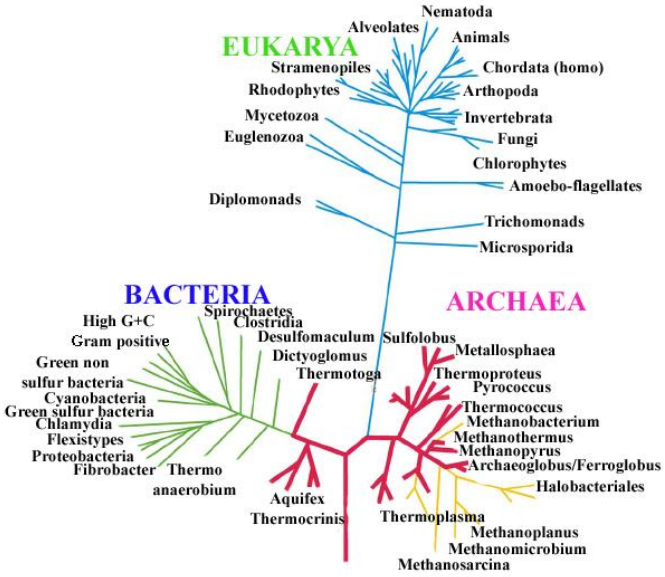


Figure 1.1: The phylogenic tree of life (LEWALTER *et al.*, 2006).

proteins in extremely hot, acidic, alkaline or saline environments. They have also unique metabolic pathways utilising low energy substrates. However, they have a wide variety of energy transducing mechanisms. Respiration, anaerobic respiration with nitrate, fermentation, anaerobic photorespiration, sodium motive methyl transfer reactions and proton reducing hydrogenase are all found in archaea (MÜLLER *et al.*, 2005). There are phylogenitically three distinct groups of archaea: the Crenaarchaeota (extremely thermophilic and thermoacidophilic organisms), the Euryarchaeota (e.g. methanogenic and extremely halophilic bacteria) and the Korarchaeota

(unculturable microbes from terrestrial hot springs) (DEPPENMEIER, 2002). Archaea such as methanogens and halobacteria have chemiosmotic mechanisms of energy conservation. The process of chemiosmosis is characterized by the formation of electrochemical potential of hydrogen ions or sodium ions (SCHÄFER *et al.*, 1999). The process of chemiosmosis can be represented by the following equation for the proton motive force:

$$\Delta\mu_{H}^{+} = RT \cdot \ln([H^{+}_{i}]/[H^{+}_{O}]) + F \cdot \Delta\Psi$$

where $\Delta\mu_{H+}$ is the proton gradient of electrochemical potential, F is the Faraday constant, $\Delta\Psi$ is the membrane potential, R is the gas constant, T is the absolute temperature, and the subscripts "*i*" and "*o*" refer to the inside and outside respectively, of the cell or other membrane-enclosed compartment. In archaea the electrochemical potential of hydrogen or sodium ions can be translocated by redox systems, by methyl transfer reaction as in methanogenesis or by light as in photophosphorylating halobacteria (SCHÄFER *et al.*, 1999).

Methanogenesis is a unique pathway carried out by a group of strictly anaerobic archaea. They convert hydrogen and carbon dioxide to methane. Studies on methanogenesis have revealed that biological production of methane occurs via a series of reactions involving unique enzymes, cofactors and electron carriers found only in methanogens (Figure 1.2). The process is accompanied by a free energy change of -131 KJ/mol CH_4 allowing the synthesis of two mol of ATP under standard conditions (DEPPENMEIER, 2002). In nature, less than one third of an ATP can be produced per mol of CH_4 due to the marginal hydrogen concentration in the atmosphere. Thus, methanogens live close to the thermodynamic limit and have a unique energy conservation mechanism by using both sodium and proton gradients (DEPPENMEIER, 2002) across the membrane for the synthesis of ATP by a unique class of ATP synthase, the A-ATP synthases. These energy producers have unusual structural and functional features which make them a

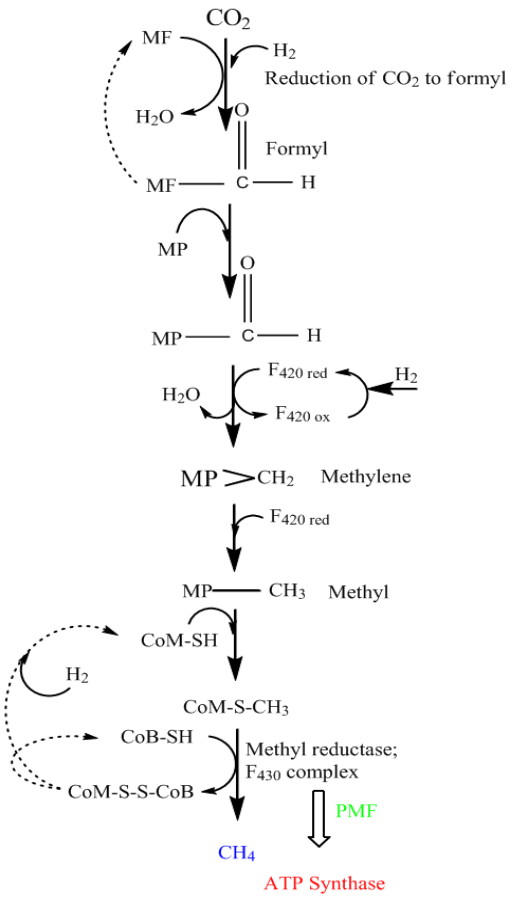


Figure 1.2: The process of methanogenesis in archaea (RUDOLF *et al.*, 2008).

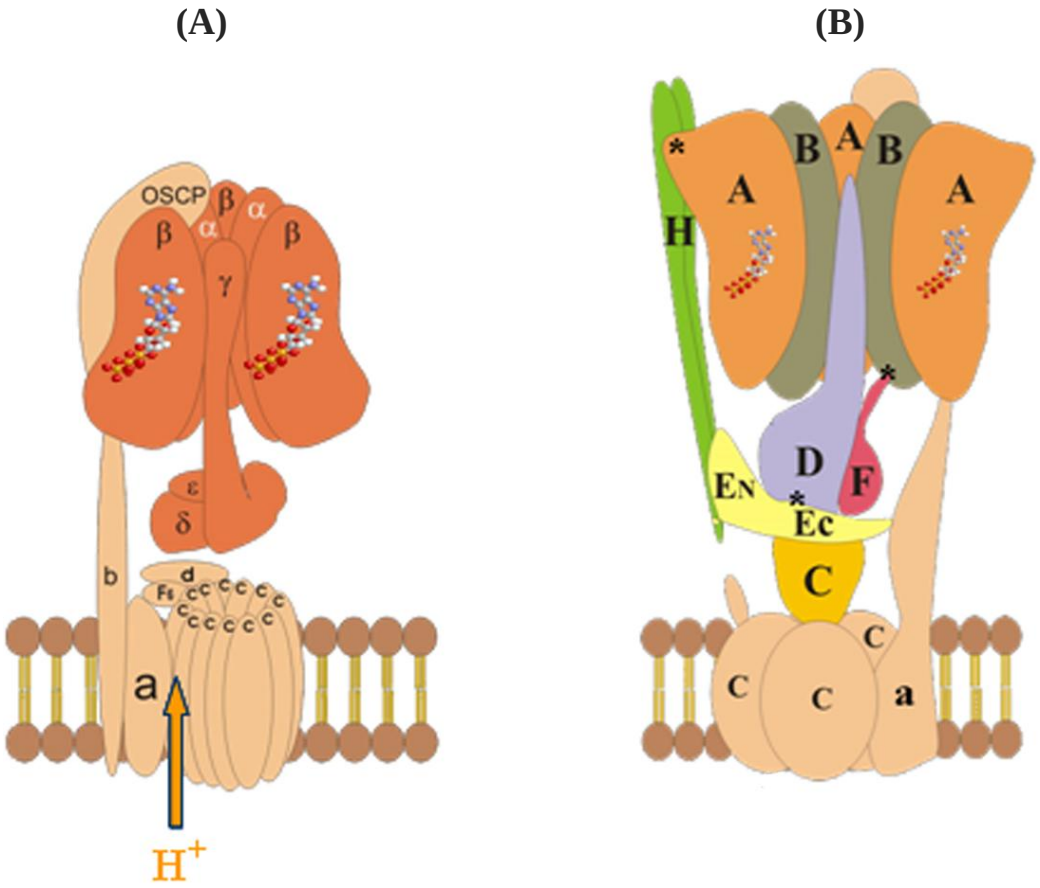


Figure 1.3: The arrangement of subunits in the F-ATP synthase (A) (MÜLLER *et al.*, 2003) and A-ATP synthase (B) (GRÜBER *et al.*, 2008). One α subunit has been removed from the F_1 section to reveal the subunit γ within the $\alpha_3\beta_3$ headpiece. The same has been done in case of A-ATP synthse where one B subunit was removed to reveal subunit D. The A subunit is attached to subunit N-terminal domain (E_N) by a peripheral stalk composed of subunit H, whereby the C-terminal part of E (E_C) is in close proximity to the coupling subunit D. *, show some of the cross-links that have been generated to probe the function and location of these subunits (SCHÄFER *et al.*, 2006b).

distinct class of enzymes from F_1F_O ATP synthase (F-ATP synthase) of bacteria, chloroplasts and mitochondria (MÜLLER *et al.*, 2003). A-ATP synthases are only functionally realted to F-ATP synthases as both of them can synthesize ATP (GRÜBER, 2003; GRÜBER *et al.*, 2008), but in terms of amino acid sequences of major subunits and overall subunit composition (Figure 1.3) they are different from F_1F_O ATP synthases.

1.2 Overview on ATP synthases

1.2.1 F-ATP synthases

F_1F_O ATP synthases are the best understood enzymes in the family of ATP synthases. They are found in the plasma membrane of bacteria, chloroplast thylakoid membrane in plants, mitochondrial inner membrane in animals, plants and also in the plasma membrane of human endothelial cells (CAPALDI *et al.*, 2002). The 550-650 kDa complex has eight subunits in

prokaryotes and in mammals it is composed of 16-18 subunits (CAPALDI *et al.,* 2002). The additional subunits are seen in generally stalk region of the enzyme. F_1F_O ATP synthases are composed of a soluble F_1 domain and a membrane bound F_O domain. The F_1 part has five subunits in the stoichiometry of α_3:β_3:γ:δ:ε. The α_3:β_3 head piece of the F_1 domain contains three catalytic nucleotide binding sites of β subunits at the interface of the adjacent α subunit. ATP synthesis / hydrolysis occurs in the α_3:β_3 sector and energy provided for or released is transferred to the F_O domain. Subunits *a*, *b*, *c* in the stoichiometry 1:2:10-14 form the F_O part, which is responsible for mediating the ion flow through the cytoplasmic membrane (STOCK *et al.,* 2000).

The F_1 and F_O part are connected by a central stalk, consisting of subunits γ and ε, which forms the key rotary element in the enzymes catalytic mechanism. Crystallographic models show that the coiled-coil structure of γ is located asymmetrically in the α_3:β_3 headpiece (HAUSRATH *et al.,* 1999; GIBBONS *et al.,* 2000). The bottom of γ subunit is in contact with the external loop of subunit *c*. The second rotary subunit ε consists of a β sandwich and a helix-loop-helix domain. It is reported that the subunit ε of F-ATP synthase plays dual functions in the coupling of catalysis to proton transport as well as regulation of catalysis events (FENIOUK *et al.,* 2006; KATO *et al.,* 2007; YAGI *et al.,* 2007; IINO *et al.,* 2009; TSUMURAYA *et al.,* 2009). The N-terminal β sandwich domain connects the *c*-oligomer to the γ subunit, whereas α-helical C-terminus has the regulatory role in ATP synthesis by adopting either a closed or extended conformation (Figure 1.4) (FENIOUK *et al.,* 2006; YAGI *et al.,* 2007; TSUMURAYA *et al.,* 2009). In the extended conformation, the C-terminal domain of subunit ε interacts with the $\alpha_3\beta_3$ hexamer, thereby inhibits the catalysis. The interaction is caused by the positively charged residues in C-terminal domain of ε with the negatively charged acid residues of βDELSEED segment in the C-terminus of subunit β (FENIOUK *et al.,* 2006). Crystal structure refined at 2.1 Å resolution of a complex of subunit ε and the central domain of the subunit γ (residues 11-258) sheds a new light on how the catalytic site events in the F_1 part are coupled to proton translocation in the F_O part (RODGERS *et al.,* 2000). In the complex structure, the β-sandwich domain of subunit ε is aligning parallel to the long pair of coiled-coil α-helices consisting of N-terminal residues 12-56 and C-terminal residues 212-247 and plays an important role in the binding of γ-ε complex to subunit *c* ring. Together with the ecentric rotation of the coiled-coil helices and the helix-sheet-helix domain of subunit γ, which is in the opposite side of the coiled-coil from β-sandwich domain of subunit ε, may influence the conformations and nucleotide binding affinities at the catalytic sites.

The peripheral stalk consists of subunits δ and *b* and functions as a stator unit, to experience the elastic strain imposed by the rotation (SENIOR *et al.*, 2006; WALKER *et al.*, 2006; WEBER, 2006). In *E. coli* two copies of *b* subunit and one copy of δ subunit form the

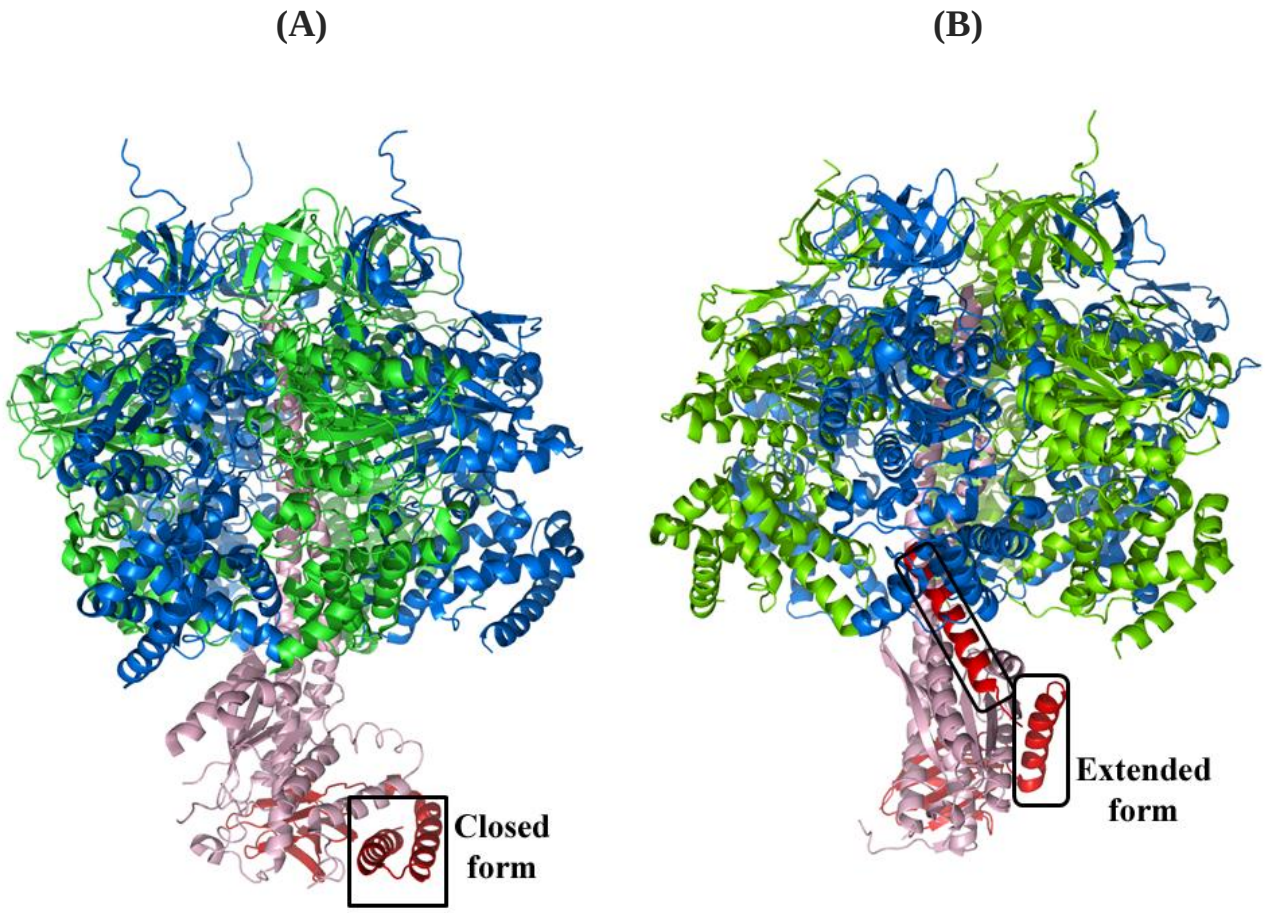


Figure 1.4: Crystal structures of the F_1 ATPase (subunit α, β and γ are in blue, green and gray color respectively) showing two diffrenet conformations of subunit ε (red): closed form (PDB code: 1E79) (GIBBONS *et al.*, 2000) (A) and extended form (PDB code: 1JNV) (HAUSRATH *et al.*, 2001) (B). The C-terminal α-helices of the ε subunit (enclosed by black lines).

peripheral stalk. The elongated shape and helical character of the *b* subunit in dimeric form have been explained experimentally in a number of studies (DEL RIZZO *et al.*, 2006; WEBER, 2006; WOOD *et al.*, 2007; PRIYA *et al.*, 2008). It extends from the top of the F_1 domain into and across the membrane to associate with subunit *a* (DUNN *et al.*, 2000; BHATT *et al.*, 2005; WEBER, 2006). The soluble part of the subunit *b* has three distinct domains: first the tether domain corresponding to the region between the membrane and beneath the F_1 sector; next, the dimerization domain, which extends up to the interface of subunits α and β; third, the C-terminal F_1 binding domain required for the interaction of the peripheral stalk and F_1. The crystallographic structure of the major part of the dimerization domain, $b_{62\text{-}122}$, revealed an α-

helix with a length of 9.0 nm (DEL RIZZO *et al.,* 2006). Small angle X-ray scattering and analytical ultracentrifuge studies indicated that the soluble domain of subunit *b* is dimeric in solution (PRIYA *et al.,* 2008). Subunit *b* is described to be connected to the δ subunit via the C-termini (RODGERS *et al.,* 1997; MCLACHLIN *et al.,* 2000). Most recently, the NMR solution structure of the segments $b_{30\text{-}82}$ and $b_{140\text{-}156}$, which interacts with the C-terminus of subunit δ (δ91-177), has been determined by NMR spectroscopy (PRIYA *et al.,* 2008; PRIYA *et al.,* 2009). Structural features of the N-terminal domain of subunit δ (residues 1-134) available from NMR studies, show a six helix bundle (WILKENS *et al.,* 1997). Immuno-electron microscopy and NMR titration experiments showed that subunit δ sits at the top of the F_1 near the N-termini of subunits α and β, and holds the F_1 with respect to F_O part in proper position for the efficient rotational energy coupling (WILKENS *et al.,* 1997; WILKENS *et al.,* 2005).

1.2.2 A-ATP synthase

The A_1A_O ATP synthases are made of ten subunits in the stoichiometry of A_3:B_3:C:D:E_x:F:G:H_2:*a*:c_x (GRÜBER *et al.,* 2001; MÜLLER *et al.,* 2003; GRÜBER *et al.,* 2008). Like F_1F_O ATP synthase, they contain a water soluble (A_1) domain and a membrane anchored (A_O) domain (MÜLLER *et al.,* 2003; WEBER, 2006; GRÜBER *et al.,* 2008). ATP synthesis and ATP cleavage occurs in the A_3B_3 headpiece of the A_1 sector. The energy provided for or released during the process of ATP hydrolysis or synthesis is provided to the membrane-bound A_O domain to facilitate ion conduction. The stalk parts are responsible for the energy coupling between the catalytic site events in the A_3B_3 hexamer with the ion conduction through A_O domain (MÜLLER *et al.,* 2003; GRÜBER *et al.,* 2008; VONCK *et al.,* 2009).

1.2.2.1 Structure of A_1A_O ATP synthase by electron microscopy from *M. jannaschii*

Despite several attempts in the past years, the structure of an intact A_1A_O was determined only in 2004 by electron microscopy by Coskun *et al*. The first two dimensional projection of the enzyme (Figure 1.5) revealed several important structural basics toward a detailed understanding of the functions of the enzyme (COSKUN *et al.,* 2004). All of the projections from the microscopy clearly indicated the tripartite structure. There is a headpiece, a membrane embedded domain and a connecting stalk region. The connecting stalk region consists of a central stalk and two peripheral stalks. There is a hexagonal and alternate arrangement of subunit A and subunit B densities. The most important organizational feature for the hexagon (10.2 nm in diameter and 10.8 nm in height) is asymmetric from the top. One of the interesting features in the structure of

A_3B_3 head piece is the knob like structures formed by the N-terminal non homologous region (NHR; 80-90 amino acids) of the three catalytic A subunits. These knob like structures are

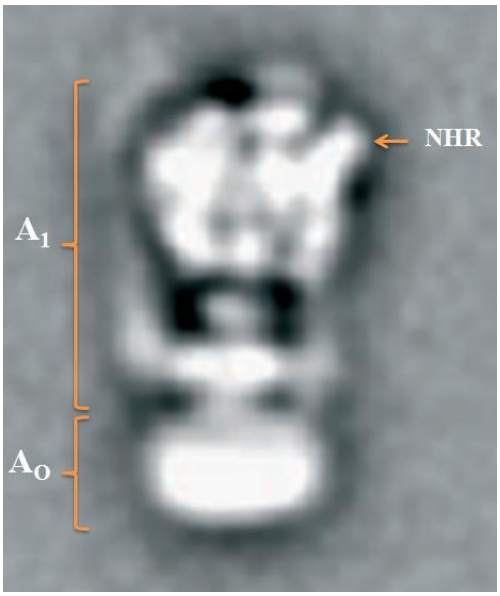


Figure 1.5: Structure of A_1A_O ATP synthase by electron microscopy. 2-D Projection of the of the A_1A_O complex from *M. jannaschii* (COSKUN *et al.,* 2004).

believed to connect the peripheral stalk to the A_3B_3 headpiece. They are absent in the amino acid sequence of homologous β subunits of F-ATP synthases (WILKENS *et al.,* 1998; KARRASCH *et al.,* 1999; BÖTTCHER *et al.,* 2000). The length of the central stalk is around 8 nm. More insight in the structure reveals a collar structure, which interacts with the two peripheral stalks. This structure together with some biochemical analysis resulted in a model in which the subunits C, D, and F form the central stalk domain (Figure 1.3B). The small angle X-ray scattering structure of the A_1 ATPase from *Methanosarcina mazei* Gö1 (GRÜBER *et al.,* 2001) pointed out that the shape and length of the central stalk in the A_1 ATPase is significantly different from F_1 ATPase and it is more elongated in the A_1 ATP synthase. This explains also the distinct class of A-ATP synthase enzyme from F-ATP synthase.

1.2.2.2 Structure and function of individual domains of the A_1A_O ATP synthase

1.2.2.2.1 The A_3B_3 headpiece

The head domain consists of a nucleotide binding subunits A and B (corresponding β and α subunits from F-ATP synthase) arranged in a pseudohexagonal fashion (COSKUN *et al.,* 2002; COSKUN *et al.,* 2004; COSKUN *et al.,* 2004; GRÜBER *et al.,* 2008). Although the high

resolution structure for the head piece of A_1A_O ATP synthase has not been solved yet, the structures of subunit A and subunit B have been solved individually by X-ray crystallography at 2.55 Å and 1.5 Å resolution, respectively (MAEGAWA *et al.,* 2006; SCHÄFER *et al.,* 2006a). In general the structure of subunit A and B consists of three domains: N-terminal β barrel, an α-β domain and a C-terminal α-helical bundle. The important difference between the structures of the two subunits is the presence of a knob-like structure in catalytic subunit A which is made from an insertion of 90 amino acids. This insertion, called non homologous region (NHR) is absent in the homologous β subunit of F_1F_O ATP synthases. It was first seen in the electron micrographs of the intact A_1A_O ATP synthase and believed to connect the peripheral stalk to the A_3B_3 headpiece (COSKUN *et al.,* 2004). Comparison of subunit A and B sequences reveals that subunit A has a phosphate-binding loop (P-loop) sequence GXXGXGKTV, which is conserved at the nucleotide binding sites of the related α and β subunits of F-ATP synthases (WALKER *et al.,* 1982; NELSON *et al.,* 1989; SARASTE *et al.,* 1990; WILMS *et al.,* 1996). Subunit B can also bind to nucleotides, although it does not have an ideal P-loop sequence (SCHÄFER *et al.,* 2006a; KUMAR *et al.,* 2009; MANIMEKALAI *et al.,* 2009). Nucleotide binding capability of both subunits A and B have been tested by photo-affinity ATP labeling and fluorescence correlation spectroscopy (FCS) experiments. The results showed that both subunits can bind to nucleotides and there is a higher affinity for subunit A. It is still unclear about the function of nucleotide binding events to subunit B. Cross-linking study with EDC in presence of nucleotides indicates that coupling of nucleotide binding to subunit B might cause structural alterations in the neighboring subunit B and F (SCHÄFER *et al.,* 2006b; GAYEN *et al.,* 2007). This B-F interaction is proposed to play an important role linking the nucleotide binding events to the ion translocation in the A_O part (GAYEN *et al.,* 2007; RAGHUNATHAN *et al.,* 2009).

1.2.2.2.2 The stalk domain (C:D:F:H_2:E_X)

The central stalk of the A_1A_O ATP synthases descends from the bottom of the headpiece A_3B_3 to the inside of the *c*-ring. The central stalk composed of C, D and F subunits with molecular masses of 36, 25 and 12 kDa, respectively. Subunit C is characteristic only for A_1A_O ATP synthases and was described as a spacer unit that plays important roles in the coupling and rotational events for the synthesis of ATP and ion translocation (MÜLLER *et al.,* 2003; IWATA *et al.,* 2004; NUMOTO *et al.,* 2004; GRÜBER *et al.,* 2008). The overall structure of subunit C contains twenty helices and is separated into three domains (IWATA *et al.,* 2004; NUMOTO *et al.,* 2004). The domains have a similar fold and topology consisting of two long helices and four

shorter helices. All the helices run roughly parallel to form a tapered cylinder. The taper of the subunit C penetrates to the central cavity up to the mid point of the *c*-ring. This is further supported by the electrostatic charge distribution of subunit C. Topside is hydrophobic and the bottom side is charged, which is suitable to bind with the *c*-ring surface (NUMOTO *et al.,* 2004). Docking study showed that some structural changes are necessary for efficient interaction of subunit C with the *c*-ring (MURATA *et al.,* 2005). Subunit D has been described as a structural and functional homolog of subunit γ of F_1F_O ATP synthases (ABRAHAMS *et al.,* 1994). Its central α-helix is surrounded by the three A and B subunits, arranged alternately like the segments of an orange. It passes through a large internal cavity connected to the crevices formed by the A and B subunits (Figure 1.3B). The function of this subunit is supposed to change the conformations of the nucleotide binding pockets in the A_3B_3 headpiece upon rotation, like the role of γ subunit in the F_1F_O ATP synthases (GRÜBER *et al.,* 2001; COSKUN *et al.,* 2002). Subunit F is comprised of a globular domain with a length of 4.5 nm and hook like domain with a length of 3.0 nm as revealed by the small angle X ray scattering structure (SCHÄFER *et al.,* 2006a). Thus, subunit F with a length of 7.5 nm would partially span the width of the central stalk. The shape is almost identical to subunit ε of *Escherichia coli* F_1F_O ATP synthase (UHLIN *et al.,* 1997). Cross linking studies with subunit F of A_1A_O ATP synthase from *M. mazei* revealed that the hook like domain of F subunit is interacting with subunit B in a nucleotide binding dependant manner (SCHÄFER *et al.,* 2006b). Other than the central stalk, the first structure of the complete A_1A_O ATP synthase obtained showed the presence of two peripheral stalks, running from the knobs on A_1 to a level of the A_O domain (COSKUN *et al.,* 2004). The potential candidates of these two parts are subunits H and *a*, respectively. Subunit H (~12 KDa) is the smallest subunit in enzyme complex. Sequence based predictions suggest that the subunit H has an α-helical structure and can form a homo dimeric coiled-coil structure (BIUKOVIĆ *et al.,* 2007). The experimental data also revealed the presence of a collar-like structure that interacts with both peripheral stalks and is in direct neighborhood to the central stalk via the rotary subunit D. The remaining hydrophilic subunit E is the most likely candidate for constituting this collar domain (BERNAL *et al.,* 2004; COSKUN *et al.,* 2004). The size of the E subunit is around 23 kDa having two domains, a highly helical N-terminal domain and a C-terminal domain composed of short helices and β sheets. The C-terminal domain of subunit E ($E_{81\text{-}198}$) from *Pyrococcus horikoshii* OT3 has been solved by crystallography at 1.85 Å resolution (LOKANATH *et al.,* 2007). The structure consists of four antiparallel β strands and

six helices, revealing an elongated shape with a distance of 75 Å. The stoichiometry of subunits H and E in A-ATP synthase is still not established. Some studies have revealed that subunits H and E form the hetero dimer and constitute the peripheral stalks of the enzyme (KISH-TRIER *et al.,* 2008; KISH-TRIER *et al.,* 2009; VONCK *et al.,* 2009). Both the peripheral stalks and the collar-like structure are believed to be involved in constituting the stator region of the A_1A_O ATP synthase, which is believed to have important roles in enzyme assembly as well as energy coupling between A_1 and A_O (SCHÄFER *et al.,* 2006b). They may even have potential regulatory roles in interaction with other subunits within the enzyme complex (LOKANATH *et al.,* 2007).

1.2.2.2.3 The ion pumping domain A_O

The A_O domain represents the membrane embedded rotor, containing only two subunits subunit *a* (stator) and subunit *c* (rotor). It has a molecular mass of 75 kDa with a hydrophilic N terminal and a hydrophobic C-terminal domain. The hydrophilic sector is believed to be similar in function to subunit *b* of F_1F_O ATP synthases and are made of helices; the hydrophobic C-terminus has seven transmembrane helices similar to subunit *a* from F_1F_O ATP synthase (MÜLLER *et al.,* 2003). The most important feature is a conserved arginine residue which is believed to be an essential positive charge for ion translocation (GRAHAM *et al.,* 2003; QI *et al.,* 2007). In most of the archaea, subunit *c* has two transmembrane helices (IHARA *et al.,* 1997; STEINERT *et al.,* 1997). Four and six transmembrane helices of subunit *c* were also seen for *M. thermoautotrophicus*, *M. jannaschii* respectively, because of the gene duplication and triplication events, followed by the fusion of the gene copies (RUPPERT *et al.,* 1999). It is still not completely understood how the *c* subunit plays a role in ion translocation and its coupling to ATP synthesis. In 2005, Murata *et al.* has solved the crystallographic structure of subunit *c* from *Enterococcus hirae* A-ATP synthase (MURATA *et al.,* 2005). The subunit *c* consists of five helices (H_0-H_4), out of which H_1-H_4 are transmembrane segments. Ten *c* subunits are arranged in two concentric rings where H_1 and H_3 are alternate in the inner ring and H_2 and H_4 are alternate in the outer ring (Figure 1.6). One of the interesting features in the structure of the *c*-ring, is that the lower and upper part of the helices H_2 and H_4 bend outward from the axis from the ring, whereas bends in the H_1 and H_3 helices are negligible. These structural features broaden the radius of the cytoplasmic and the periplasmic side with respect to the central region. H_0 helices are pointing inward on the opposite side of the membrane, narrowing the internal diameter of the ring. The ring diameter differs greatly from the *c* ring of bacterial F ATP synthases (STOCK *et*

al., 1999; VONCK *et al.,* 2002). This fact points out the differences of central stalks between these ATP synthases and its association with the upper part of the ring. In ATP synthases subunit C is supposed to be interposed between the central stalk (IWATA *et al.,* 2004) and the ring, whereas in F ATP synthases there is a direct contact between the central stalk (γ, δ, ε subunits) and the ring (STOCK *et al.,* 1999). The structure of subunit *c* also gives important aspects towards the ion specificity of A_1A_O ATP synthases. Residues from H_2, H_3, H_4 (T64, Q65, Q110 and E139) are important for the sodium binding pocket. The change of torsional angle of E139 might allow the exchange of sodium ion with the bulk phase (MURATA *et al.,* 2005). The clockwise rotation of the *c*-ring brings an occupied sodium binding site to the *c*-ring – subunit *a* interface. R573 of subunit *a*, which is very close to the sodium binding pocket, disrupts the hydrogen bonding network and releases the ion into the periplasm. Further rotation will allow the binding of another sodium ion by disrupting the arginine – glutamic acid interaction (MURATA *et al.,* 2005).

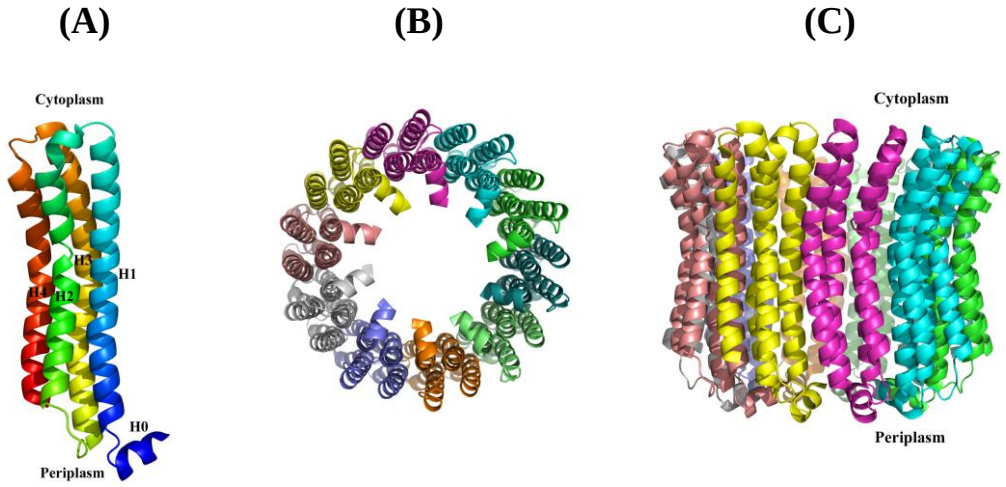


Figure 1.6: Structure of subunit *c* from *Enterococcus hirae*. (A) Ribbon representation of the structure of the *c* subunit. 10 protomers of the *c* subunit form the ring structure. (B) shows a view from the cytoplasmic side of the membrane and (C) a side view of the *c*-ring (MURATA *et al.,* 2005).

1.2.2.3 Conformational changes due to nucleotide binding in the A_1A_O ATP synthase

Small angle X-ray scattering studies showed that significant structural alterations occur to subunits A and B due to nucleotide binding in the A_1A_O ATP synthase from *M. mazei* Gö1 (COSKUN *et al.,* 2002). The value of the radius of gyration decreases slightly after binding to Mg-AMP-PNP and increases after binding Mg-ATP, Mg-ADP (COSKUN *et al.,* 2002). Such conformational changes of the quaternary structure are consistent with alterations in the closely related F_1-ATPase, as demonstrated in a recent crystallographic structure of the bovine F_1-

ATPase (MENZ *et al.,* 2001). In the F_1-ATPase significant changes in the lower part of the nucleotide binding domain and the C-terminal domain of the catalytic β subunit have been observed due to the nucleotide binding. Cross linking of central stalk subunit D to catalytic subunit A depends also on type of nucleotide present in the solution. Mg-ADP and Mg-ATP favors the cross-link formation, whereas Mg-ADP + Pi, Mg-AMP-PNP and the absence of nucleotide disrupts the cross link (COSKUN *et al.,* 2002). These results point out the close proximity of subunit A and D and will allow the coupling between the catalytic site events in A via D to the other central stalk subunit C and F. Subunit F can be cross-linked to subunit B depending on the type of nucleotide. The similar nucleotide dependant movement is also seen in case of subunit H and subunit A (SCHÄFER *et al.*, 2006b). Therefore, the conformational changes in the A_1 headpiece depending on the type of the nucleotide might correlate with structural alterations in the stalk region. This can be expected when the stalk couples ATP hydrolysis/synthesis to ion conduction (MÜLLER *et al.*, 2003).

Thus, more structural and biochemical investigations are needed in order to understand this fascinating enzyme especially the coupling stalk regions. In the present study isolated stalk subunits were cloned, expressed and purified. Structural characterizations were done with nuclear magnetic resonance (NMR) spectroscopy. NMR spectroscopy has been used as a tool to determine the three dimensional structure, dynamics as well as protein-protein interactions.

1.3 Outline of the thesis

The A_1A_O adenosine triphosphate (ATP) synthase from archaea uses the ion gradients generated across the membrane sector (A_O) to synthesize ATP in the A_3B_3 domain of the A_1 sector. The crystal structure of a large portion of A_1 has been solved. However, the structure of the central portion of the enzyme, though which conformational changes in A_O are communicated to A_1, has until now remained elusive. To get a detailed picture in this region the solution structure of the subunit F, N-and C-terminal domain of subunit H and N-terminal domain of subunit E ($E_{1\text{-}52}$) of the A_1A_O ATP synthase were determined. Subunit F from *M. mazei* Gö1 exhibits a distinct two-domain structure, with the N-terminal having 78 residues and residues 79-101 forming the flexible C-terminal part. The two domains are loosely associated with more flexibility relative to each other. The flexibility of the C-terminal domain is confirmed by NMR dynamics studies. In addition, NMR structure of the isolated C-terminus of subunit F points out that the flexibility of the C-terminus enables subunit F to switch between an elongated and closed state and thereby provides new insights into energy coupling between A_O and A_1.

Subunit H of the A_1A_O ATP synthase is a component of one of the two peripheral stalks connecting the A_1 and A_O domain. The structural and functional role of this subunit is not known properly since its precise arrangement inside the A_1A_O complex is still an unresolved issue. The goal of this project is to over express, produce the recombinant protein, which could be used to derive the structural and biochemical information. Circular dichroism (CD) spectroscopy revealed that subunit H is comprised of 78% α-helix and a coiled-coil arrangement. To understand the importance of C-terminus for the stabilization of the quaternary structure of the protein, structure of a short length peptide (H_{85-104}) in the C-terminus of subunit H has been solved and three mutants of subunit H (H_{I93C}, H_{L96C} and H_{L98C}) were expressed, purified, and analyzed by cross-links and CD experiments. To understand how the N-terminal domain of H is formed the solution NMR structure of H_{1-47} was solved, revealing an α-helix between residues 15–42 and a flexible N-terminal stretch. The α-helix includes a kink that would bring the two helices of the C-terminus into the coiled-coil arrangement.

The structural and functional role of another stalk subunit, subunit E is also essential in understanding the function of A_1A_O enzyme. To get into details, NMR titration experiments were performed revealing that the N-terminal residues M1–6, K10, E11, A15, Vl20 and E24 of $H_{1–47}$ interact specifically with the N-terminal domain $E_{1–100}$ of subunit E. A more detailed picture regarding the residues of $E_{1–100}$ involved in this association was obtained by titration studies using the N-terminal peptides $E_{1–20}$, $E_{21–40}$ and $E_{41–60}$. These data indicated that the N-terminal tail $E_{41–60}$ interacts with the N-terminal amino acids of $H_{1–47}$. In order to understand the E-H assembly and the N-terminal structure of E, the truncated construct $E_{1–52}$ of *M. jannaschii* A_1A_O ATP synthase was determined by NMR spectroscopy. The protein is 60.5 Å in length and forms an α helix between the residues 8–48. The molecule is amphipathic with a strip of hydrophobic residues, which could lead to a possible helix-helix interaction with neighboring subunit H. Analysis of $^{1}H–^{15}N$ heteronuclear single quantum coherence (HSQC) spectra of the central stalk subunit F in the presence and absence of $E_{101–206}$ show no obvious interaction between the C-terminal domain of E and subunit F. The data presented provided, for the first time, structural insights into the interaction of subunits E and H, and their arrangement within A_1A_O ATP synthase.

2.1 Materials

2.1.1 Chemicals

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

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)
$^{15}NH^{4}Cl$	Cambridge Isotopes Lab (USA)
^{13}C Glucose	Cambridge Isotopes Lab (USA)
LB Media	BD (Sparks, MD, USA)
Electrophoresis Chemicals	
(Agarose, SDS, Glycine, APS etc.)	Bio-Rad (Hercules, CA, USA)
Antibiotics	Calbiochem, Sigma and Gibco (Invitrogen)
IPTG	Fermentas
BSA	GERBU (Heidelberg, Germany)

2.1.2 Molecular biology materials

Genomic DNA	
M. jannaschii (43067D)	ATCC, Manassas, VA 20108, USA.
Primers	1st Base and Research Biolabs (Singapore)
Peptides	NTU proteomics core facility, School of Biological Sciences, Nanyang Technological University, Singapore.
Pfu and T4 DNA Polymerase	Fermentas (Glen Burnie, MD, USA)
NcoI, SacI	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)
pET-9d1	GRÜBER *et al.*, 2002.
pET-24 (+)	Novagene (Darmstadt, Germany)

2.1.3 Chromatography

2.1.3.1 Ion exchange

RESOURCE™ Q (6 ml)	Amersham Biosciences (New York city, USA)

2.1.3.2 Gel filtration

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

2.1.3.3 Instruments and accessories

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

2.1.3.4 Protein concentration, estimation

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

2.1.4 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

2.1.5 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)
NMRpipe	Delaglio *et al* (NIH, USA)
nmrDraw	Howard Hughes Medical Institute, USA
PyMOL v0.99	DeLano Scientific LLC, USA

2.2 Methods

2.2.1 Cloning of subunit H of A_1A_O ATP synthase from *M. jannaschii*

To amplify the Atp H coding region, oligonucleotide primers 5'-GGA ACC ATG GGC GTT AGT GTT ATG-3' (forward primer) and 5'-AAT CAC CTG AGC TCT TAA ATC TCA AGA ATC-3' (reverse primer), incorporating *Nco*I and *Sac*I restriction sites (underlined) were designed. Isolated pure genomic DNA of *Methanocaldococcus jannaschii* ATCC® 43067D was used as a template for the amplification of the subunit H gene. Polymerase chain reaction (PCR) was set up in a total volume of 50 µl on ice as mentioned below with the appropriate concentration of constituents as:

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 pipetted into a PCR tube kept on ice all the time, mixing of the reagents was properly done by short centrifugation. PCR thermocycler (Biometra T personal) was preheated to 95 ºC before reaction tubes were placed inside the machine. The following PCR thermocycler program was used for the amplification. The cycle denaturation, annealing and extension steps were repeated for another 29 cycles. Total time taken for the PCR program was 1h 43 min 49 sec.

Cycle steps	Temperature	Time	
Lid	99 ºC		
Initial denaturation	96 ºC	3 min	
Cycle denaturation	96 ºC	30 sec	30 cycles
Annealing	58 ºC	45 sec	
Extension	72 ºC	60 sec	
Final extension	72 ºC	10 min	
End	4 ºC		

The amplified PCR product was stored at -20 °C. In order to check the quality of the PCR amplified product, 5 µl of the samples were applied onto an analytical agarose gel (1%) and after that, the gel was kept in ethylene dibromide solution for 20 min and observed under UV light. The size of the PCR product was confirmed with the appropriate DNA marker. Remaining reaction mixture was applied onto a preparative agarose gel. 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 Tris-HCl buffer of pH 8.0. 1 µl of gel extracted PCR product was applied onto 1% analytical agarose gel for estimating the purity of the sample. Purified DNA was double digested overnight at 37 ºC by taking appropriate amount 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 2.2.1) and amplified PCR product was setup at room temperature. The reagents used were as follows:

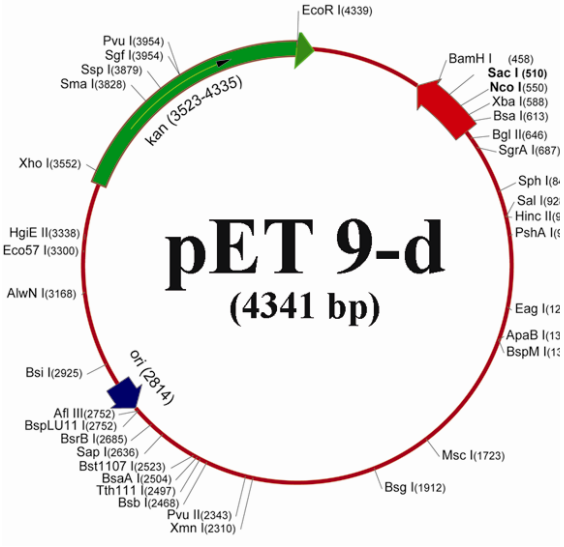


Figure 2.2.1: Vector Map modified pET-9d1 (+) (GRÜBER *et al.*, 2002) vector showing full map and unique multiple cloning sites (MCS) that has been used to clone various genes.

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

The ligation mixture was kept at room temperature for 60 min. The reaction was terminated by precipitation of DNA from rest of the reaction misture by adding 85 µl of Milli Q water to make a final volume of 100 µl. 1 ml of butanol was added into the mixture and mixed thoroughly. Then the mixture was centrifuged at 13,000 x *g* (using Eppendorf mini-centrifuge) for 10 min. Supernatant was discarded gently and the pellet was resuspended in 150 µl of 70%

ethanol and centrifuged again for 5 min. Supernatant was removed and the pellet was re-dissolved in 10 µl of elution buffer (10 mM Tris-HCl, pH 8.5) (EB buffer, QIAGEN) or Milli Q water. 5 – 10 µl of ligation mixture was used for the transformation into *Escherichia coli* DH5α cells. On the next day colonies were picked from the transformation plate and grown at 37 °C. The plasmid DNA was isolated using standard protocol (QIAGEN mini-prep kit) and subsequently double digested and applied onto 1% agarose gel, in order to confirm the ligated insert. Size of insert and vector were compared with appropriate controls and markers. 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).

2.2.2 Cloning of other constructs

Similar strategy was used for the cloning of three mutants (H_{I93C}, H_{L96C} and H_{L98C}) of subunit H of A_1A_O ATP synthase from *M. jannaschii*. The primers used for current study are described below. Since the mutations sites are in the C-terminal of subunit H, I have introduced the mutations in the reverse primers (marked underline) in each case. PCR reactions and PCR program were set up according to the methods described in section 2.2.1, except for the annealing temperature of the PCR reaction, which in case of H_{L98C} mutant was set at 55 °C.

Construct	Primers
I93C	Forward 5'-CGC GAC ACC ATG GGC GTT AGT GTT ATG GAA GCA ATA-3' Reverse 5'-CCG AGC TCT TAA ATC TCA AGA ATC TCA GAC AAT TTC AAC GAA AGG CAC TTA ACC T-3'
L96C	Forward 5'-CGC GAC ACC ATG GGC GTT AGT GTT ATG GAA GCA ATA-3' Reverse 5'-CCG AGC TCT TAA ATC TCA AGA ATC TCA GAC AAT TTG CAC GAA AGT ATC-3'
L98C	Forward 5'-CGC GAC ACC ATG GGC GTT AGT GTT ATG GAA GCA ATA-3' Reverse 5'-CCG AGC TCT TAA ATC TCA AGA ATC TCA GAA CAT TTC AAC GA-3'

2.2.3 Protein production and purification of subunit H and its mutants of A_1A_O ATP synthase from *M. jannaschii*

2.2.3.1 Induction test

From the selection plate, five colonies were taken and grown till the OD_{600} reaches to 0.6 at 37 °C by constant shaking at 180 - 200 r.p.m. (Infors HT Minitron shaker). Induction test was carried out by using 1 mM IPTG for a period of 3 h at 37 °C. Respective controls under identical conditions were also taken, where IPTG was not added. Then the cells were centrifuged at 11,000 x g for 5 min at 4 °C and resuspended properly in 40 µl lysis buffer (98 µl 4 x Laemmli buffer, 4.1 µl DTT and 423.5 µl H_2O). Resuspended samples were heated at 95 °C for 5 min and were loaded in 17 % SDS polyacrylamide gel (Figure 2.2.2).

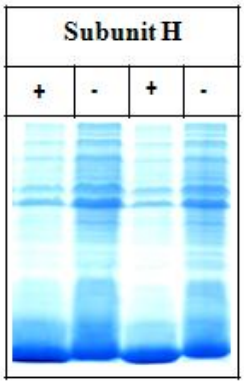


Figure 2.2.2: Induction test of subunit H: Uninduced (-) and induced samples for two colonies of subunit H from *M. jannaschii*. Induction test was done with 1 mM IPTG at 37 °C. Whole cell lysate was heated in 1 x lysis buffer at 95 °C for 5 mins and then applied onto 17 % SDS gel.

2.2.3.2 Solubility test of recombinant proteins

The solubility of the protein was tested in different buffers. 50 ml cultures were harvested and appropriate concentration of IPTG was added when the OD_{600} of the culture reached to 0.6 to 0.7. Cells were separated into five equal fractions and pelleted down by centrifugation at 10,000 x g for 10 min. In order to prevent degradation, appropriate concentrations of Pefabloc^{SC} (2-8 mM) and/or PMSF (1-2 mM) were added. In the case of H mutant proteins, which have cysteine residues, reducing agent DTT (0.5 mM in the case of Ni^{2+}-NTA resin column and 1 mM in the case of Resource Q and Superdex 75 columns) was used. The cell pellet was completely resuspended and sonicated at 20% power with KE76 tip of sonicator (Bandelin Sonoplus) three times for one minute each. A cooling interval of two minutes was kept between each cycle. After sonication, the pellet and the supernatant were separated by centrifugation and the pellet was

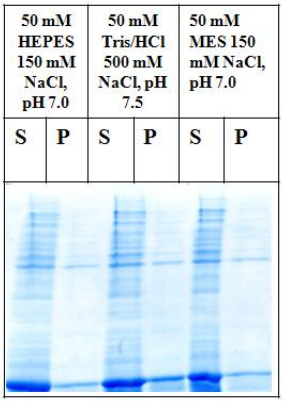


Figure 2.2.3: Solubility test results from H subunit of A_1A_O ATP synthase from *M. jannaschii*. Pellet (p) and supernatant (s) in various buffers is shown in the gel.

again suspended in the same buffer. Samples from both the pellet and supernatant were loaded onto a 17 % SDS gel (Figure 2.2.3).

The *E. coli* cells having the gene atp H was grown on 30 μg/ml kanamycin-containing Luria-Bertoni (LB) agar plates. To express subunit H, liquid cultures were shaken in LB medium containing kanamycin (30 μg/ml) for about 20 h at 30 °C until an optical density OD_{600} of 0.6-0.7 was reached. To induce expression, the cultures were supplemented with IPTG to a final concentration of 1 mM. Following incubation for another 4 h at 30 °C, the cells were harvested by centrifugation at 6,000 x *g* for 20 min at 4 °C. Afterwards the protein was purified by following the scheme described below.

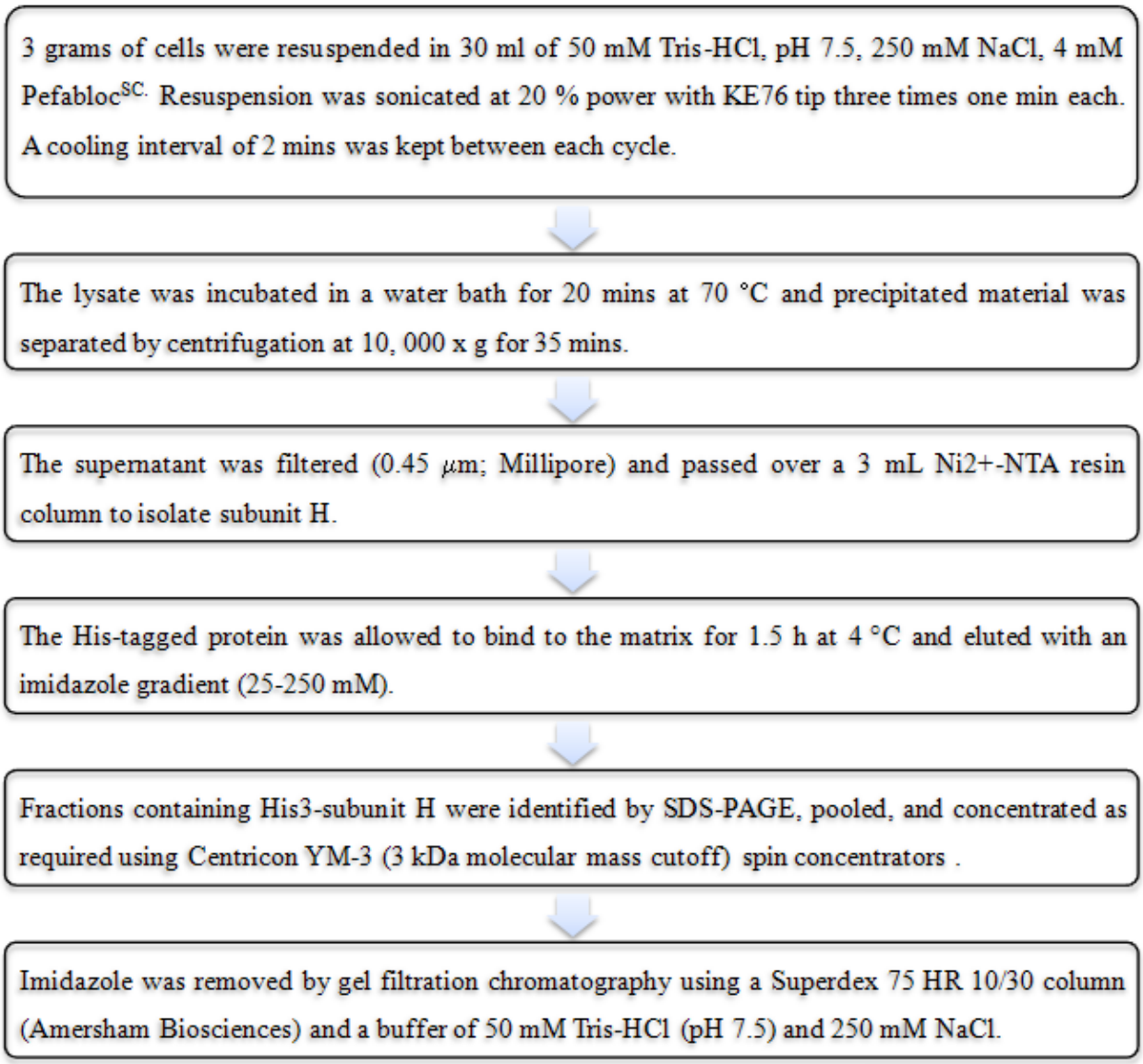


Figure 2.2.4: Scheme for the purification of subunit H of A_1A_O ATP synthase from *M. jannaschii*.

The mutants of subunit H (H_{I93C}, H_{L96C} and H_{L98C}), were isolated using a Ni^{2+}-NTA resin column as described above using the buffer of 50 mM Tris-HCl (pH 7.5), 250 mM NaCl and 0.5 mM DTT. After that the sample was applied on an ion-exchange column (Resource Q, 6 ml; Amersham Biosciences). The protein was eluted by a linear NaCl-gradient (0 M - 1 M) with a flow rate of 1 ml/min. Finally the sample was applied to a size exclusion column (Superdex 75 HR 10/30; GE Healthcare).

2.2.4 Protein quantification by bicinchoninic acid method

The bicinchoninic acid (BCA) assay is available in kit form from Pierce (Rockford, IL, USA.). The protein concentration was estimated as per the manufacturer's instructions. BSA (0 – 250 µg/ml) was used as standard. In general, two dilutions of protein samples were taken and each of them was measured in triplicate to avoid any error in the measurement. Optical density was measured at 562 nm against blank.

2.2.5 Circular Dichroism spectroscopy

Circular dichroism (CD) is used to determine the secondary structure content of proteins and peptides. Steady-state CD spectra were measured in the far-UV light (185-260 nm) using a Chirascan spectropolarimeter (Applied Photophysics) under continuous purging of N_2 gas. 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 taken with the average of 1 sec at each wavelength, and the recorded ellipticity values were the average of three determinations for each sample at a concentration of 2 mg/ml. The spectrum for the buffer was subtracted from the spectrum of the protein. CD values were converted to mean residue ellipticity in units of deg x $dmol^{-1}$ x aa^{-1} using the software Chirascan version 1.2 (Applied Photophysics) by giving appropriate cell path length, molecular weight, number of amino acid residues and protein concentration values. This baseline corrected spectrum was used as input for computer methods to obtain predictions of secondary structure.

2.2.6 Nuclear magnetic resonance (NMR) spectroscopy

2.2.6.1 Isotopic labeling (^{15}N, ^{13}C) of proteins

Reagents:

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 % D-Glucose
30 µM $FeCl_3$
5 ng/l Thiamin
Antibiotic
LB Media

For the production of uniformly labelled ^{15}N and ^{15}N and ^{13}C recombinant proteins in *E. coli* BL21 (DE3), M9 minimal media was used. All unlabeled reagents were standard laboratory reagents grade. All solutions were prepared using Milli-Q water (Millipore, Billerica, MA). 50 ml of overnight culture was grown. The cells were pelleted down by centrifugation at 5,000 x g at room temperature. Then the pellet was washed two times with LB media following a final wash in 1 x M9 media and finally resuspended in 10 ml M9 media. 1 ml of final resuspension was inoculated per 500 ml. Cells were shaken at 37 ºC till an OD_{600} of 0.6 - 0.7 was achieved. After that the culture was induced with 1 mM IPTG for 4 h or at 37 ºC. At the end, cells were harvested and frozen in liquid nitrogen and stored at – 80 ºC until purification.

2.2.6.2 NMR data collection and processing

The NMR samples were prepared in 25 mM NaH_2PO_4/Na_2HPO_4 (pH 6.5) and 0.1% NaN_3. 10% D_2O was added for the NMR-lock signal. All NMR experiments were performed at temperatures ranging from 10 °C to 35 °C on a Bruker Avance 600 MHz and 700 MHz spectrometers equipped with 5 mm triple-resonance ($^1H/^{15}N/^{13}C$) single-axis gradient probes, and a cryo-probe for both spectrometers. Different protein concentrations were used for the measurements. For HSQC experiment, 0.01 mM to 3 mM concentration of the protein was used. For the triple resonance backbone experiments, generally 0.5-1 mM protein sample was used. Appropriate pulse calibrations and other parameters including temperature, buffer and protein concentration were optimized before making final measurements. Homo-nuclear experiments included 2D ^1H, ^1H-NOESY, and TOCSY experiments for samples in H_2O and ^2H_2O (spectral widths of about 9615.85 Hz in both dimensions, using a WATERGATE scheme to suppress the water signal for the H_2O samples and pre-saturation of the residual water signal during recovery time for the ^2H_2O samples). For 2D TOCSY, a 80 ms mixing time and for 2D NOESY, a 300 ms mixing time were used for the data collection. Hetero nuclear experiments included 2D ^{15}N-HSQC, 2D ^{13}C-HSQC, 3D HNCA, HNCO, HN(CO)CA, HNCACB, CBCA(CO)NH, HNCACO, HCCH-TOCSY HCCCONH, 3D ^{15}N-NOESY-HSQC, 3D ^{13}C NOESY-HSQC. In all

of the above experiments, spectral widths were set to 9.6 kHz for protons, 2.8 kHz for ^{15}N and 6-12 kHz for ^{13}C. The 3D ^{15}N-^{1}H NOESY-HSQC and 3D ^{13}C-^{1}H NOESY-HSQC mixing time of 100 ms were used for NOESY assignment of the subunit F protein from *M. mazei* Gö1. 3D ^{15}N-^{1}H NOESY-HSQC of 200 ms mixing time were collected for the proteins H_{1-47} and E_{1-52}. 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 (DELAGLIO *et al.*, 1995) or Bruker Avance spectrometer in-built software Topspin. Peak-picking and data analysis of the Fourier transformed spectra were performed with the SPARKY program (GODDARD *et al.*, 1997). The translational diffusion rates were measured by monitoring 1D ^{1}H signal decay due to molecular diffusion in the z-direction of the sample using PFGs at variable concentrations of H_{1-47}, H_{1-61}, H_{1-69} and H_{1-80}, respectively. In each experiment, the PFG strength was linearly incremented from 2 to 95 G/cm, with a translational diffusion delay of 200 ms and total encoding and decoding gradient durations of 5 ms. The diffusion rates were estimated using TopSpin (Bruker BioSpin).

2.2.6.3 Resonance assignment

In the initial stage of any investigation of nuclear magnetic resonance data of a protein, each resonance associated with a particular nucleus in the protein molecule must be assigned. The information useful for the sequence specific assignment for a protein is of three types based on scalar couplings, dipolar coupling and the chemical shift. Different strategies are employed for the resonance assignments depending on the type of the sample: unlabeled proteins or isotopically labeled proteins (CAVANAGH *et al.*, 1995).

Even for a small protein it is impossible to resolve all resonance lines in a one-dimensional (1D) spectrum. Assignments can be obtained using the sequential assignment approach based on homonuclear two-dimensional (2D) correlation spectroscopy (COSY) or total correlation spectroscopy (TOCSY) (Braunschweiler *et al.*, 1983) and 2D nuclear Overhauser effect spectroscopy (NOESY). TOCSY experiment is 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 are generally the best resolved set of resonances and resonate between 6-12 ppm. The side chain resonance assignment are generally done by examining the peaks in the backbone amide regions. This process is known as the sequential assignment strategy developed by Wüthrich and his co workers (WÜTHRICH, 1990).

For the sequential assignment strategy of larger molecules (>10 kDa.) this above assignment approach is not useful because ovelaps of the NMR signals complicates the assignment. Such systems require the addition of a third dimension by the replacement of the naturally abundant isotope ^{14}N with ^{15}N and ^{12}C by ^{13}C. Here, 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. The HNCA experiment gives the chemical shifts of the $^{13}C^{\alpha}$ and the previous amino acid's $^{13}C^{\alpha}$,

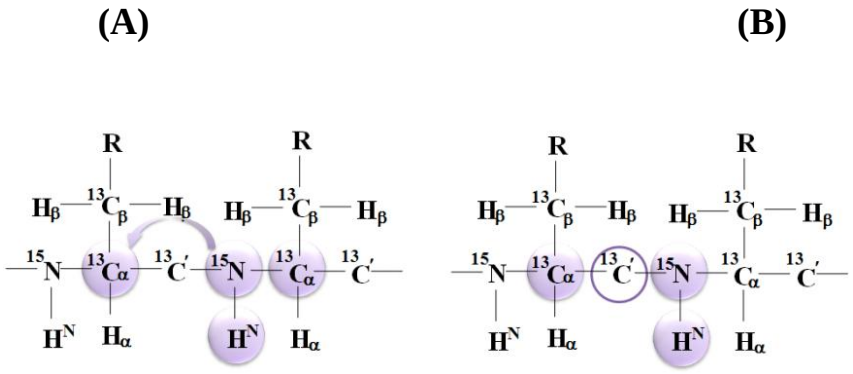


Figure 2.2.5: Transfer of magnetization in the HNCA (A) and HN(CO)CA (B) experiments. HNCA experiment provides sequential connectivities by making use of ^{15}N-$^{13}C\alpha$ J-coupling (7 - 11 Hz). HN(CO)CA experiment provides sequence correlations by transferring the coherence via the intervening ^{13}CO spin (CAVANAGH *et al.*, 1995).

while the HN(CO)CA gives only the previous amino acid's $^{13}C^{\alpha}$. Sequential assignment can then be undertaken by matching the shifts of each spin system's own and previous carbons. Figure 2.2.5 shows the transfer of magnetization in the HNCA and HNCOCA. Figure 2.2.6 shows N = 128.682 ppm plane of HNCA and HN(CO)CA experiments. In HNCA plane each amide ^{1}H chemical shift corresponds to two $^{13}C^{\alpha}$ chemical shifts. The peak with the stronger intensity is the amino acid's own $^{13}C^{\alpha}$, while the weaker intensity peak belongs to the previous amino acid's $^{13}C^{\alpha}$. In case of HN(CO)CA experiment only one peak is appearing for each amide ^{1}H shifts. An example of the assignment by these two experiments is also shown in Figure 2.2.6.

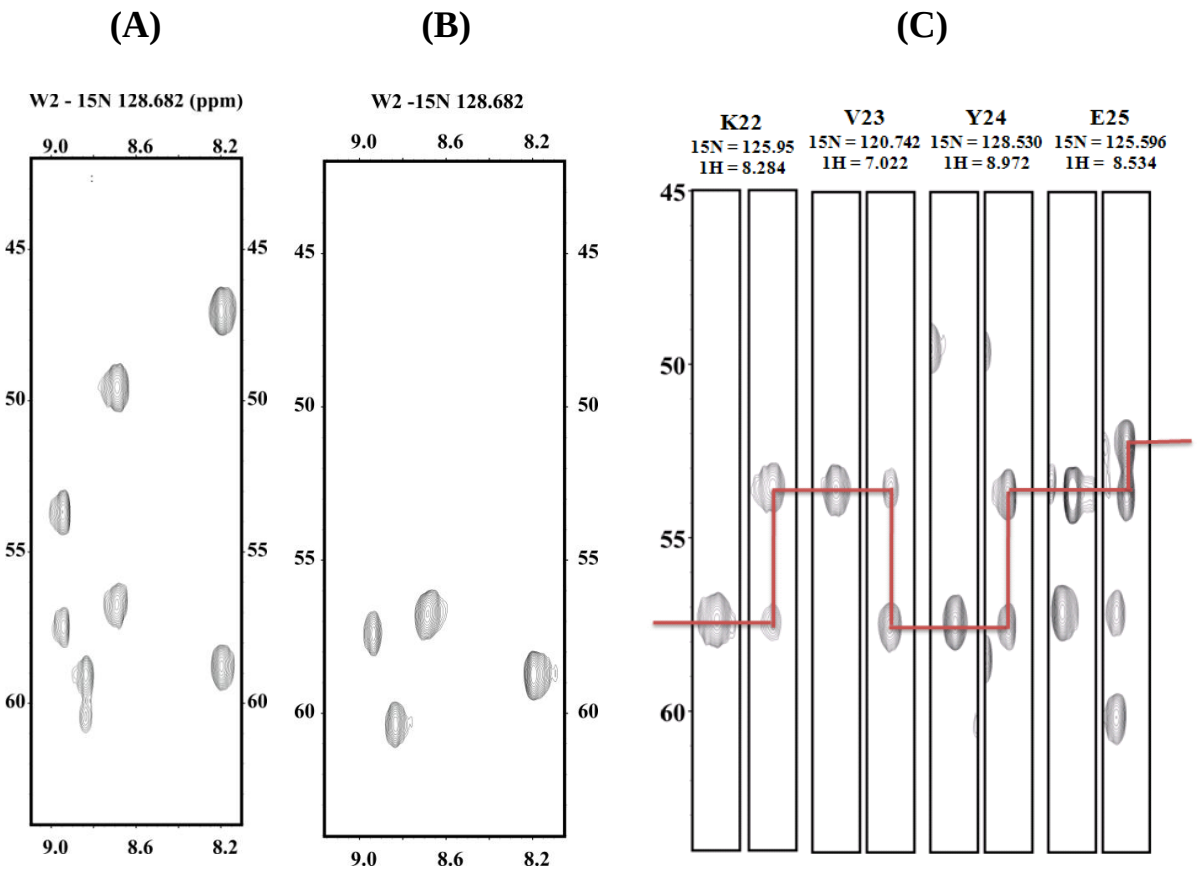


Figure 2.2.6: ^{15}N = 128.682 ppm plane of the HNCA experiment (A); ^{15}N = 128.682 ppm plane of the HN(CO)CA experiment(B); An example of assignment of 4 residues (K22,V23, Y24 and E25) by HNCA and HN(CO)CA experiments in case of subunit F of A_1A_O ATP synthase from *M. mazei* Gö1(C).

The next four experiments are grouped in pairs in a similar way, providing a mean of double-checking the chemical shift data. 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 2.2.7 in case of HNCACB and CBCA(CO)NH experiments. These two experiments take into account both the $^{13}C^{\alpha}$ 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. In the same way, the experiments HNCO, HN(CA)CO are also used in the backbone assignment of the proteins. HNCO experiments correlates the amide ^{1}H and ^{15}N chemical shifts of one amino acid to the ^{13}CO chemical shifts of the previous residue by using ^{15}N-^{13}CO J-coupling (~15Hz). HN(CA)CO experiment can give correlations of ^{1}H and ^{15}N spins to the ^{13}CO chemical shifts of the previous and same residue. Thus, HNCO experiment when used in conjunction with HN(CA)CO experiment, it provides a method to sequentially assign the amide ^{1}H, ^{15}N and ^{13}CO resonances.

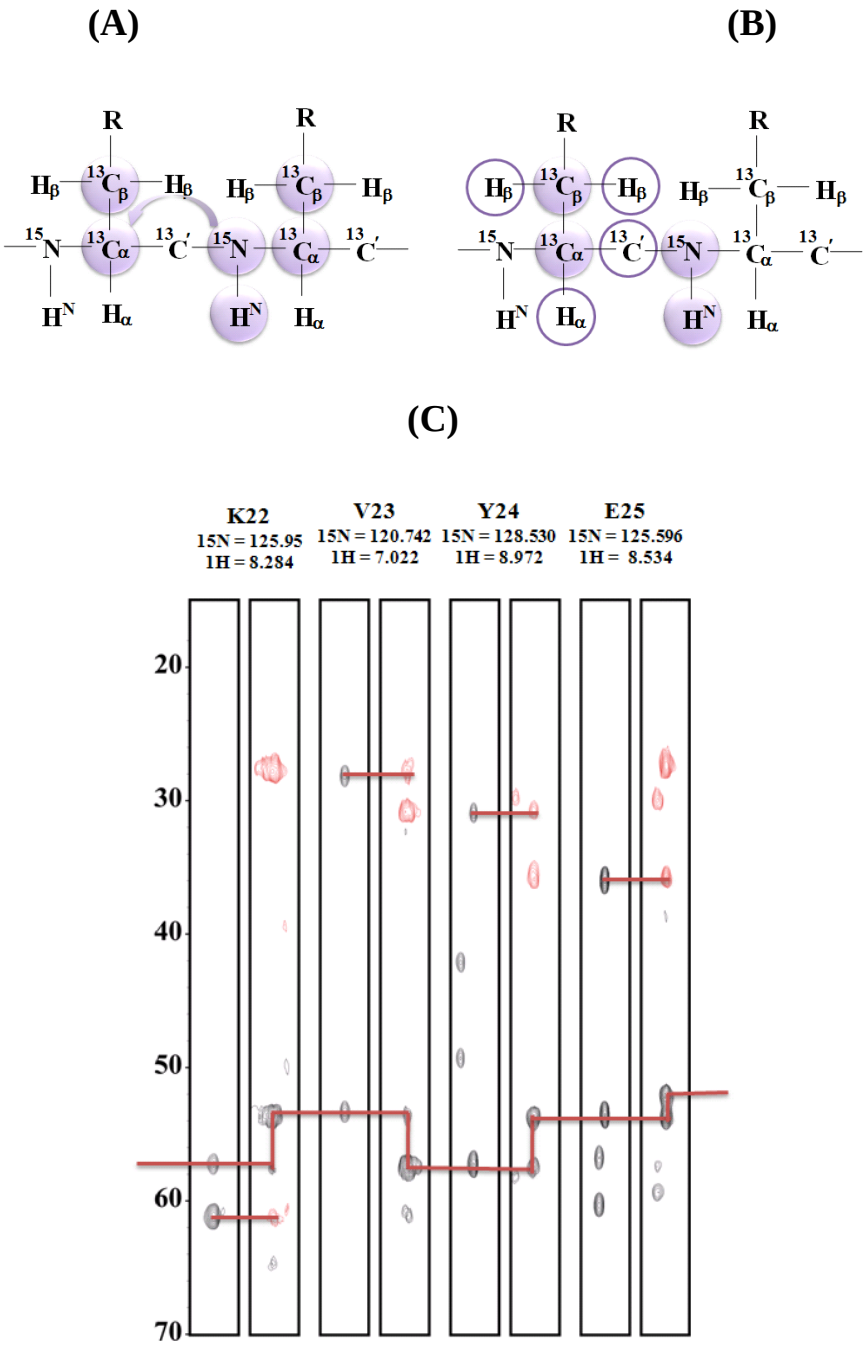


Figure 2.2.7: Transfer of magnetization in the HNCACB (A) and CBCA(CO)NH (B). HNCACB experiment correlates the $^{13}C\alpha$ and $^{13}C\beta$ resonances with the amide ^{1}H and ^{15}N resonances of the same and the previous residue whereas the CBCA(CO)NH experiment correlate the $^{13}C\alpha$ and $^{13}C\beta$ resonances with the amide 1H and 15N resonances of the previous residue (CAVANAGH *et al.*, 1995). An example of assignment of 4 residues (K22,V23, Y24 and E25) by HNCACB and CBCA(CO)NH experiments in case of subunit F of A_1A_O ATP synthase from *M. mazei* Gö1(C).

After successfully assigning the backbone of residues in ^{1}H-^{15}N HSQC spectrum by using three sets of experiments (HNCA and HN(CO)CA; HNCACB and CBCA(CO)HN; HNCO and HN(CA)CO) side chain (^{13}C and ^{1}H) chemical shifts were assigned by using mainly three experiments: (H)CC(CO)NH-TOCSY, H(CC)(CO)NH-TOCSY and HCCH TOCSY (Figure

2.2.8). The first two experiments (H)CC(CO)NH-TOCSY, H(CC)(CO)NH-TOCSY provide the correlation of ^{1}H and ^{15}N amide resonances of one residue with the ^{13}C side chain and ^{1}H side chain resonances, respectively, of the previous residue. The 3D HCCH-TOCSY experiment is specifically designed to correlate side-chain aliphatic proton and ^{13}C resonances via ^{13}C-^{1}H and ^{13}C-^{13}C J-coupling constants (CAVANAGH *et al.*, 1995).

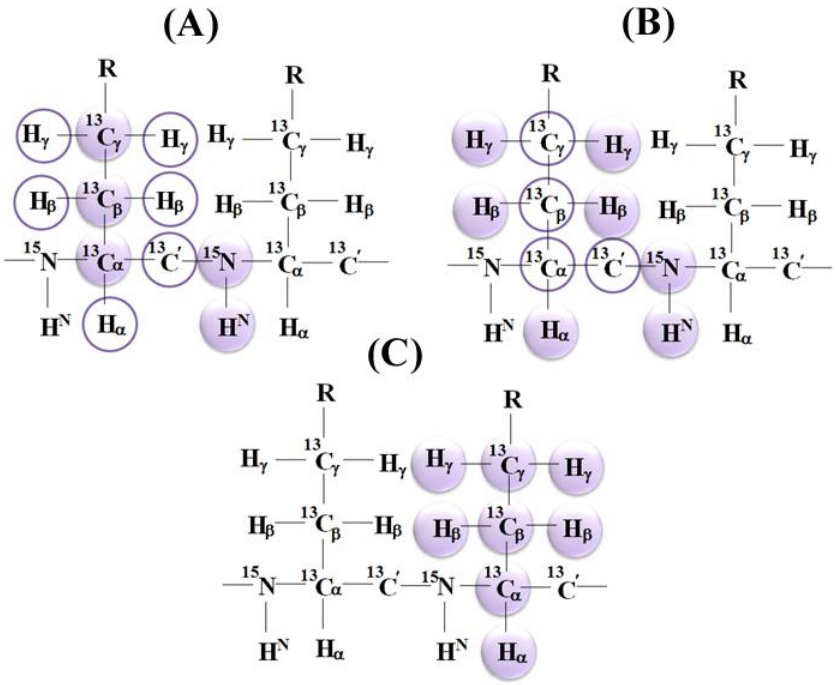


Figure 2.2.8: Transfer of magnetization in the (H)CC(CO)NH-TOCSY (A), H(CC)(CO)NH-TOCSY (B) and HCCH TOCSY (C) (CAVANAGH *et al.*, 1995).

2.2.6.4 NMR spectroscopy and structure calculation

The process of protein structure determination by NMR spectroscopy (Figure 2.2.9) involves three basic steps: first, the assignment of every resonance to a nuclear spin, second, the measurement of structural restraints and third, the calculation and subsequent refinement of structures that satisfy the experimental restraints. The most important restraint is NOE. NOE cross peak intensities depends on the cross relaxation rate, proportional to the inverse sixth power of the distance between the two interacting protons (CAVANAGH *et al.*, 1995). Thus, NOE intensities are used as distance restraints in the structure calculation. In NOESY experiment mixing time is very important. Intra residue and sequential NOE intensities were obtained from 100 ms NOESY spectra for subunit F and 200 ms NOESY spectra for $H_{1\text{-}47}$ and $E_{1\text{-}52}$ in order to minimize the effects of spin diffusion. The additional His-residues of the His-tag at the N-

terminus, which are essentially unstructured, were not used in the structure calculation. In case of subunit F, the ^{1}H sequential assignments were achieved mainly using homo-nuclear 2D NOESY (τ_m = 100 ms) and TOCSY spectra in both H_2O and $^{2}H_2O$ solutions at 600 MHz at 15 °C; this approach was adopted because labeled protein samples became available only later in the work. Once labeled protein is available, assignments were extended to include backbone ^{15}N and ^{13}C signals using ^{15}N NOESY-HSQC, ^{13}C NOESY-HSQC, as well as triple resonance backbone experiments (HNCA, CBCA(CO)NH). The distance restraints for the structure calculation were collected from 3D ^{15}N-^{1}H NOESY-HSQC and ^{13}C-^{1}H NOESY-HSQC by manual and automatic assignments for which CYANA 2.1 was used. The experiments used for H_{1-47} and E_{1-52} were ^{15}N HSQC, HNCACB, CBCA(CO)NH, H(CC)(CO)NH-TOCSY and ^{15}N NOESY-HSQC. Dihedral angle restraints were calculated from chemical shifts using Torsion Angle Likelihood Obtained from Shift and sequence Similarity (TALOS) (CORNILESCU *et al.*, 1999) and the overall secondary structure were predicted from Chemical Shift Index (CSI) and NOE pattern. Upper bound for all NOE distance restraints were initially set to 5 Å and adjusted for non-stereo-specifically assigned methylene and methyl protons using the method described originally for DYANA and detailed by Güntert (GÜNTERT *et al.*, 1997). Experimental evidence for hydrogen bonds was obtained from hydrogen/deuterium exchange in per-deuterated buffer at 15 °C over a period of 24 h for subunit F. Amide protons that are associated in regulatory secondary structural elements are generally longer lived than those present in the flexible part of the protein. Thus, the slow amide exchange rate can be used as a measure for the presence of hydrogen bond. For the determination of the H/D exchange factors, subunit F sample was lyophilized and resuspended in 100% D_2O. A 2D ^{1}H-^{15}N-HSQC spectra were obtained immediately after the lyophilized protein was dissolved in 0.5 ml of D_2O. The dead time of the experiment was about 30 min. The structure calculations were performed with the CYANA 2.1

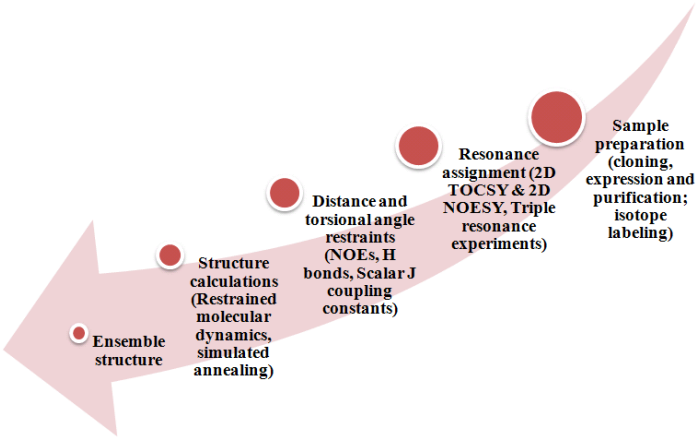


Figure 2.2.9: A flow chart illustrating the general strategy for determining the three dimensional structure by high resolution NMR spectroscopy. Structure calculations are done in an iterative manner where calculated structures are used to identify more NOEs.

program package (GÜNTERT *et al.*, 1997), which uses simulated annealing in combination with molecular dynamics in torsion angle space. With the first round of structure calculations, only unambiguous long-range NOE constraints were used to generate a low-resolution fold for the structure. Assignments of ambiguous NOE cross-peaks were made by applying a structure-aided filtering strategy in repeated rounds of structure calculations. Starting *ab initio*, 100 conformers were calculated in 8,000 annealing steps each. The program MOLMOL (KORADI *et al.*, 1996) was used to visualize the result of ensemble of minimized conformers.

2.2.6.5 Correlation time (τ_c) and relaxation measurements

The correlation time was estimated using the equation:

$$\tau_c = ([6(T_1/T_2) - 7] /4)^{1/2} /\Omega_N 2\pi$$

where T_1 and T_2 are the mean longitudinal and transverse relaxation times, and Ω_N is the Larmor frequency of ^{15}N at the field strength of 14.6 T and 16.4 T were used here. The ^{15}N longitudinal and transverse relaxation time constants, T_1 and T_2, respectively, were determined by collecting a time series of ^{15}N-HSQC spectra with sensitivity enhancement. ^{15}N relaxation measurements were acquired on Bruker Avance 600 and 700 MHz spectrometer equipped with cryo probe at appropriate temperature using the triple-axis gradient enhanced and sensitivity-enhanced ^{1}H-^{15}N HSQC experiment provided in the Bruker pulse sequence library.

For T_1 measurements, the spectra were collected with relaxation delays of 5, 40, 80, 130, 210, 330, 470, 630, 800 and 1000 ms, with repeat experiments at 40 ms and 130 ms for error estimations. For T_2 measurements, data were acquired with delays of 14.4, 28.8, 43.2, 57.6, 72.0, 86.4, 100.8, 115.2, 129.6, 144.0 and 158.4 ms with duplicate points at 43.2 and 72.0 ms. The delay between 180 pulses in the Carr–Purcell–Meiboom–Gill (CPMG) pulse train for T_2 measurements were fixed to 0.9 ms. A total of 2,048 complex data points with 128 complex increments were collected for the relaxation experiments. The relaxation rate constants were determined by fitting the cross-peak intensities to a mono exponential function.

T_1 and T_2 spectra were recorded with spectral widths of 9,615.85 Hz sampled over 2,048 complex points in the ω_2 (^{1}H) dimension, and 2,405.12 Hz over 128 complex points in the $\omega 1$ (^{15}N) dimension with 16 scans for each increment in the indirect dimension. The relaxation delay for both T_1 and T_2 measurement was 1.5 sec.

The steady-state NOE spectra were acquired with a spectral width of 9,615.85 Hz over 2,048 Hz complex points in ω_2 (^{1}H) dimension and 2405.12 Hz over 128 complex points in the

ω_1 (^{15}N) dimension. ^{15}N decoupling during acquisition was achieved using a GARP–4 pulse sequence (SHAKA *et al.*, 1985). The field strength of the CPMG refocusing train about 3.3 kHz and 1.2 ms delay was used between the refocusing pulses (PERVUSHIN *et al.*, 1997). The effects of cross relaxation between ^{15}N-^{1}H dipolar and ^{15}N chemical shift anisotropy were removed by applying 180° pulses during relaxation delay. ^{1}H-^{15}N steady-state NOE values were measured with two different data sets, one collected with no initial proton saturation and a second with initial proton saturation. The proton saturation period was set to 3 sec.

2.2.6.5.1 Data processing and analysis

The HSQC spectra were processed using nmrPipe software (DELAGLIO *et al.*, 1995). Peak intensities were measured and analyzed using the peak picking routine built into SPARKY (GODDARD *et al.*, 1997). The T_1 and T_2 spectra were processed using a sine bell apodization function shifted by 90° for both dimension. The final sizes of the matrices were 2048 × 128 real points after zero filling in both the dimension and the Fourier transformation. An automated baseline correction was applied in both dimensions. The spectra were referenced to DSS signal.

T_1 and T_2 values were obtained by fitting peak intensities using single exponential decay:

$$\mathbf{I\,(t) = I_0}\ \exp\,(-t/\mathbf{T_{1,2}})$$

Where I (t) is the peak intensity, t is the time, I_0 is the intensity at time 0.

The heteronuclear steady-state ^{15}N-^{1}H NOE values were obtained from the ratios of the 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 600 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.

2.2.6.6 Binding studies with NMR

NMR spectroscopy is extensively used to identify the important residues for protein protein interaction (TAKEUCHI *et al.*, 2006; VAYNBERG *et al.*, 2006). With respect to other technique NMR have the advantage of monitoring interactions at amino acid level. ^{1}H-^{15}N HSQC spectrum of labeled protein was used as starting point at optimized condition. Then the ligand protein/peptide was 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 binding partner. 100 - 500 µM of ^{15}N labeled protein was used for the reference spectrum. Unlabeled binding partner was added in increasing amounts to a molar ratio of 1:1 to 1:2 and chemical shift perturbations were monitored in ^{1}H ^{15}N HSQC spectrum (Figure 2.2.10). Experiments were performed on Bruker Avance 600/700 machine using Topspin for acquisition and processing of spectra. Respective spectra were overlapped to monitor chemical shift changes, further analysis were done in SPARKY (KNELLER *et al.*, 1997).

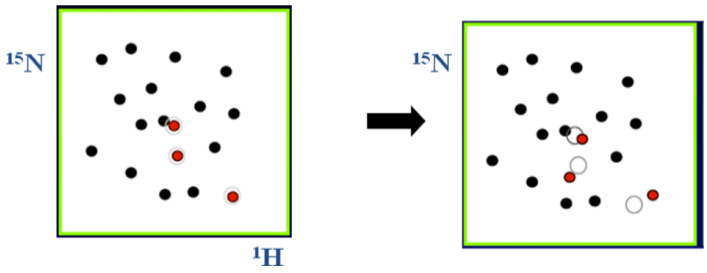


Figure 2.2.10: Illustration of peak position change (red color) in ^{1}H-^{15}N HSQC spectrum upon addition of binding partner (CAVANAGH *et al.*, 1995).

1.2.7 *Fluorescence measurements*.

A *Varian Cary Eclipse* spectrofluorimeter was used, and all experiments were carried out at 20 °C. The samples were excited at 295 nm, and the emission was recorded from 310 to 380 nm with excitation and emission bandpasses set to 5 nm. For titration of the tryptophan fluorescence of peptide $F_{81\text{-}101}$ with ATP, the emission wavelength was 338 nm. Before use, peptide $F_{81\text{-}101}$ and increasing amounts of ATP were incubated in a buffer of 50 mM Tris/HCl (pH 8.0) for 5 mins.

3. Results

3.1 Solution structure of subunit F of A_1A_O ATP synthase from *Methanosarcina mazei* Gö1

3.1.1 Introduction

Catalytic-site events (ATP synthesis and ATP hydrolysis) are connected to ion translocation through conformational changes of the subunits in the stalk region of the enzyme, A_1A_O ATP synthase. Subunit F, the smallest subunit in A_1 ATPase, is a part of the central stalk domain of A_1A_O ATP synthase (COSKUN *et al.,* 2002; SCHÄFER *et al.,* 2006b) and interacts with subunit B in the A_1 part, depending on the type of nucleotide bound to the enzyme. This important observation was noticed from the cross linking studies of the enzyme A_1A_O ATP synthase from *Methanocaldococcus jannaschii* (SCHÄFER *et al.,* 2006b). When the enzyme was incubated with 5 mM MgAMP-PNP at 4 °C before EDC treatment, a 63 kDa B-F product was formed. By contrast there was lower cross-link formation in the presence of MgATP. The cross-link product completely disappeared when the enzyme was incubated with MgADP or MgADP + Pi. MALDI mass spectroscopy identified that the zero length cross-link product was formed between the C-terminal region $_{388}$DLVAVVGEEALTDR$_{401}$ and $_{82}$IDPVKELIR$_{90}$ of subunit B and subunit F, respectively. Interestingly, the peptide $_{388}$DLVAVVGEEALTDR$_{401}$ is at a similar position to the so called DELSEED region of the nucleotide binding subunits α and β of F_1F_O ATP synthases (AGGELER *et al.,* 1998). The cross-link study with EDC identified that subunit F is also in close contact to subunit C and subunit D. Thus, structure characterization of subunit F is very essential, in order to understand its functions in the A_1A_O complex. Most recently, the shape of subunit F of the A_1A_O ATP synthase from the archaeon *Methanosarcina mazei* Gö1 in solution was determined by small angle X-ray scattering, indicating that the protein is elongated and organized into two well-defined domains (SCHÄFER *et al.,* 2006b). Here, the three dimensional structure of F subunit of the A_1A_O ATP synthase from *M. mazei* Gö1 as well as dynamical properties of this subunit in solution has been determined by NMR spectroscopy.

3.1.2 Resonance assignments of subunit F of the A_1A_O ATP synthase

In order to determine the three dimensional structure it is important to do resonance assignment for ^{1}H, ^{15}N and ^{13}C atoms. The sequential resonance assignments of the backbone amide groups are presented in a 2D ^{1}H/^{15}N-HSQC spectrum (Figure 3.1.1), indicating backbone HN and ^{15}N cross-peaks of individual residues. A well-dispersed HSQC (Heteronuclear Single Quantum Coherence) spectrum suggests a folded structure for the protein. The folding of subunit F was also evident from the 1D-^{1}H NMR spectra of the protein, showing a good dispersion of amide peaks as well as up-filled shifted methyls (Figure 3.1.2). Assignments for the backbone HN, ^{15}N, $^{13}C^{\alpha}$, 13C$^{\beta}$ and 13C′ resonances were achieved by a combined analysis of the triple-

resonance HNCA, HN(CO)CA, CBCA(CO)NH, HNCACB and HNCO spectra. The HSQC peaks were identified for all residues except for F15 and G83, S84, E89, D98 which are present in the flexible C-terminal domain. The presence of glycine and serine residues in the C-terminal favors rapid amide exchange rate and thus, amides of these residues (G83, S84, E89, D98) were not observable in the HSQC spectrum. The calculated structure of subunit F also confirms and supports the missing assignments, as this protein exhibits a rigid N-terminal domain consisting of the first 78 residues and a detached flexible C-terminal domain consisting of the last 20 amino acid residues. The absence of the N-terminal F15 amide resonance could be due to high solvent exchange rate at this region of the helix. The hydrogen deuterium exchange rate at this region also clarifies that the amide exchange rate is faster in these regions (Figure 3.1.3).

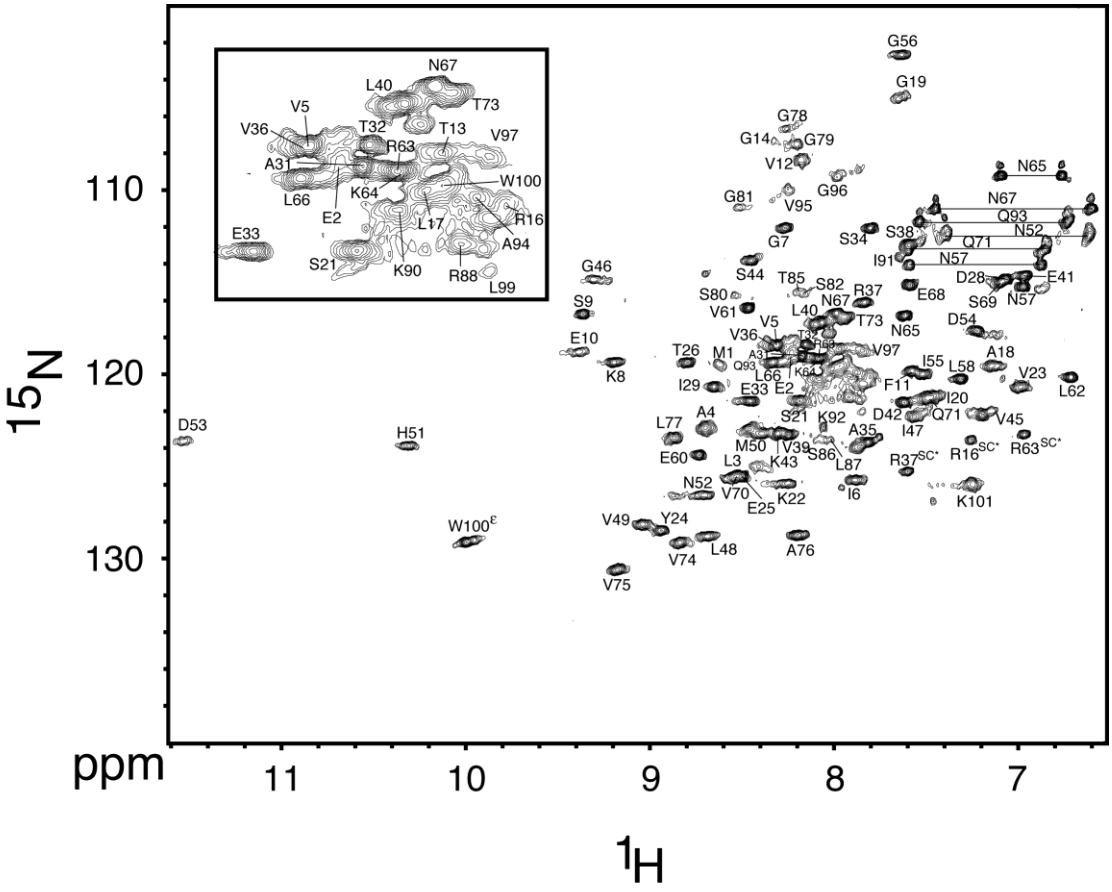


Figure 3.1.1: ^{1}H-^{15}N HSQC spectrum of subunit F of the methanogenic A_1A_O ATP synthase in 25 mM sodium phosphate buffer (pH 6.5) at 288 K. Backbone and side chain (sc) amide assignments are shown for each residue. HSQC cross-peaks of side chain of residues R_{16}, R_{37} and R_{63} are folded in the ^{15}N dimension and indicated by "SC*". Signals from side chain NH2 groups are connected by horizontal lines. The overlapping region is expanded (GAYEN *et al.*, 2007).

The $C^{\alpha}H$ and $C^{\beta}H$ proton resonances and side-chain carbon and proton resonances were assigned from the H(CCO)NH, C(CO)NH and HCCH-TOCSY spectra. In the ^{1}H-^{15}N HSQC spectrum, different line widths were observed for a number of residues. In order to investigate in detail, HSQC spectra at various concentrations ranging from 0.1 to 1 mM were taken at identical NMR conditions. Analysis of these spectra indicated that there is no measurable difference either in the line-widths or in the chemical shifts of ^{1}H and ^{15}N resonances. Moreover, small-angle X-ray scattering (SAXS) data also indicated that the protein is monomeric in the solution (SCHÄFER *et al.*, 2006b) at the concentration used. These experiments clearly suggest that the broadness of the peak in 2D HSQC is due to the conformational changes of the protein, altering the environment of a nuclear spin and modulating chemical shift. When this process is on the microsecond to millisecond timescale and the chemical shift changes are large, the NMR signal broadens.

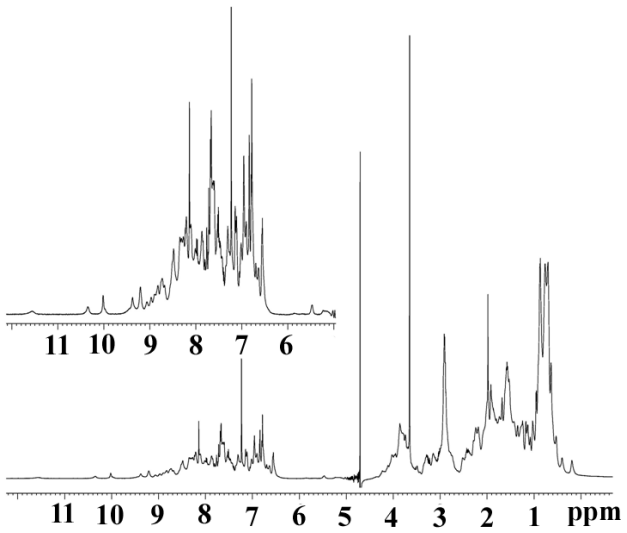


Figure 3.1.2: One dimensional ^{1}H NMR spectrum of subunit F recorded at 288 K in 25 mM PO_4 buffer (pH 6.5) on Bruker Avance 600 MHz machine. Nice dispersion of resonance peaks in HN region, α-protons and uphill shifted methyl peaks are the characteristics of natively folded protein. Strong peaks at 4.7 ppm represents water signal.

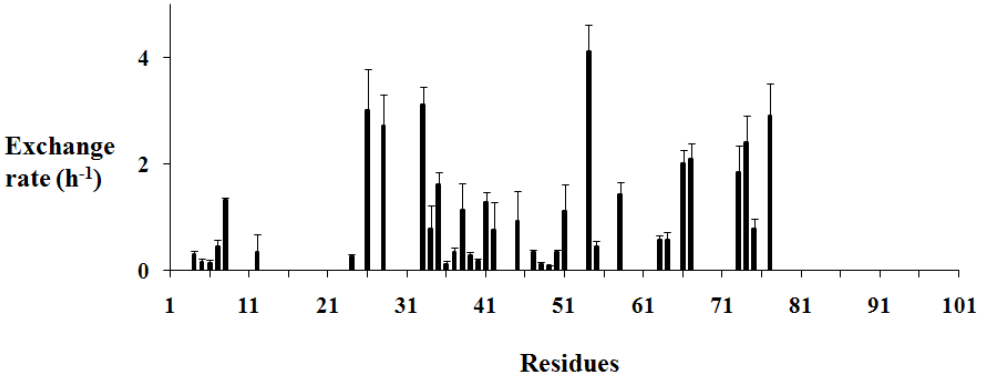


Figure 3.1.3: Exchange rate (h^{-1}) was plotted against residues for subunit F. The exchange rate has been calculated from the hydrogen deuterium exchange experiment where the protein was dissolved in 100% D_2O and the loss of intensity of the signal was monitored through ^{1}H-^{15}N HSQC experiment over time.

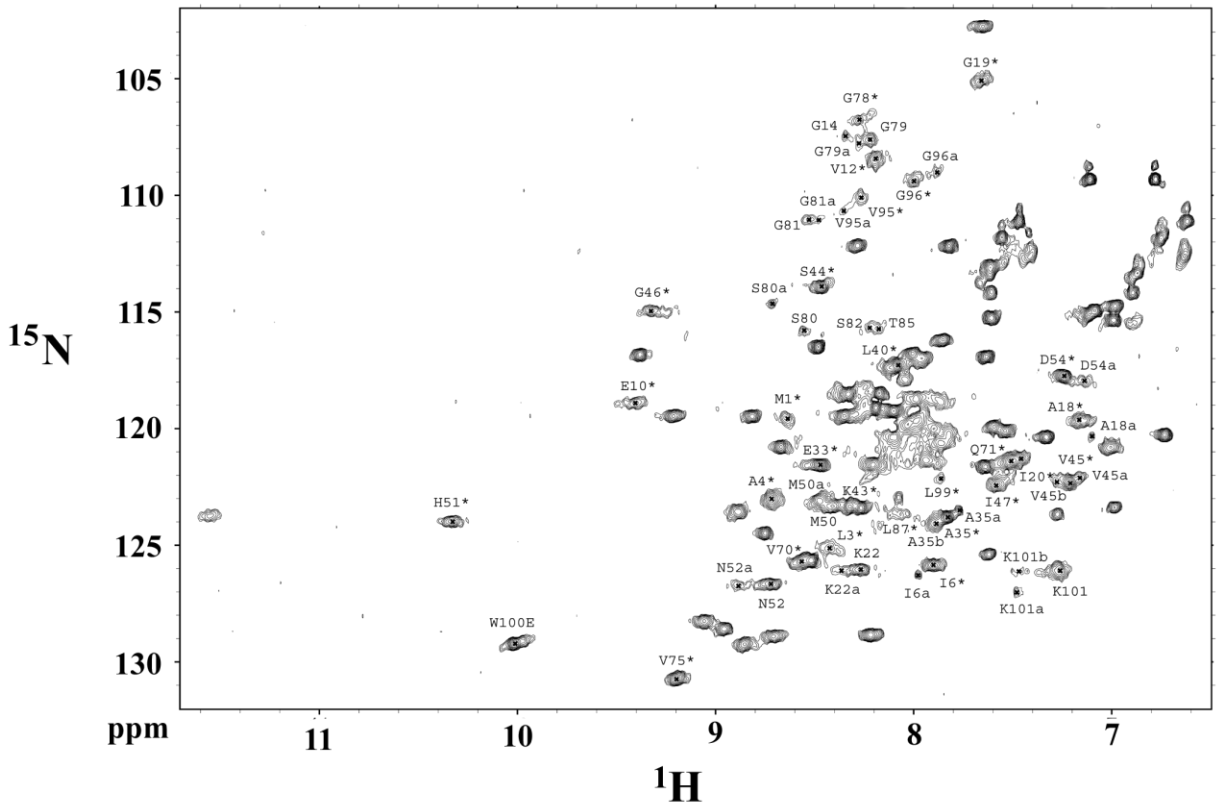


Figure 3.1.4: 2D ^{1}H-^{15}N HSQC of subunit F with different peak shapes and intensities are indicating chemical exchange phenomena. The unlabelled peaks are single sharp resonances having line width of about 25-30 Hz. The peaks marked in single letter amino acid residue with asterisks indicate residues undergoing chemical exchange. The peaks marked in single letter amino acid residue with subscripts *a* and *b* indicate multiple peaks for the same residue. For example, A35, K101, and V45 have multiple peaks and are marked with subscripts *a* and *b* (GAYEN *et al.*, 2007).

Three different types of line shapes of the peaks were observed in the HSQC spectrum. There are single intense peaks and single peaks with exchange broadening and two or more peaks appear with skewed populations. Figure 3.1.4 represents the 2D ^{1}H-^{15}N HSQC of subunit F, where the unlabelled peaks are single sharp resonances. 1D projection of the representative HSQC peaks are shown in Figure 3.1.5. Single sharp peaks have been shown for residues S9, V61, L62 and E68 peaks with exchange broadening as well as two or more peaks for the same residue have been shown for E10, A18, K22, G46 and A35, N52, G78, G96 respectively. Interestingly, all these single sharp peaks are found to be clustered in the middle of the protein sequence representing residues in β2, α2, β3, α3 and β4 secondary structural elements (Figure 3.1.5B). The intensities of the peaks are very low for the C-terminal residues, indicating that this region is flexible compared to the N-terminal domain. This difference in intensity of HSQC peaks is clear evidence that the protein undergoes conformational exchange on the microsecond to millisecond time scale. Structures of subunit F were calculated for one major conformer having intense peaks.

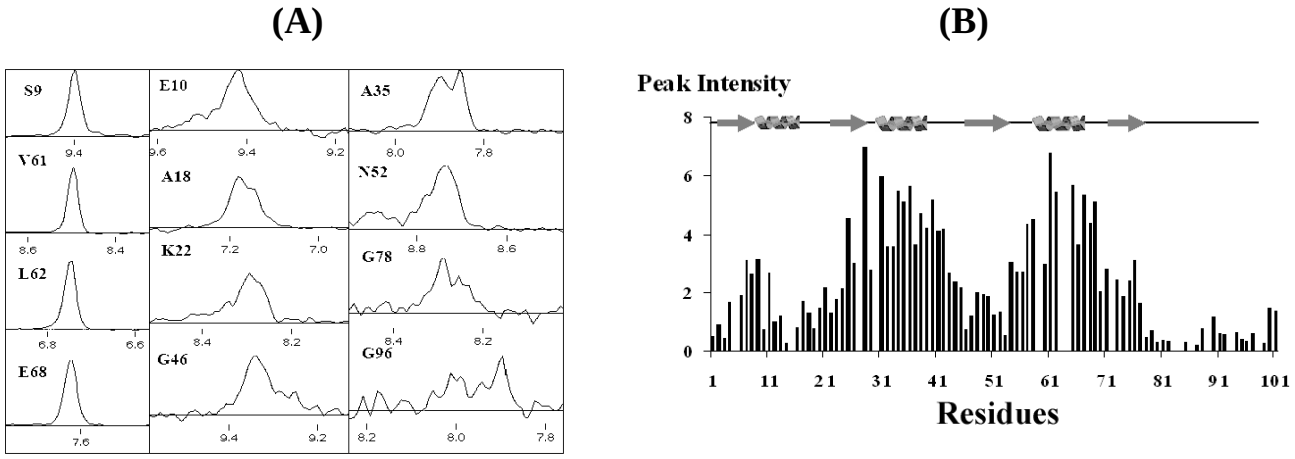


Figure 3.1.5: (A) 1D projection of representative ^{1}H-^{15}N HSQC peaks of subunit F are plotted in the same Hz/pt scale. The column 1 shows single sharp peaks, column 2 single peak with exchange broadening and column 3 shows two or more peaks for the same residue. The residue along with the sequence number is marked in the left side of the peak. (B) Bar diagram showing intensities of the HSQC cross peaks as a function of residue position of the subunit F. Intensities were measured using SPARKY (GAYEN *et al.*, 2007).

3.1.3 Solution structure of subunit F of A_1A_O ATP synthase from *M. Mazei* Gö1

The three dimensional solution structure of subunit F shows that the protein is an elongated molecule. The structure was calculated on the basis of a total of 1399 nontrivial NMR-derived distance restraints. Figure 3.1.6 shows an overlay of the 10 lowest energy structures of the protein and the statistics are given in table 3.1.1.

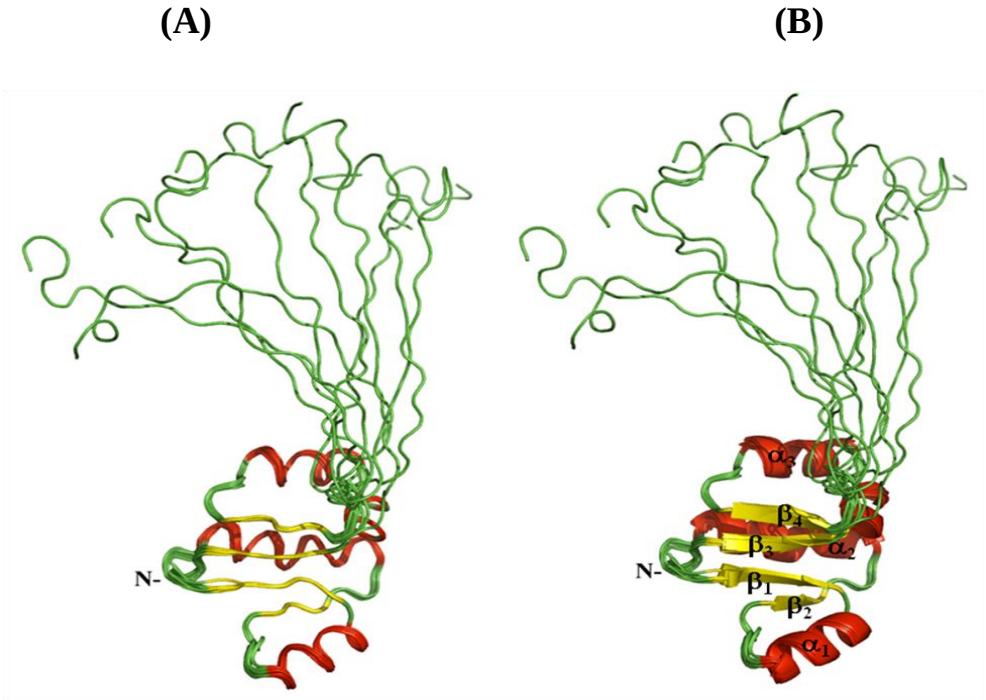


Figure 3.1.6: Backbone (N, Cα, C′) (A) and ribbon (B) diagrams of 10 low-energy NMR-derived structures of subunit F of the A_1A_O ATP synthase from *M. mazei* Gö1 (GAYEN *et al.*, 2007).

Total number of NMR restraints	1399
Unambiguous NOE peaks	
Intraresidue (i=j)	301
Sequential (|i-j|=1)	377
Medium-range (2 ≤|i–j|≤ 4)	261
Long-range (|i–j| > 4)	297
Dihedral angle constraints	163
Hydrogen-bond distance restraints	16
No. of restraint violations[1]	
Total number of restraint violations > 0.5 Å	0
Total number of dihedral angle violations > 5 °	0
Ramachandran plot[2] (%)	
Residues in most favored regions	70.6
Residues in additionally allowed regions	28.9
Residues in generously allowed regions	0.4
Residues in disallowed regions	0.1
Average r.m.s.d. to Mean (Å)	
Backbone (residues 1–78)	0.39 ± 0.09
Heavy atoms (residues 1–78)	0.78 ± 0.08

[1]There are no distance violations greater than 0.5 Å or dihedral-angle violations greater than 5º. All residues are included in the final ensemble. [2]Ramachandran plot was for all the residues.

Table 3.1.1: Structural statistics for the ten selected structures of subunit F of the A_1A_O ATP synthase after energy minimization (GAYEN *et al.*, 2007).

These structures have an overall r.m.s.d. of 0.39 ± 0.09 Å for backbone atoms and 0.78 ± 0.08 Å for all heavy atoms in the well-ordered regions of the three-dimensional structures (residues 1-78) of the protein. High r.m.s deviations with loose C-terminal domain of subunit F reflect the difficulty in determination due to the lack of distance and dihedral restraints. The dimensions of the elongated subunit F (78.7 Å) are approximately 24.2 Å x 28.5 Å x 38.7 Å for the N-terminal part. The C-terminal tail has a length of about 40 Å. All these structures have energies lower than 3.7 kcal mol^{-1}, no NOE violations greater than 0.3 Å and no dihedral violations greater than 5°.

Ramachandran plot of residues ϕ and Ψ angles of subunit F has been shown in figure 3.1.7. The secondary structure prediction from chemical shift index (CSI) method in figure 3.1.8A indicates that there is an alternating arrangement of four α-helices and four β-strands and the residues 86-94 have a low probability to form a helix. However, no NOEs has been detected in this region. The location of relevant NOE restraints for α-helices and β strands are shown in figure 3.1.8B and 3.1.8C. From the figure 3.1.8C, it is clear that the structure of subunit F has been calculated mostly using the NOE restraints from the N-terminal domain (GAYEN *et al.*, 2007).

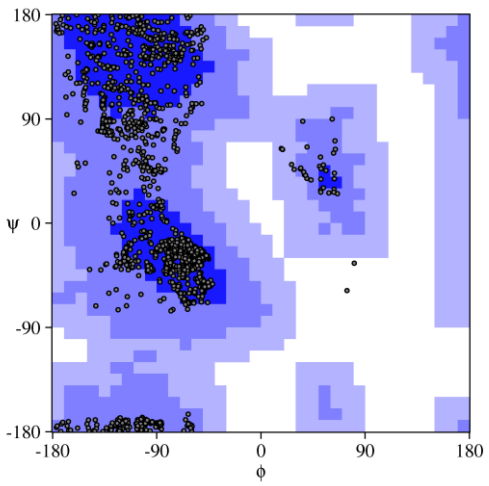


Figure 3.1.7: Ramachandran plot of residues ϕ and ψ angles of subunit F. 70.6% in most favored regions, 28.9% in additionally allowed regions, 0.4% in generously allowed regions and 0.1% in disallowed regions. This plot has been made from CYANA (GÜNTERT *et al.*, 1997).

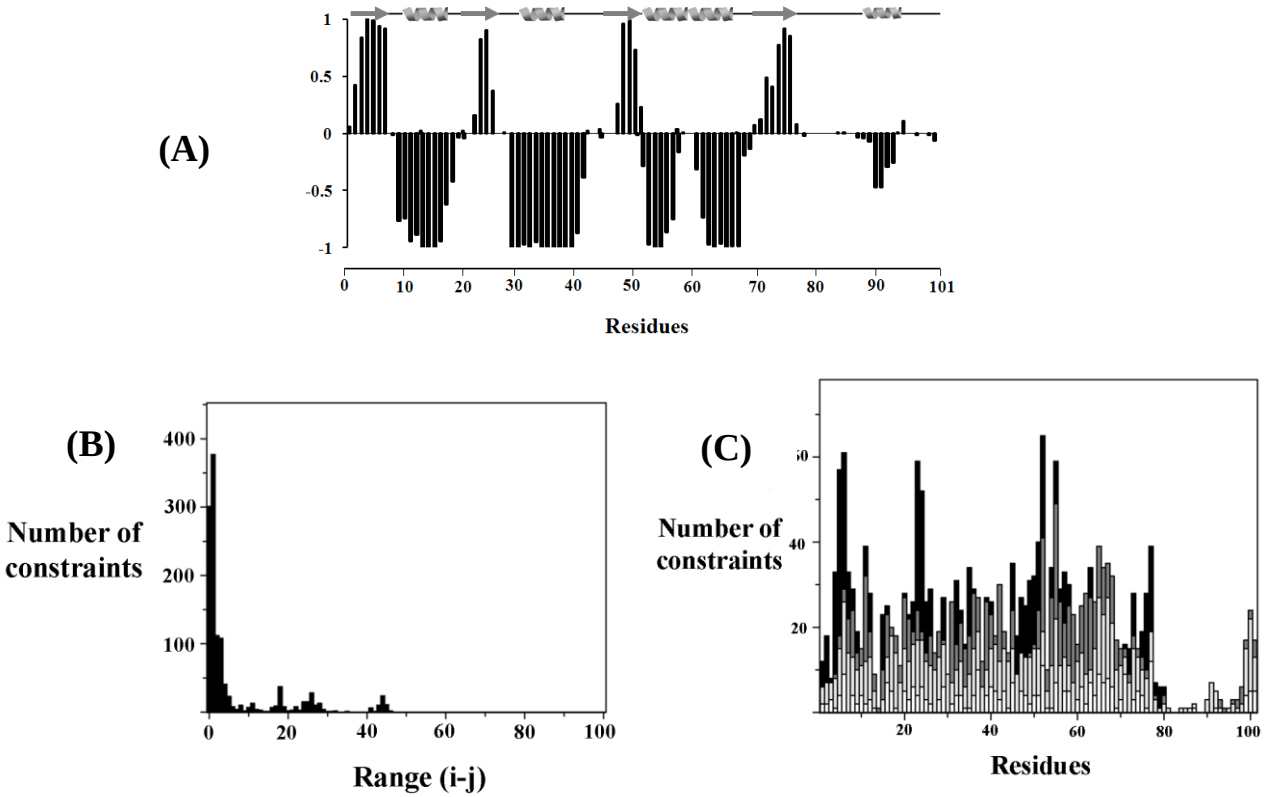


Figure 3.1.8: (A) Secondary structure prediction from chemical shifts of subunit F from *M. mazei* using PECAN software (EGHBALNIA *et al.*, 2005). The positive vertical axis represents the probability of helix and the negative vertical axis represents the probability of sheet structure. Values near zero represents the random coil values. (B) and (C) Plot of relevant distance restraints used in the structure calculation by CYANA (GÜNTERT *et al.*, 1997).

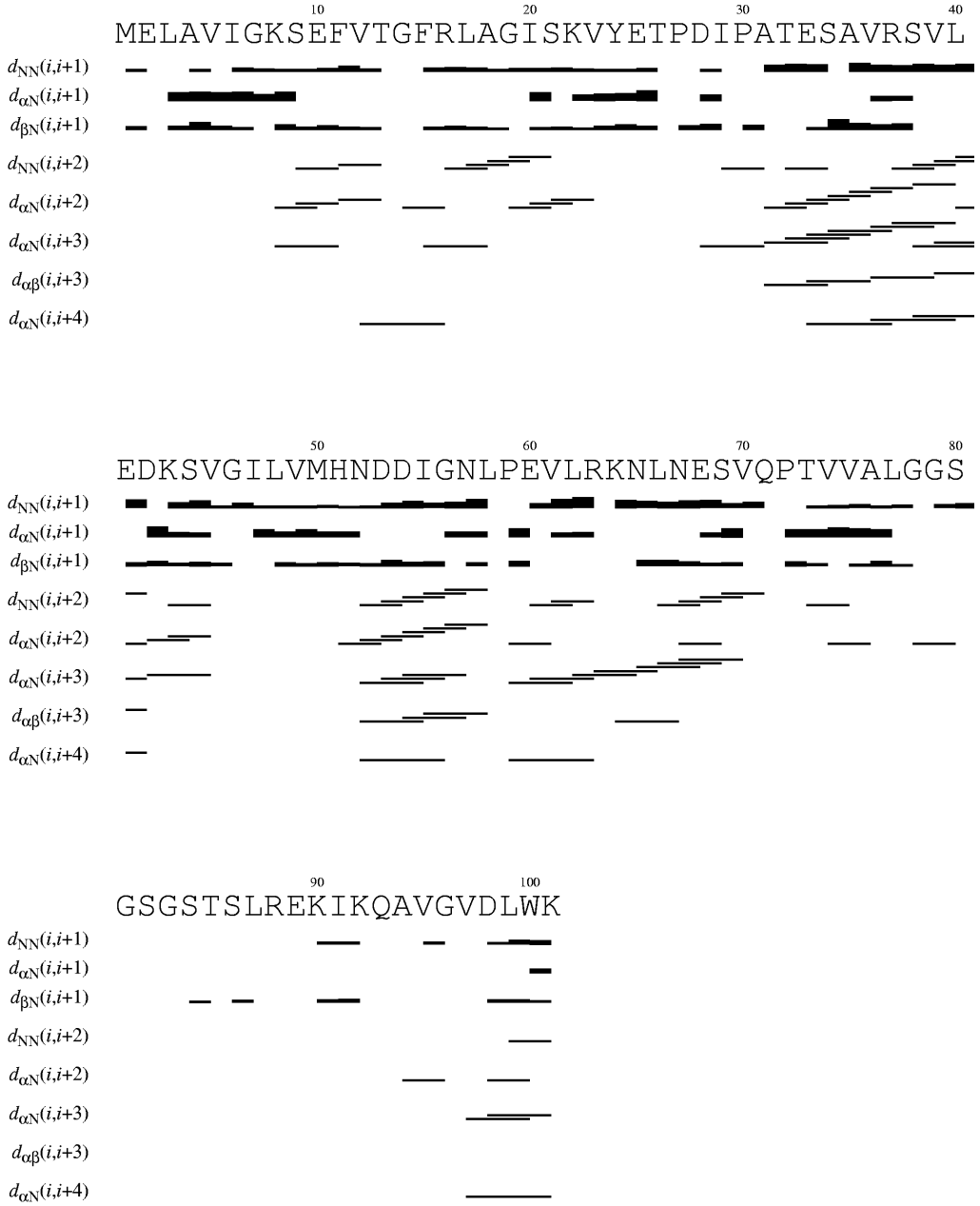


Figure 3.1.9: Amino acid sequence of the subunit F with a survey of the short- and medium range NOE connectivities, which were used to establish the sequence-specific ^{1}H NMR assignment and to identify elements of regular secondary structure (GAYEN *et al.*, 2007)

The overall topology of the solution structure of subunit F is divided into a well-defined N-terminal domain, formed by an α/β fold and an extended C-terminal tail at its very end (Figure 3.1.7). The four-stranded β-sheet of the N-terminal part are oriented parallel to each other and are separated due to the alteration with the three α-helices. These α-helices are made up by the residues S9 - A18, A31 - V39 and V61 - E68, whereby the residues L3 - G7, K22 - T26, V45 - N52 and T73 - A76 form the β-strands (β1-β4) (shown in figure 3.1.10).

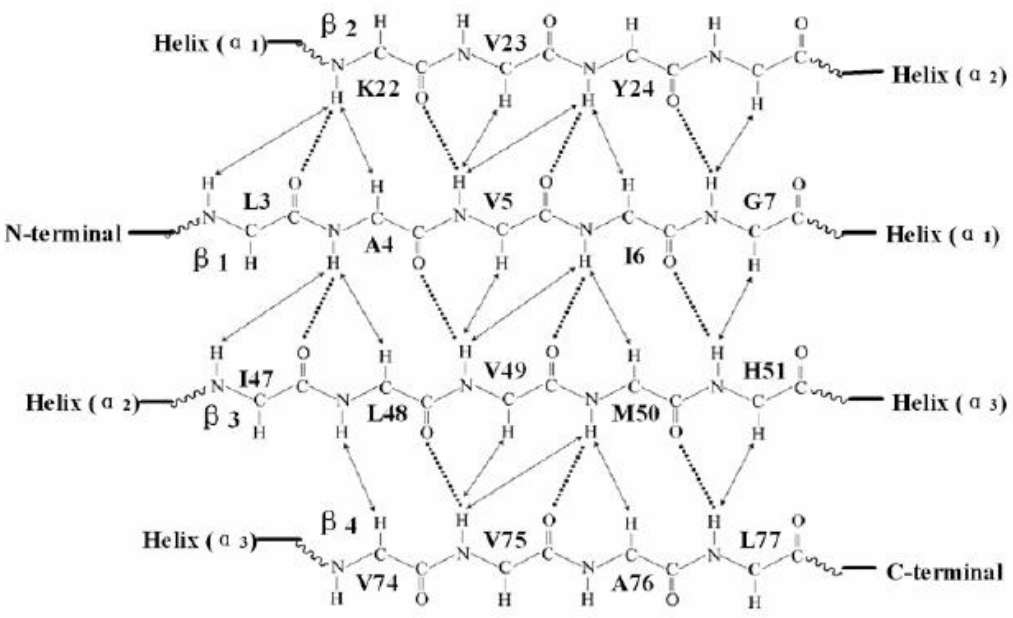


Figure 3.1.10: Schematic representation of the secondary structure of subunit F is shown. Dotted lines indicate inter strand hydrogen bonds and solid lines indicate long-range ^{1}H-^{1}H distances in 4 parallel β-strands of the subunit F (GAYEN *et al.,* 2007).

3.1.4 Intramolecular dynamics of the C-terminal domain

NMR relaxation experiments of proteins provide uniquely detailed and quantitative information about the conformational dynamics on the pico- and nanosecond time scale. There are NMR spin relaxation experiments of ^{15}N, ^{13}C or ^{2}H labeled protein, which are used to determine the longitudinal T_1, transverse T_2, heteronuclear NOE and other relaxation pathways. ^{15}N amide relaxation analysis of subunit F was done in order to get details about the flexibility of the C-terminal domain relative to the $(\alpha/\beta)_4$ part in more detail, which might be important for the coupling function of this subunit (GAYEN *et al.,* 2007). A first indication that the N- and C-terminal domains of subunit F are in loose contact in the solution structure is the observation that there are no long range NOEs between residues of the N-and C-terminal domains. Furthermore, a backbone ^{15}N-^{1}H NOE spectrum was recorded and backbone relaxation parameters were measured. As shown in figure 3.1.5B, the peaks in the C-terminal regions are weakened, indicating that the two domains of the F subunit tumble as two different entity. Figure 3.1.11B-D, F-H illustrates the longitudinal R_1 and transverse R_2 relaxation rate constants and R_2/R_1 ratios. R_1 and R_2 relaxation rates as well as ^{15}N-^{1}H NOE were determined at two different field strengths of 14.1 and 16.4 T respectively. Both R_1 and R_2 rates are fairly constant across

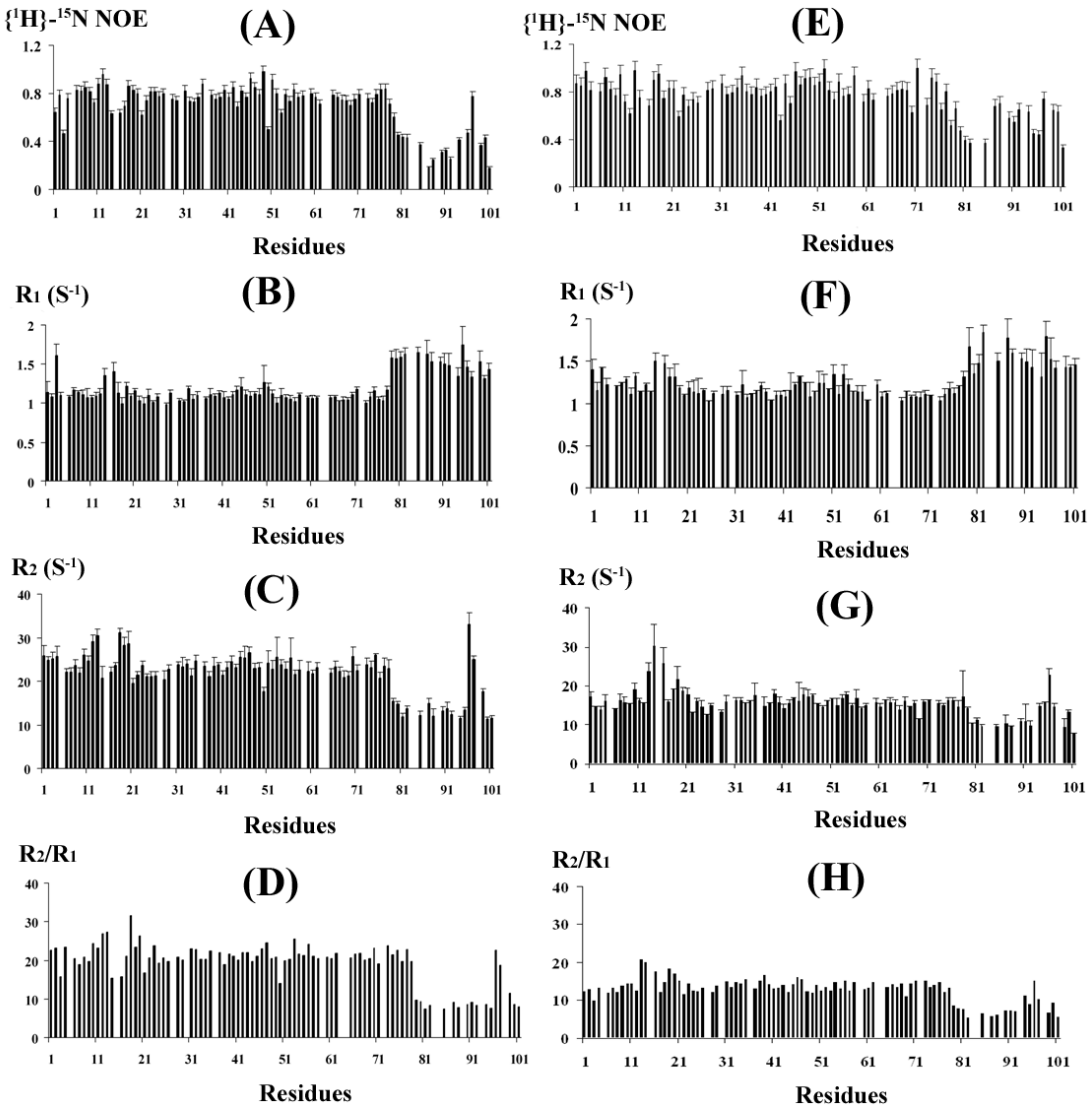


Figure 3.1.11: Comparison of NMR data collected on 600 MHz (A-D) and 700 MHz (E-H) spectrometers. (A, E) Steady-state ^{1}H-^{15}N NOE enhancement data. The less intense bar graph indicates that this portion (80-101) of the protein is very dynamic. (B-D, F-H) Experimental longitudinal (R_1) and transverse (R_2) ^{15}N relaxation rates and *R*2/*R*1 ratios obtained for subunit F*Mm* at 600.13 and 700.13 MHz, respectively. Almost all residues in the well-ordered region exhibit a rather uniform dynamic behavior, suggesting low backbone mobility for this region (1-78) of the protein structure. All C-terminal residues show a rather non uniform dynamic behavior, suggesting high backbone mobility for this C-terminal region of the protein structure. The error bars represent the SD (GAYEN *et al.*, 2007).

the primary sequence for the N-terminal $(\alpha/\beta)_4$ part. But in the C-terminal end of the protein (around 20 amino acids) there is a discernible increase in R_1 and a decrease in the R_2 values was observed. The average values of the relaxation parameters R_1 and R_2 and NOEs in the ordered regions (residues 1-78) were calculated at 14.6 and 16.4 T are 1.104, 23.469 and 0.774 sec^{-1} and 1.174, 16.334 and 0.809 sec^{-1} respectively. The non uniformity in the R_1 and R_2 relaxation rate constants signifies that the protein is tumbling anisotropically. Overall correlation times of 14.5 ns for 14.6 T and 9.94 ns at 16.4 T were obtained for this subunit F protein according to the equation described in the section 2.2.6.5. The backbone dynamics of subunit F was also

investigated by model-free analysis taking the relaxation parameters at two different field strengths of 14.1 and 16.4 T. The results for the model-free analysis are shown in figure 3.1.12A-C. The residues in the N-terminal globular part have S^2 ~0.8 indicating fairly rigid overall structure. Whereas, the flexible residues from the C-terminal loop has $S^2 < 0.6$. To examine further the presence of conformational exchanges on the micro- to millisecond time scale in subunit F, a rough estimation of the amount of conformational motions in the backbone of subunit F has been made from the difference in the R_2 values obtained for the two extreme values of ν_{CPMG}^{min} - ν_{CPMG}^{max}. As shown in figure 3.1.12D, a difference between the two experiments was observed in a number of residues in subunit F with significant positive values of $R_2(\nu_{CPMG}$ min) - $R_2(\nu_{CPMG}$ max), indicating that the conformational exchange occurs at multiple backbone sites in the protein. The corresponding residues probed by CPMG experiments were located primarily in the three different regions of the protein, in the N-terminal, loop1-α1-loop2 (residues 8-23), in the middle part, β3-loop 3 (residues 47-53) and in β4 (residues 75, 76, and 79) and the C-terminal part of subunit F, suggesting a chemical exchange phenomenon in these regions.

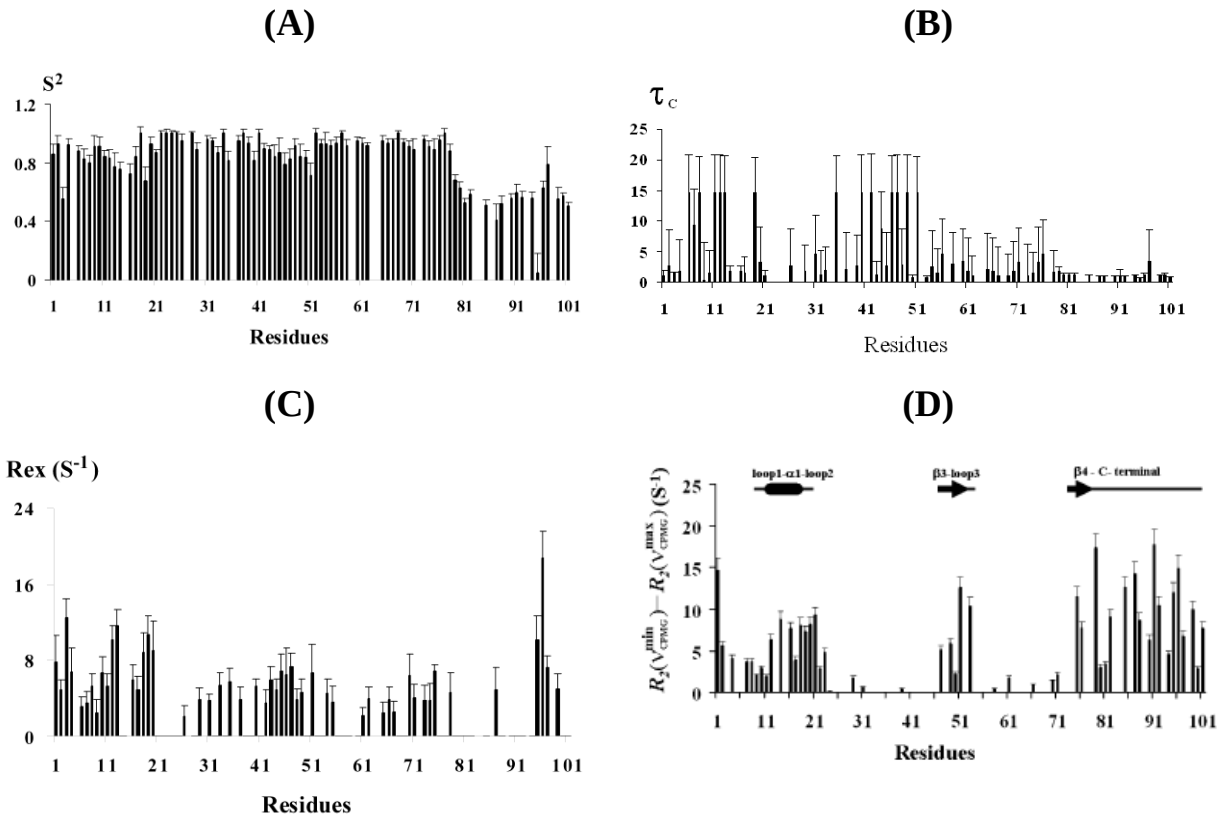


Figure 3.1.12: (A-C) Plots from model free analysis. (D) Conformational exchange in subunit F of the methanogenic A_1A_O ATP synthase. Measurement of backbone conformational motions from the difference values of R_2(CPMG) at minimal and maximal CPMG pulse repetition rates. $R_2(^{CPMG}{}_{min})$ - $R_2(^{CPMG}{}_{max})$ are displayed as a function of the subunit F sequence. Positive values are expected for residues that undergo conformational exchange.

3.1.5 Solution Structure of C-term of Subunit F (F_{81-101})

The NMR structure of subunit F revealed a globular N-terminal and an elongated C-terminal (GAYEN *et al.*, 2007). However, the secondary structure prediction from chemical shifts shows that there is a propensity to form a helix structure for the residues 86-94 in the entire subunit F (Figure 3.1.8A). Here, I investigated the structure of the 21 mer peptide from the C-terminal end of F subunit (F_{81-101}) using NMR spectroscopy.

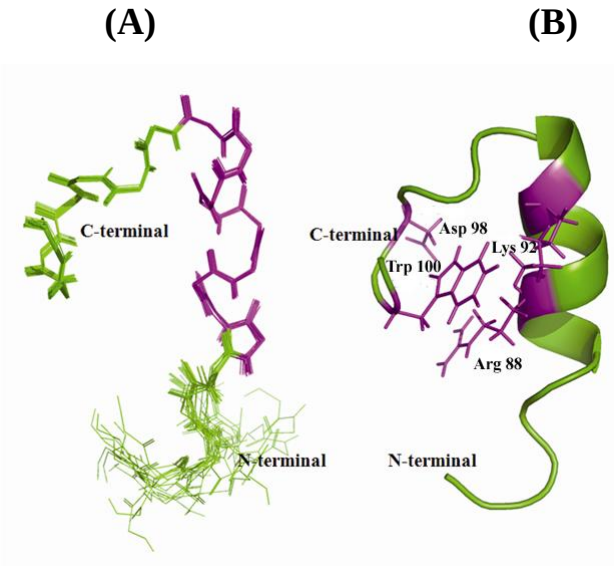


Figure 3.1.13: NMR structure of the peptide F_{81-101} (A) superimposition of twenty lowest energy structures. (B) Cartoon representation of a single peptide structure showing orientations of R88, K92, D98 and W100 side chains.

Distance restraints	
Total	255
Intraresidue (i – j = 0)	53
Sequential (|i-j| = 1)	102
Medium-range (2 ≤ |i–j| ≤ 4)	66
Long-range (|*i*–*j*| ≥ 5)	34
Average number of violations	
Distance violations > 5 Å	0
Ramachandran plot (%)	
Residues in most favoured regions	63.2
Residues in additionally allowed regions	36.8
Residues in generously allowed regions	0.0
Residues in disallowed regions	0.0
Average RMSD to Mean (Å)	
Residue 3-21, backbone (C^{α}, C′, and N)	0.339 ± 0.31

Table 3.1.2: Structural statistics for F_{81-101} (RAGHUNATHAN *et al.*, 2009).

Using the standard procedure for sequential assignment based on homonuclear TOCSY and NOESY experiments, all the residues of the peptide were assigned. Figure 3.1.13 shows an overlay of the twenty lowest energy structures of the peptide (F_{81-101}), which have an overall r.m.s.d. of 0.491 ± 0.31. All these structures have energies lower than -100 kcal mol^{-1}, no NOE violations greater than 0.3 Å and no dihedral violations greater than 5° were observed. The statistics are given in Table 3.1.2. The secondary structure prediction using the ^1H chemical shifts is shown in figure 3.1.14A.The connectivity diagram of the peptide is indicative of a helical conformation with the sequential HN---HN, Hα---HN(*i, i+3*), Hα---HN(*i, i+4*), and Hα---Hβ(*i, i+3*) connectivities (Figure 3.1.14B & C). The peptide displays a short flexible N-terminal tail of the amino acids G81 – T85 and helix structural motif formed by the residues S86 to A94 (Figure 3.1.13).

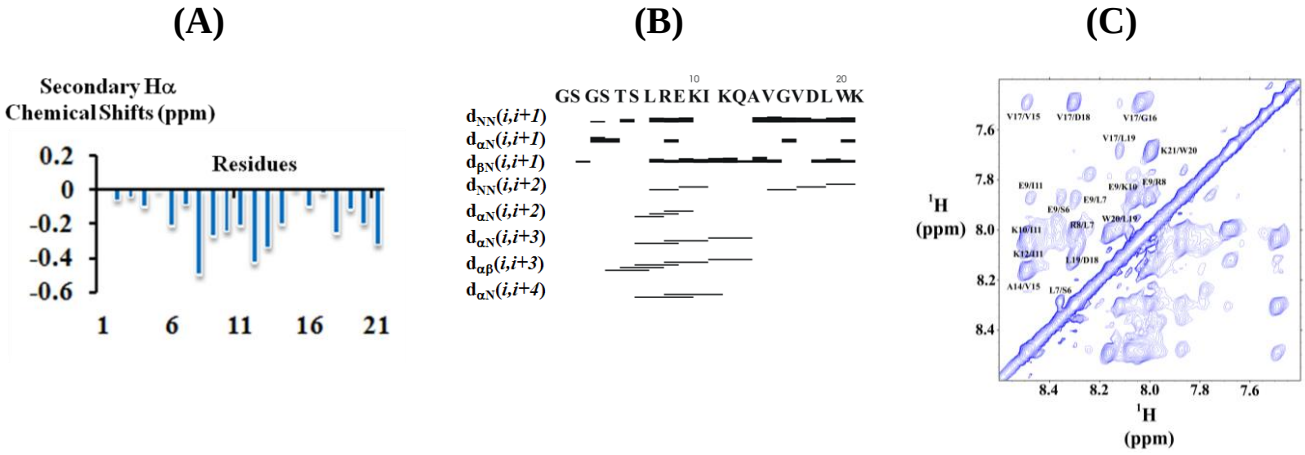


Figure 3.1.14: (A) secondary structure prediction using Hα chemical shifts. (B) NOESY connectivity diagram and (C) HN-HN region of the NOESY spectra for F_{81-101} (RAGHUNATHAN *et al.*, 2009)

The solution structure of F_{81-101} presented exhibits a defined loop formation at its very N-terminus, being part of the glycine rich sequence $_{78}$GCSGSGSTS$_{86}$, reflecting a phosphate loop (P-loop) motif GXXGXGKTV (WALKER *et al.*, 1982). Surface representation of the NMR structure showed that there is both positive and negatively charged surface area. Residues R88, K92 and K101 formed the positively charged area whereas residues E89 and D98 formed the negatively charged surface area. Zero-length cross-link data (SCHÄFER *et al.*, 2006b) and

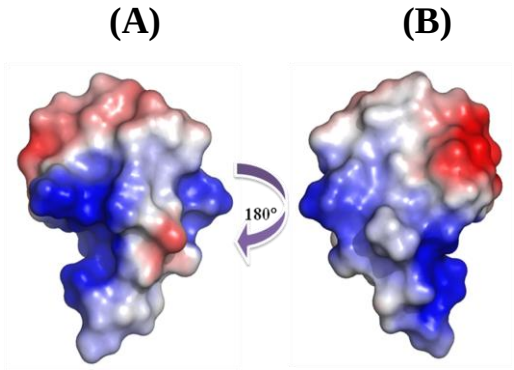


Figure 3.1.15: Electrostatic potential surface representation of the NMR structure of F_{81-101}. Red and blue represent negatively and positively charged areas, respectively.

intrinsic tryptophan fluorescence spectroscopy (GAYEN *et al.*, 2007) showed that the association of subunit B and F occurs via the C-termini of both subunits. Thus, it might be possible that cationic regions on the extended C-terminal of subunit F may interact with the anionic patches on the ventral face of the subunit B. In order to examine whether ATP can bind to isolated $F_{81\text{-}101}$, $^{1}H\text{-}^{15}N$ HSQC experiment was done in absence and presence of different concentration of ATP. A overlaid $^{1}H\text{-}^{15}N$ HSQC spectra of $F_{81\text{-}101}$ in absence and presence of increasing concentration of ATP is shown in figure 3.1.16. HSQC peak positions of most of the residues are shifted with the addition of ATP. The maximum chemical shift change occurs for residues G83, S84, T85, V95, V97, W100 and K101, whereas residues L87, R88, E89, D98 and L99 have medium chemical shift change and residues S86, K90, I91 and A94 have no/mimimum chemical shift change. The change of peak position can occur due to ATP binding as well as binding induced conformational change. Thus, it is difficult to conclude from the results which residues are directly involved the binding. However, the results suggest that the isolated peptide $F_{81\text{-}101}$ can bind to ATP. A model fitting of ATP into the $F_{81\text{-}101}$ structure reveals close contacts of the nucleotide to the P-loop as well as cation-π and π-π interaction of R88 and W100 with the adenine-ring respectively (Figure 3.1.16). α-, β- and γ-phosphate are in close neighborhood to the residues S84, T85 and S82 with distances of distance 3.04 Å, 2.30 Å and 4.02 Å, respectively.

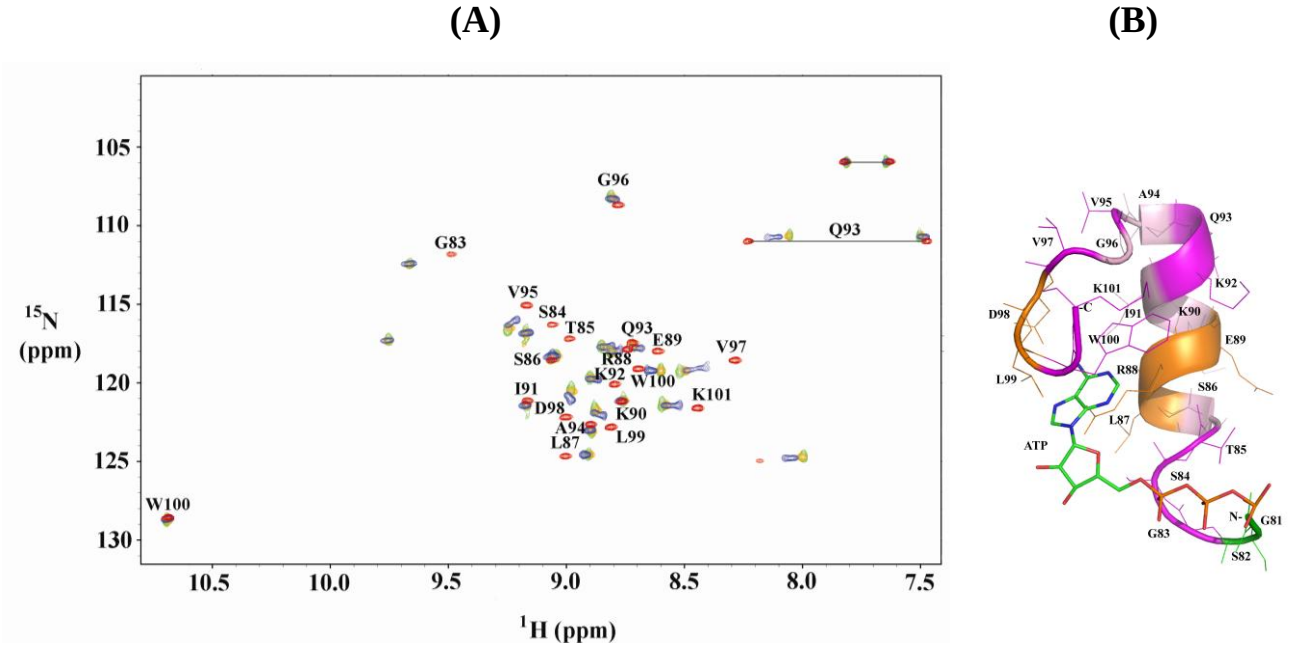


Figure 3.1.16: (A) Overlay of 2D $^{1}H\text{-}^{15}N$-HSQC spectra of $F_{81\text{-}101}$ in the absence (*red*) and presence (*blue, yellow and green*) of 8.33 mM, 16.6 mM and 30 mM ATP. The spectrum was collected in a Bruker Avance 600 MHz spectrometer in 25 mM sodium phosphate buffer (pH 6.5), 40% TFE at 25°C. (B) ATP binding model for the peptide: pink, orange and light pink color residues represent the maximum, medium and no/mimimum chemical shift change respectively in the HSQC spectrum upon ATP addition (RAGHUNATHAN *et al.*, 2009).

The NMR-titration experiment of $E_{85\text{-}104}$ with ATP has shown that the intrinsic Trp residue in $F_{81\text{-}101}$ undergoes a shift in the HSQC spectra after addition of ATP. In order to quantitatively evaluate the binding of the nucleotide the intrinsic tryptophan fluorescence quenching at 338 nm has been performed and the result is shown in Figure 3.1.17. The titration curve has a hyperbolic shape from which a binding constant (K_d) of 560 μM could be determined.

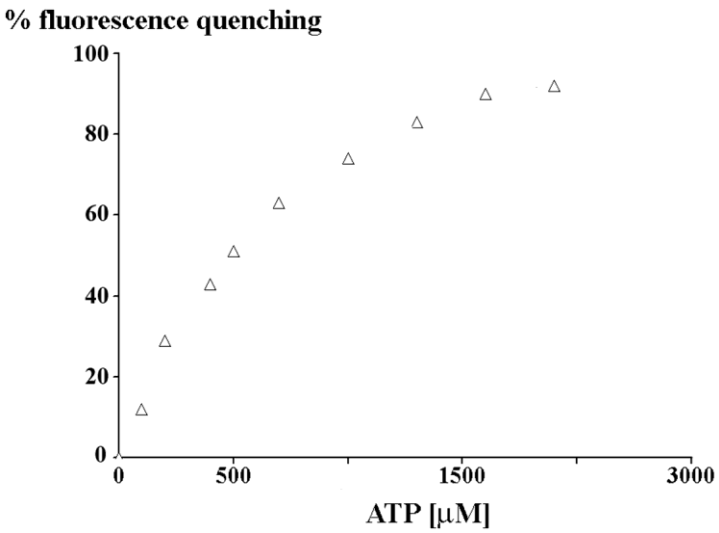


Figure 3.1.17: Fluorescence titration of F81-101 with the ATP. Excitation was at 295 nm and the emission was measured at 338 nm.

3.2 Structural characterization of subunit H of A_1A_O ATP synthase from *Methanocaldococcus jannaschii*

3.2.1 Introduction

The energy coupling between the A_1 headpiece, A_3:B_3 and membrane-embedded ion translocating part A_O domains occurs via the stalk part(s) (GRÜBER *et al.,* 2008). The central stalk is made of subunits C, D and F whereas the subunit H is believed to be a part of the peripheral stalk of the enzyme (MÜLLER *et al.,* 2003; COSKUN *et al.,* 2004). Cross-linking studies of A_1A_O ATP synthase subunits isolated from *M. jannaschii* using the cross-linker ehtyl-3-(dimethylaminopropyl)-carbodiimide (EDC), which can induce cross-linking between carboxyl and amino groups, demonstrated cross-linking of subunits B-F, D-E, A-H, and A-B-D. Among them, crossing-linking between subunits H and A, B and F were found to be nucleotide dependent (SCHÄFER *et al.,* 2006b). Cross-linking of D-E and A-B-D is independent of the nucleotide present. The nucleotide-dependency of A-H crossing-linking that occurs through the C-terminus of subunit H and the N-terminal domain of subunit A is opposite to that of B-F, as the presence of MgADP + Pi seems to induce A-H binding but inhibit B-F binding, while the presence of the non-hydrolyzable ATP analogue, MgAMP-PNP readily induce cross-linking of B-F but not A-H (SCHÄFER *et al.,* 2006b). Understanding the structural and functional role of H subunit is essential since the precise arrangement of this subunit inside the A_1A_O complex is still an unresolved issue (GRÜBER *et al.,* 2008). The goal of the present work was to produce and purify the peripheral stalk subunit H of the A-ATP synthase and to examine the structural and biochemical features of this stalk subunit in solution.

3.2.2 Expression and purification of subunit H

Subunit H was cloned into a modified pET9d-His_3 (GRÜBER *et al.,* 2002) expression vector by using the genomic DNA from *M. jannaschii* ATCC® #43067D™ as the template for PCR amplification. The details of the PCR program and reaction mixture was described under section 2.2.1. Figure 3.2.1 shows a double restriction digestion with *Nco*I and *Sac*I enzymes. Recombinant H subunit was purified as described under section 2.2.3. The protein was found entirely within the soluble fraction. A Ni^{2+}-NTA resin column and an imidazole-gradient (25 – 200 mM) in buffer consisting of 50 mM Tris/HCl (pH 8.5) and 250 mM NaCl was used to separate subunit H from the main contaminating proteins (Figure 3.2.2). Subunit H eluting at

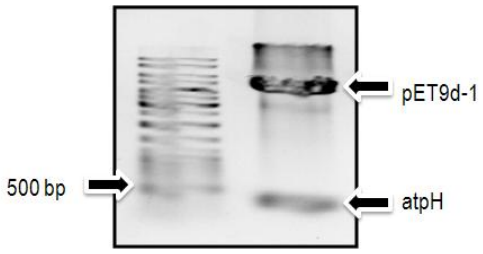


Figure 3.2.1: Cloning of subunit H. DNA agarose gel (1 %) picture showing double restriction digestion with NcoI and SacI enzymes of atpH obtained after ligation of PCR fragment atpH to pET9d-1 (released insert and vectors are shown by arrow).

125 to 200 mM imidazole fractions were collected and subsequently applied to a Superdex 75 HR 10/30 column (GE Healthcare), in order to remove traces of minor contaminants and imidazole. Analysis of the isolated protein by SDS-PAGE revealed the high purity of the subunit. SDS-gels were stained with Coomassie Brilliant Blue G250. Protein concentrations were determined by the bicinchoninic acid assay (BCA; Pierce, Rockford, IL. USA). The SDS-PAGE of the produced recombinant subunit H revealed a prominent band of 12.5 kDa revealing the high purity of the subunit as shown in figure 3.2.3.

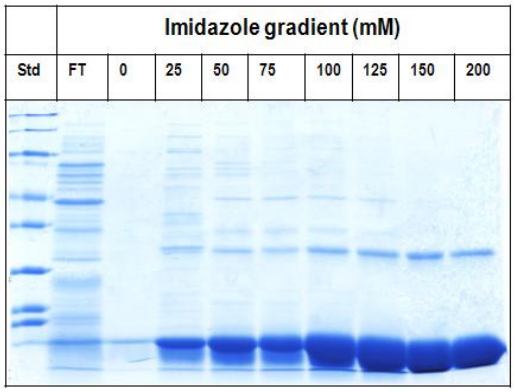


Figure 3.2.2: Ni^{2+}-NTA affinity enrichment of protein from whole cell lysate in buffer 50 mM Tris/HCl, 250 mM NaCl, pH 8.5 using an imidazole gradient (0 – 200 mM) for elution of Ni^{2+}-NTA bound protein. Fractions from 125 – 200 mM imidazole were pooled together for further purification step.

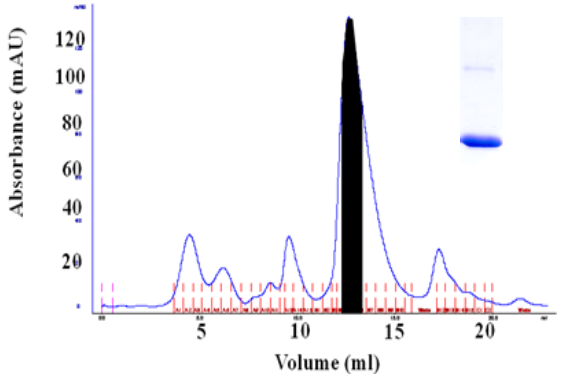


Figure 3.2.3: Purification of protein by gel filtration column Superdex 75 HR 10/30. 125 - 200 mM Ni^{2+}-NTA fractions were pooled and applied and the fractions under the half width of the peak (black) in the chromatogram were collected. Insert shows 1 µl of get filtration purified protein on a 17 % SDS gel

3.2.3. Secondary structure determination of subunit H

The secondary structure content of this subunit was determined from circular dichroism spectra, measured in a far uv range between 185-260 nm. The CD spectra of subunit H measured at room temperature is shown in figure 3.2.4. Spectra were collected in a 60 µl quartz cell (Hellma) at 18 °C at a step resolution of 1 nm. CD spectroscopy of subunit H was performed in a buffer 50 mM Tris/HCl (pH 7.5) and 250 mM NaCl. The spectrum for the buffer was subtracted from the spectrum of the protein. This baseline corrected spectrum was used as input for computer methods to obtain predictions of secondary structure (BOHM *et al.*, 1992). The minima at 222 and 208 nm and the maximum at 192 nm indicate the presence of α-helical structures in the protein. The secondary structure content was calculated to be 78% α-helix and 10% random coil. This result is consistent with secondary structure predictions based on subunit H amino-acid sequence. The molar ellipticity values at 208 nm (corresponding to the α-helix

parallel-polarized amide π-π* electronic transition) and at 222 nm (corresponding to the n-π* transition) are in a ratio of 0.96.

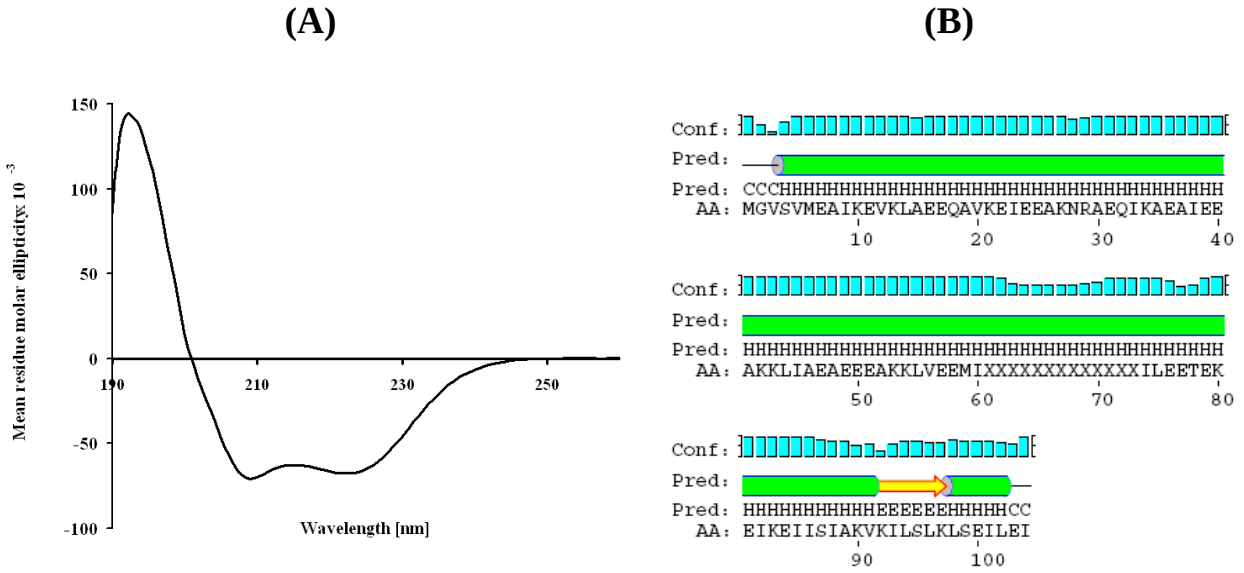


Figure 3.2.4: (A) CD Spectroscopy of subunit H. CD spectrum measured at 18 °C in a Hellma quartz cuvette (60 µl volume) on Chirascan spectropolarimeter (Applied Photophysics) at a step resolution of 1 nm in the far UV range of the spectrum (BIUKOVIĆ *et al.*, 2007) (B) predicted secondary structure by online PSIPRED software, which uses a position specific scoring matrices generated by PSI-BLAST for secondary structure calculation by a two stage neural network algorithm (MCGUFFIN *et al.*, 2000).

Since non-interacting helices typically give ratios of around 0.8, whereby interacting ones have ratios close to 1, the CD spectrum presented indicates, that many of the residues in subunit H are in coiled-coil interaction. Figure 3.2.4B shows the predicted secondary structure by online PSIPRED software, which uses a position specific scoring matrix generated by PSI-BLAST for secondary structure calculation by a two stage neural network algorithm (MCGUFFIN *et al.*, 2000).

The computer-based prediction method "MultiCoil" has been used to identify a coiled-coil motif in the protein (Figure 3.2.5), whose amino acid sequence shows a clear repeating pattern of hydrophobic residues (Figure 3.2.6). The prediction results clearly indicate that this sequence has high propensity to form a two-stranded instead of three-stranded coiled-coil motif in between amino acid residues 25-98.

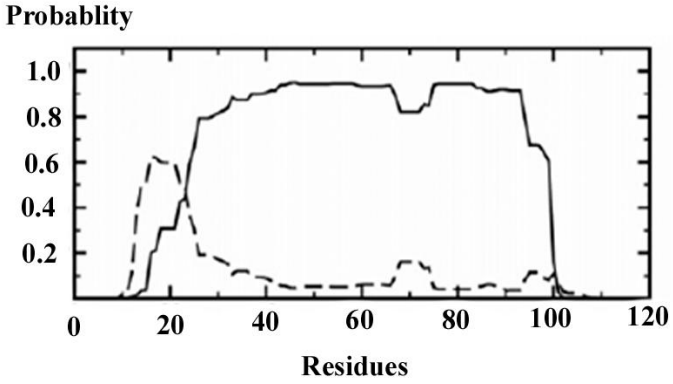


Figure 3.2.5: Probability of two stranded coiled-coil formation (solid line) and of three stranded coiled-coil formation (dashed line) predicted by the multi-coil program (BIUKOVIĆ *et al.*, 2007).

Figure 3.2.6: Sequence alignment of subunits H (A-ATP synthase) and G (V-ATPase) and the soluble domain of subunit *b* (F-ATP synthase) from *M. jannaschii*, *S. cerevisae*, and *E. coli*, respectively. Alignment was generated using AlignX (Vector NTI v9 InforMax). Identical and homologous amino acids are highlighted in black and gray shading, respectively (BIUKOVIĆ *et al.*, 2007).

3.2.4 Low resolution shape and domain structure of subunit H

X-ray solution scattering patterns of subunit H and truncated form of subunit H, $H_{1\text{-}98}$ (BIUKOVIĆ *et al.*, 2007) were kindly measured and analyzed by Prof. Grüber. The radius of gyration (R_g) of subunit H and $H_{1\text{-}98}$ were found to be 5.7 ± 0.2 nm 5.5 ± 0.2 nm respectively, and maximum dimensions (D_{max}) is 20.1 ± 0.3 and 15.1 ± 0.3, respectively. Comparison of the forward scattering with the values obtained from a reference solution of bovine serum albumin

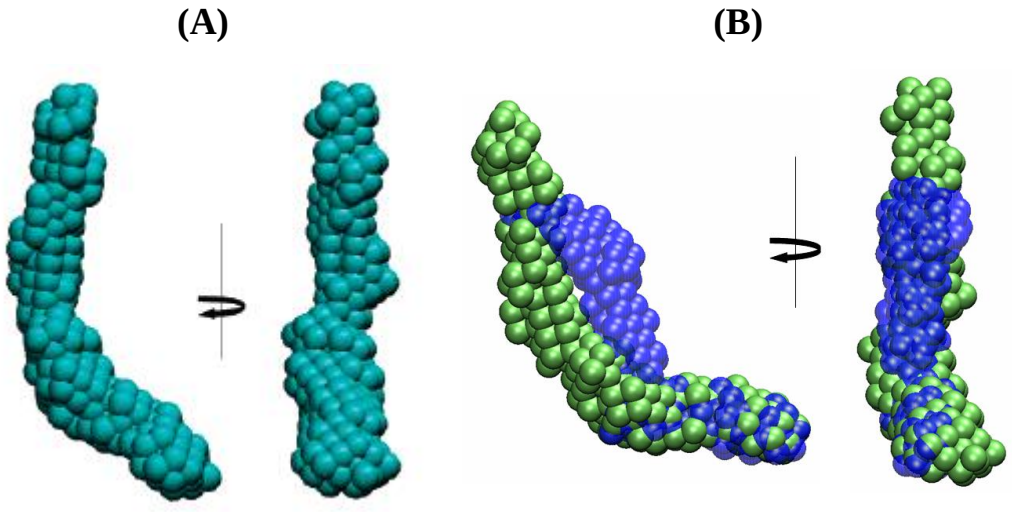


Figure 3.2.7: (A) Low-resolution structure of subunit H determined from SAXS data. (B) Superposition of the shapes of subunit H (green) and the truncated $H_{1\text{-}98}$ form (blue), determined from solution X-ray scattering data (BIUKOVIĆ *et al.*, 2007).

(BSA; 66 ± 2 kDa) yields a molecular mass of 25 ± 2 kDa, independently of the concentration used (2, 4, and 12 mg/ml) for subunit H. The determined mass is about twice the value calculated from the amino acid sequence and indicates that subunit H is dimeric in solution. The obtained shape for subunit H yields a good fit to the experimental data in the entire scattering range. Subunit H appears as a boomerang-shaped molecule of two distinct parts, with about 12.0 and 6.8 nm in length. The low-resolution shape of the $H_{1\text{-}98}$ (BIUKOVIĆ *et al.*, 2007) form is also determined and superimposed on subunit H in figure 3.2.7B. Like the entire H subunit, the $H_{1\text{-}98}$

form has an extended boomerang shape. However, comparison of the two models indicates that the truncation of six amino acids at the C-teminus of subunit H results in a shortening of the upper arm of the molecule, which seems to induce a twisting of the upper part. The shortening of the truncated protein H_{1-98} is reflected by its decreased D_{max} value from X-ray scattering measurement. An intriguing observation is that the dimensions and shape of the lower domain of both molecules do not alter, implying that the C-terminus of subunit H is located in the upper domain and that this domain is important for the stabilization of the quaternary structure of the protein.

3.2.5 Solution structure of C-terminal of Subunit H (H_{85-104})

To further characterize the C-terminus of subunit H, which is interacting with the N-terminal region of subunit A (K_d, 206 nM) (BIUKOVIĆ *et al.*, 2007), attention was focused to the short length peptide of the very C-terminus of subunit H (residues 85-104). Secondary structure prediction of this region of protein shows presence of both helix and strand in the structure (MCGUFFIN *et al.*, 2000). The peptide was dissolved in phosphate buffer pH 6.5 and NMR experiments were recorded on 600 MHz Bruker Avance spectrometer at a peptide concentration of 1 mM.

3.2.5.1 Assignment of the peptide H_{85-104}

Amino acids in the primary sequence of peptide H_{85-104} were sequentially assigned as per the standard procedure. 2D TOCSY and NOESY experiments were recorded at a mixing time of 80 ms and 300 ms respectively. 2D TOCSY and 2D NOESY raw data was processed using in-built Topspin software 1.3 (Bruker biospin). All the amino acids were sequentially assigned. Primary sequence amino acid marking was followed by the assignment of cross peaks by overlaying TOCSY and NOESY spectra using the software SPARKY (GODDARD *et al.*, 1997). Secondary structure prediction was done by using the HA chemical shifts which shows α-helical structure in the middle of the protein (Figure 3.2.8). Identified cross peaks in HN-HN region are shown in the figure 3.2.9A. HN–HN, Hα–HN(i, i+3), Hα–HN(i, i+4), and Hα–Hβ (i, i+3) connectivities were plotted from the assigned NOESY spectrum (Figu3.2.9B) also supports α-helical formation in the C-terminus.

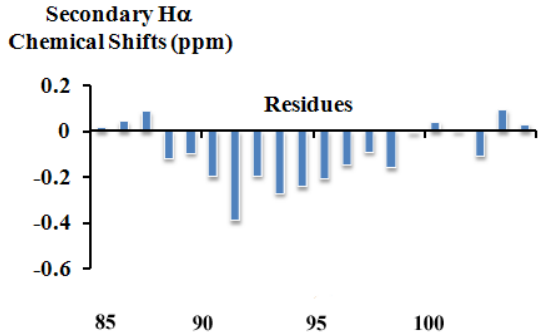


Figure 3.2.8: Secondary structure prediction using Hα chemical shifts.

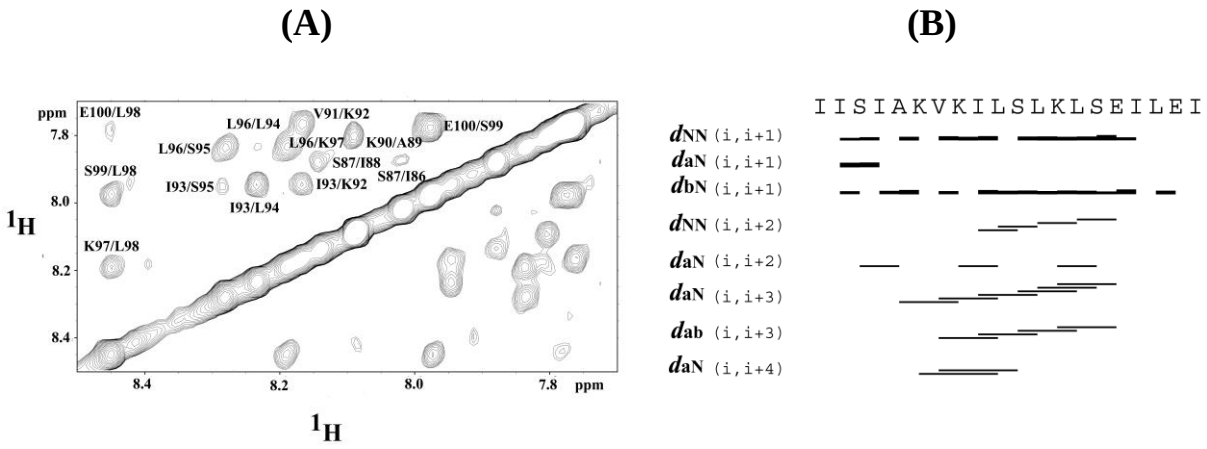


Figure 3.2.9: (A) Assignment of cross-peaks in the NOESY spectrum of $H_{85\text{-}104}$. Figure A shows NOESY cross-peaks in the HN-HN region of the spectrum. Peak picking was done in Sparky 3.110 software and cross peaks were identified based on TOCSY and NOESY spectrum. (B) The NOESY connectivity plot of peptide $H_{85\text{-}104}$ indicative of the residues connected in space revealing the presence of helical structure in the middle of the peptide (residues 90-100).

3.2.5.2 Structure calculation of $H_{85\text{-}104}$

Structure calculation was carried out by Cyana 2.1 package (GÜNTERT *et al.*, 1997). Distance restraints were obtained from the assigned 2D NOESY spectra. Pseudo atom corrections were considered when stereo specific assignments could not be made. In the helix region, the values of ϕ dihedral angles were given as 60° ± 30. In disordered regions the ϕ angles

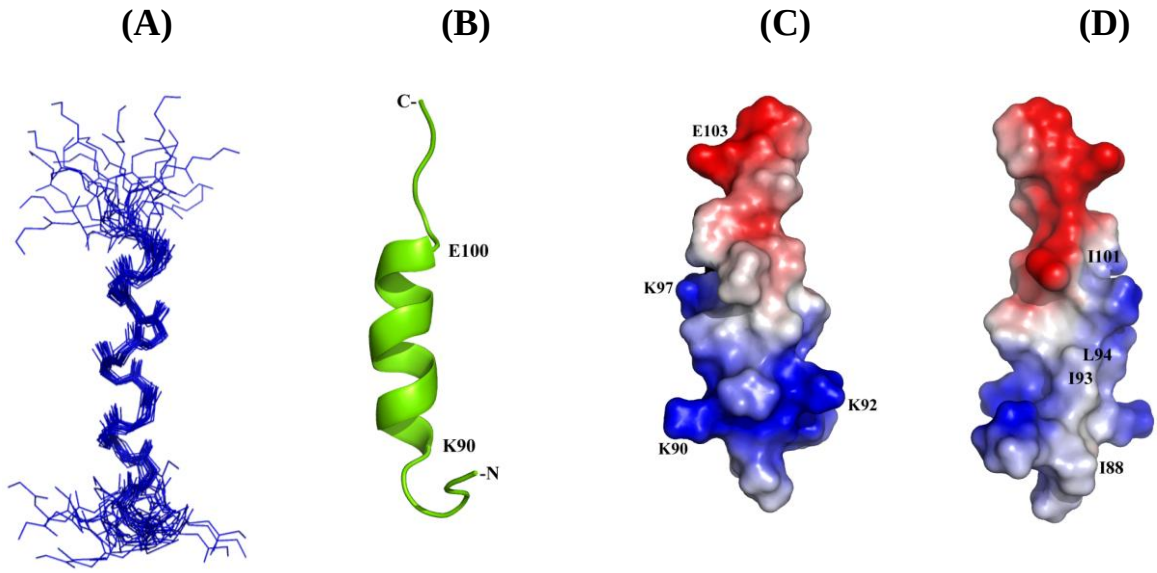


Figure 3.2.10: Superimposition of 20 lowest energy NMR structures (A) and an average structure (B) of $H_{85\text{-}104}$ calculated by the program CYANA (GÜNTERT *et al.*, 1997). The figures (C) and (D) show the molecular surface electrostatic potential of peptide $H_{85\text{-}104}$ generated by Pymol, where the positive potentials are drawn in blue and negative in red.

were constrained between -180° to -20° to maintain good stereochemistry of the calculated structures. Out of the 100 structures generated, 20 lowest energy structures were taken for further analysis. In total, an ensemble of 20 calculated structures resulted in an overall r.m.s.d. of 0.45 Å for the backbone atoms and 1.67 Å for the heavy atoms. All these structures have energies lower than -100 kcal mol^{-1} and there are no NOE violations greater than 0.3 Å and no dihedral violations greater than 5º. The summary of the statistics for 20 structures are shown in table 3.2.1. The calculated structure displays a flexible N- and C-terminal region (residues 85-89, residues 101-104 respectively) and an α-helical region (residues 90-100) as shown in figure 3.2.10A - B. Molecular surface electrostatic potential of the peptide is shown in figure 3.2.10C - D. The charge distribution of the peptide is amphiphilic, with the positive charge (K90, K92 and K97) spreads on one side while the hydrophobic surface on the other side, formed by the residues I88, I93, L94 and I101. The only negatively charged amino acid E100 is at the hydrophobic side of the helix.

Total number of residues	20
Total number of NMR restraints	164
Intraresidual (|i–j|=0)	33
Sequential (|i–j|=1)	49
Short range (|i–j|≤ 1)	82
Medium-range (2 ≤|i–j|≤ 5)	31
Long-range (|*i*–*j*| > 5)	0
Dihedral angle constraints	31
Ramachandran plot statistics (%)	
Residues in most favoured regions	85.0
Residues in additionally allowed regions	15.0
Residues in generously allowed regions	0
Residues in disallowed region	0
Structural precision for well ordered region	
r.m.s.d. backbone (residues 90–100)	0.45 Å
r.m.s.d. heavy atoms (residues 90–100)	1.67 Å

Table 3.2.1: Structural statistics for $H_{85\text{-}104}$

3.2.6 Examination of disulfide bond formation by cross linking and mutagenesis of subunit H from *M. jannaschii*

In this study, I have focused further on the C-terminus to determine the important role of specific residues to form the coiled-coil domain of subunit H. The solution structure of $H_{85\text{-}104}$ shows a helix formed by residues K90-E100. In the helix region of $H_{85\text{-}104}$, there is a triangular orientation of residues I93, L96 and L98 (Figure 3.2.11). In order to go into details, I have made three mutants of subunit H (H_{I93C}, H_{L96C} and H_{L98C}), where in each case a cysteine residue is introduced in place of I93, L96 and L98 respectively. This will nicely allow to analyze different inter subunit disulfide pattern as well as the proximity of those residues in the dimeric subunit H.

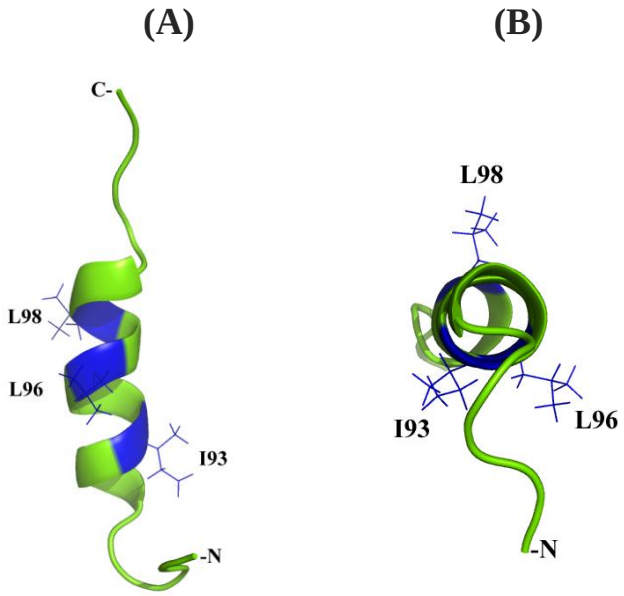
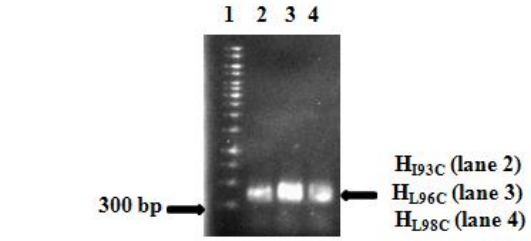


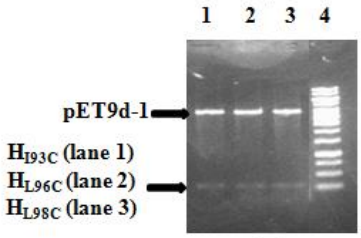
Figure 3.2.11: NMR solution structure of $H_{85\text{-}104}$ showing different orientations of three residues I93, L96 and L98 (A) side view, (B) top view. The positions of I93, L96 and L98 are shown in blue color.

3.2.6.1 Expression and purification of H_{I93C}, H_{L96C} and H_{L98C}

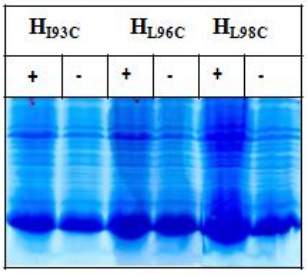
Genomic DNA of *Methanocaldococcus jannaschii* ATCC 43067D was used as template for the PCR. H_{I93C}, H_{L96C} and H_{L98C} was successfully cloned in modified pET9-d1 expression vector (GRÜBER *et al.,* 2002) as described in section 2.2.2. Expression was carried in routinely used host BL21 (DE3) *E. coli* strains. Screening tests pointed out that the protein is produced in high quantity and quality, when expressed in BL21 (DE3) *E. coli* strains at 37 °C for 4 h with 1 mM IPTG. Mutant proteins were purified as described under section 2.2.3.2. Ni^{2+}-NTA affinity chromatography was used in the first step of purification. Distinct bands of proteins were identified in the fractions 150 to 400 mM on SDS gel (Figure 3.2.12). To remove residual minor contaminations and imidazole, protein was additionally applied onto an anion-exchange column RESOURCE^{TM} Q followed by a Superdex 75 HR 10/30 column. Prior to injecting on the RESOURCE^{TM} Q column, fractions were diluted appropriately in 50 mM Tris/HCl, pH 7.5 buffer to produce a final NaCl concentration less than 100 mM. The column the flow rate was kept low (1 ml/min) to allow binding of the protein to the matrix.



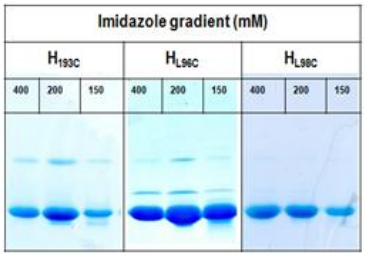
The AtpH coding region was amplified using the genomic DNA from *Methanocaldococcus jannaschii* ATCC® #43067D™.



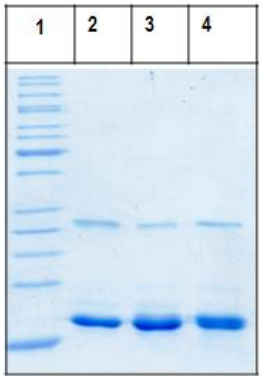
Following digestion with *Nco*I and *Sac*I, the PCR products were ligated into the pET9d-1 vector. After digestion ligated vector clearly shows the bands of insert and vector separately.



The pET9d-1 vector, containing the appropriate insert, was then transformed into *E. coli* cells (BL21 (DE3)) and grown on 30 µg/ml kanamycin-containing LB agar-plates. To induce expression of mutant proteins, the cultures were supplemented with IPTG to a final concentration of 1 mM.



Ni^{2+}-NTA elution profile of H_{I93C}, H_{L96C} and H_{L98C} shows that almost pure protein was obtained at 150 mM, 200 mM and 400 mM imidazole fractions.



SDS PAGE after ion exchange and gel filtration chromatography. Lane 1, standard; 2, H_{I93C}; 3, H_{L96C}; 4, H_{L98C}.

Figure 3.2.12: Expression and purification of subunit H mutants, H_{I93C}, H_{L96C} and H_{L98C}.

Linear gradient of buffer A (50 mM Tris/HCl, pH 7.5) and buffer B (50 mM Tris/HCl, 1 M NaCl, pH 7.5) were used for the elution of the protein. Appropriate fractions were collected and then again applied to a Superdex 75 HR 10/30 column. Fractions of constant volumes (200 µl) were collected and applied onto an SDS gel. Then, the fractions were pooled, concentrated and subsequently applied onto an SDS gel (Figure 3.2.12) to check the purity and quantity of the recombinant protein.

3.2.6.2 Dimerization interactions of subunit H mutants

To analyze the proximity of I93, L96 and L98 residue of the first and second α-helices of the dimeric H subunit, a cysteine residue is introduced in place of these residues generating the mutants H_{I93C}, H_{L96C} and H_{L98C} of the entire subunit H. Cross-link formation in the absence and in presence of $CuCl_2$ (1 µM) is shown in figure 3.2.13. In the case of mutant H_{I93C}, H_{L96C} ~15% of the protein formed a dimer. In contrast, when the residue L98, located at the opposite site of helix in $H_{85\text{-}104}$ structure (Figure 3.2.11) was exchanged by a cysteine residue, 100% of the protein formed dimer. Almost there is no difference when the mutant proteins were incubated in the absence and presence of oxidizing agent.

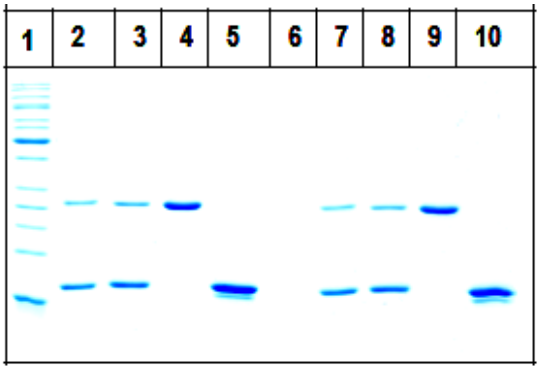


Figure 3.2.13: Cross-linking of the H_{I93C}, H_{L96C} and H_{L98C} and subunit H using $CuCl_2$. Protein were incubated with 1 µM $CuCl_2$ for 30 min at 4 °C. The reaction was stopped by the addition of EDTA. The samples were applied to a 17% total acrylamide and 0.4% cross-linked acrylamide gel. Lanes: 1, protein marker; 2, 3, 4 and 5, untreated proteins H_{I93C}, H_{L96C} and H_{L98C} and subunit H respectively; 6, blank; 7, 8, 9 and 10, treated proteins H_{I93C}, H_{L96C} and H_{L98C} and subunit H respectively.

3.2.6.3 CD spectroscopy of subunit H mutants

CD spectra of the mutant proteins (Figure 3.2.14) shows minima at Θ222 and Θ208 and maxima at Θ192, indicative of significant α-helical contents of the proteins. 2 mg/ml protein concentrations were used for the measurement of CD spectra from 180-260 nm. The details of the measurements have been described in the section 2.2.5. H_{I93C} and H_{L96C} mutants showed almost identical spectrum, suggesting the helical content and nature of helix interactions is not changed due to mutations. In contrast, there is a decrease in absorption at 208 and 222 nm for

H_{L98C} mutant in comparison with H_{I93C} and H_{L96C} mutants. The result indicate that a slight decrease in the fraction of residues in helical content for H_{L98C}.

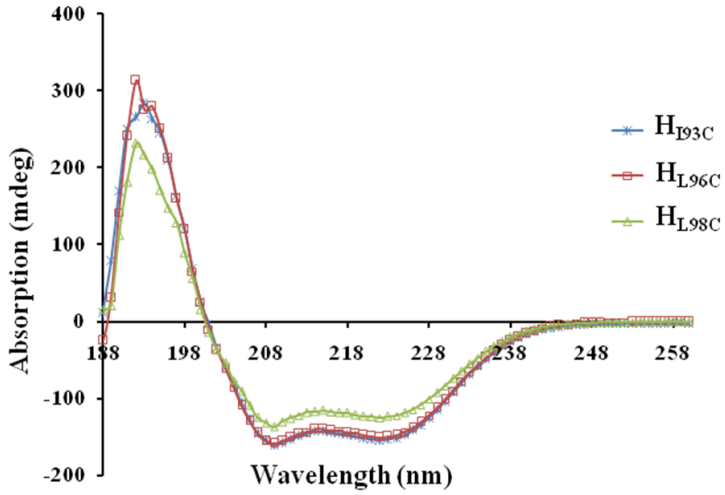


Figure 3.2.14: CD Spectroscopy of subunit H mutants, H_{I93C} (blue), H_{L96C} (brown) and H_{L98C} (green). CD spectrum of mutant proteins were measured in a buffer (50 mM Tris, 250 mM NaCl, pH 7.5) at room temperature in a Hellma quartz cuvette (60 µl volume) on Chirascan spectropolarimeter (Applied Photophysics) at a step resolution of 1 nm in the far UV range of the spectrum.

3.2.7 NMR study of the N-terminal domain of subunit H (H_{1-47})

Subunit H of A_1A_O ATP synthase from *M. jannaschii* has a boomerang like shape in solution as determined from SAXS data (BIUKOVIĆ *et al.,* 2007). It is divided into two arms 120 and 60 Å in length. Now, I focused on determining the regions of the N-terminal arm of subunit H. In this context, H_{1-47} construct was prepared by Dr. Goran Biuković, which was used for the structure determination by NMR spectroscopy. Since subunit H is proposed to be a part of the peripheral stalk of the A_1A_O ATP synthase due to its boomerang shape, this N-terminal construct H_{1-47} will enable us to determine the possible neighboring subunit(s) in the complex.

3.2.7.1 Resonance assignments of H_{1-47}

The crucial step to analyze protein by NMR spectroscopy is the process of sequential assignment. Assignments of amino acids in the N-terminal part of H_{1-47} were performed by backbone ^{15}N and ^{13}C signals using ^{15}N NOESY-HSQC as well as triple resonance backbone experiments (HNCACB, CBCA(CO)NH). The $C^{\alpha}H$, $C^{\beta}H$ proton resonances and some of the side chain methyl resonances were assigned from the ^{15}N-edited NOESY-HSQC. An example of the sequential assignment for the continuous residues (from Met_1 to Lys_{10}) is shown in figure 3.2.15. The backbone amide (^{15}N and $^{1}H_N$) resonances for 47 possible residues were completely assigned. The assignments of the resolved backbone residues of H_{1-47} are presented on a 2D ^{1}H-

^{15}N heteronuclear single quantum correlation (HSQC) spectrum (Figure 3.2.15B). Overall the limited chemical shift dispersion of only 1.5 ppm for the amide proton resonances may indicate predominantly helical conformation of H_{1-47}.

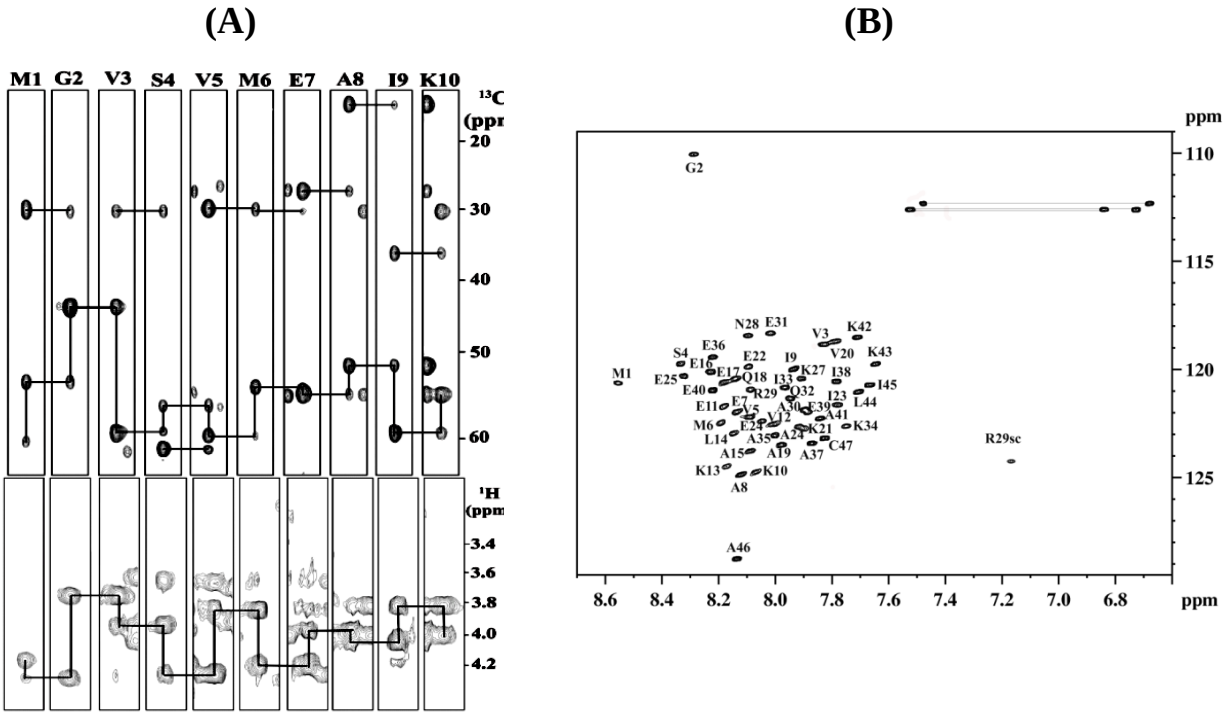


Figure 3.2.15: (A) Strip plots from the HNCACB (top) and ^{15}N NOESY-HSQC (bottom) spectra of H_{1-47}. Sequential Cα, Cβ and Hα connectivity for the 1st 10 residues M1 - K10 are shown in lines. (B) ^{1}H/^{15}N-HSQC spectrum of H_{1-47} in 25 mM sodium phosphate buffer (pH 6.5) at 288 K. Backbone and side-chain (sc) amide assignments are shown for each residue. Signals from side-chain NH_2 groups are connected by horizontal lines (GAYEN *et al.*, 2008).

3.2.7.2 Solution structure of the N-terminal domain H_{1-47}

The three dimesional NMR structure of subunit H_{1-47} was calculated based on a total of 468 unequivocally assigned NOEs-derived distance constraints including 133 intraresidual, 180 sequential and 155 medium range intermonomeric NOEs. Figure 3.2.16 shows 10 lowest energy conformers of H_{1-47} superimposed using backbone atoms of residues 15-42 and statistics is shown in table 3.2.2. The NOE patterns together with the chemical shift index (CSI) reveal a helical region in the middle of the protein. CSI consist of a single helix in middle of protein with some flexible residues (M1 to M6 and L44 to C47) at both terminus as shown in figure 3.2.17. The analysis of the ramachandran plots shows 81.5 % residues in the most favored regions, 18.5 % are in the additionally allowed regions. The regular α-helical conformation was detected for residues A15 to K42 with a small kink in the centre of the molecule. In the solution structure, H_{1-47} has a total length of about 60 Å, including the α-helical structure of 43 Å in length. Figure

3.2.18 shows a bar diagram summarizing the backbone and the backbone/sidechain NOE connectivities of H_{1-47}.

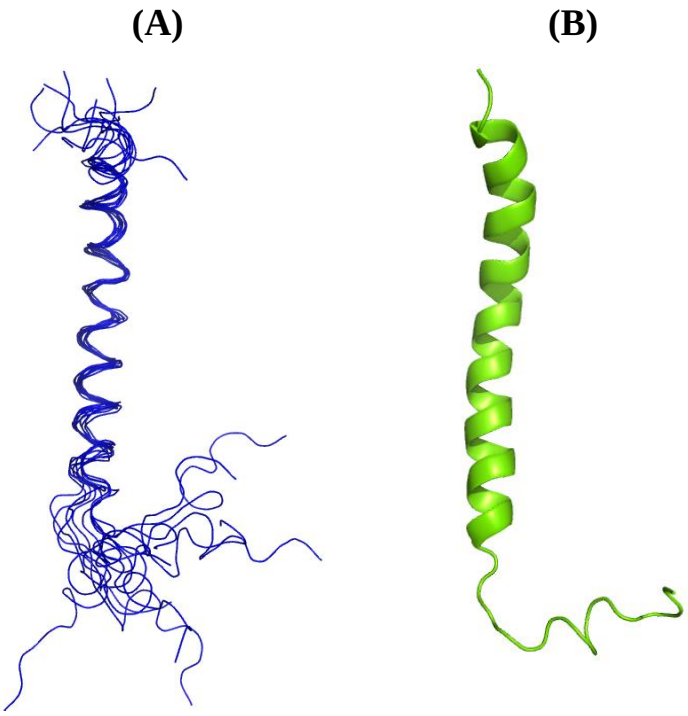


Figure 3.2.16: Ribbon diagram of the NMR solution structure of H_{1-47}. (A) best fit superimposition of the 10 lowest-energy NMR structures (B) average structure of H_{1-47} (BIUKOVIĆ *et al.*, 2009).

Total number of NMR restraints	526
Intraresidual (|i–j|=0)	133
Sequential (|i–j|=1)	180
Short range (|i–j|≤ 1)	313
Medium-range (2 ≤|i–j|≤ 5)	155
Long-range (|*i*–*j*| > 5)	0
Dihedral angle constraints	58
Ramachandran plot statistics (%)	
Residues in most favoured regions	81.5
Residues in additionally allowed regions	18.5
Residues in generously allowed regions	0
Residues in disallowed regions	0
Structural precision for well ordered region	
r.m.s.d. backbone (residues 15–42):	0.95 Å
r.m.s.d. backbone (residues 15–33):	0.68 Å
r.m.s.d. backbone (residues 33–42):	0.61 Å

Table 3.2.2: Statistics for the 10 final structural models of H_{1-47} (BIUKOVIĆ *et al.*, 2009)

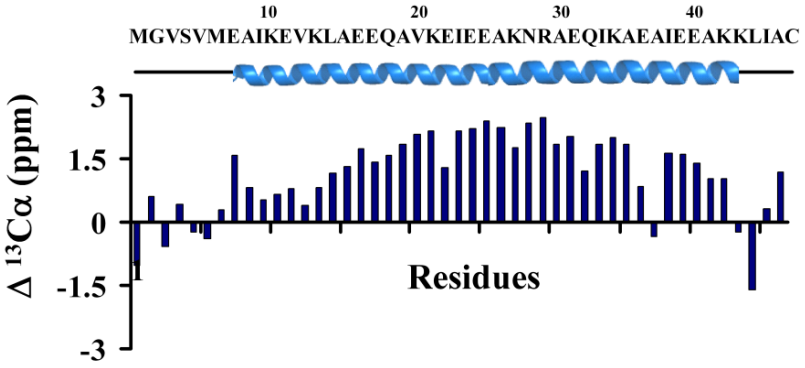


Figure 3.2.17: The amino acid sequence of H_{1-47} and the secondary structure elements, based on $^{13}C\alpha$ chemical shifts with respect to its random coil values (GAYEN *et al.*, 2008).

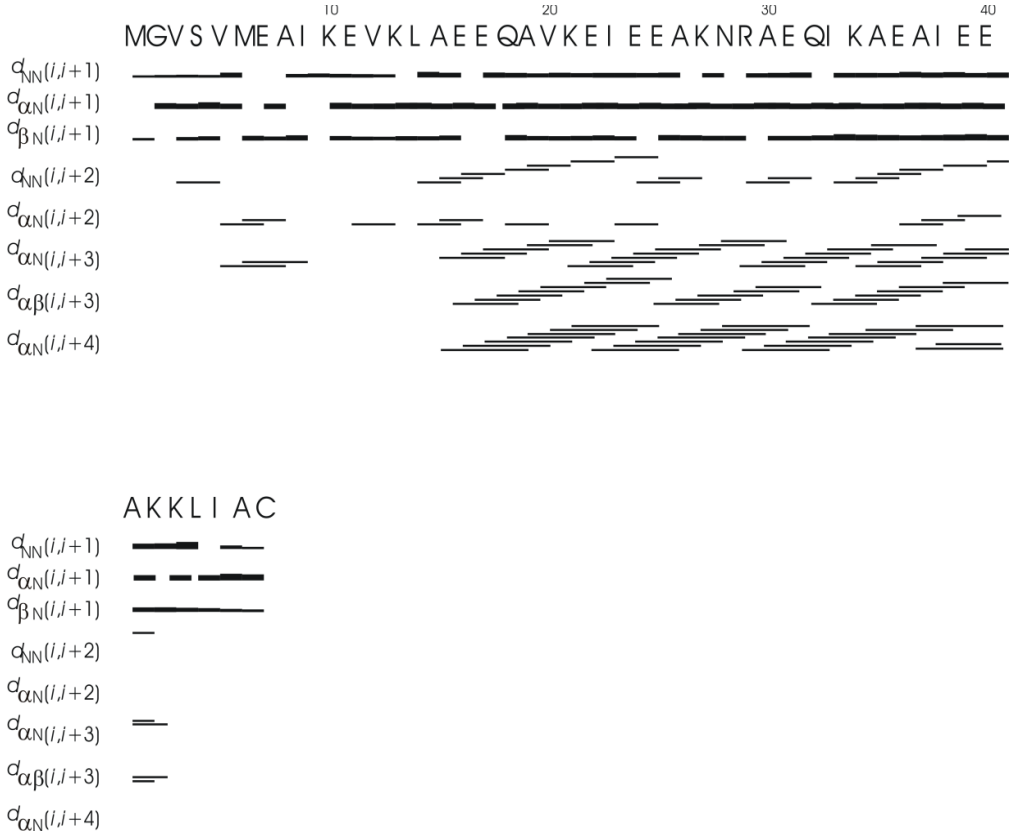


Figure 3.2.18: Summary of structurally important intra monomer NOESY connectivities of H_{1-47} (BIUKOVIĆ *et al.*, 2009).

3.2.7.3 Concentration dependence of ^{1}H, ^{15}N chemical shifts and ^{15}N relaxation parameters for H_{1-47}

The change of ^{1}H and ^{15}N chemical shifts was monitored through ^{1}H ^{15}N HSQC experiments. The spectra are taken as a function of concentration ranging from 0.125 to 3 mM.

An increase in concentration of H_{1-47} results in an overall change in the HSQC spectrum (Figure 3.2.19). The variation of the cumulative ^{1}H and ^{15}N shifts at highest and lowest studied concentrations is shown in figure 3.2.20. Since only one set of peaks is observed for all residues in the ^{1}H-^{15}N HSQC spectrum, the resonances must be average values corresponding to monomeric H_{1-47} and the oligomeric H_{1-47} in fast exchange in the chemical shift time scale. The largest shifts observed for the continuous N-terminal stretch of residues comprising M6 to A19, where structural consolidation upon dimmer formation is observed by ^{15}N relaxation data. The results showed that there are several important residues affected by the formation of dimer, *e.g.* A19, I23, E24, A26, K27, Q32, I33, K34, E36, A37 and E39.

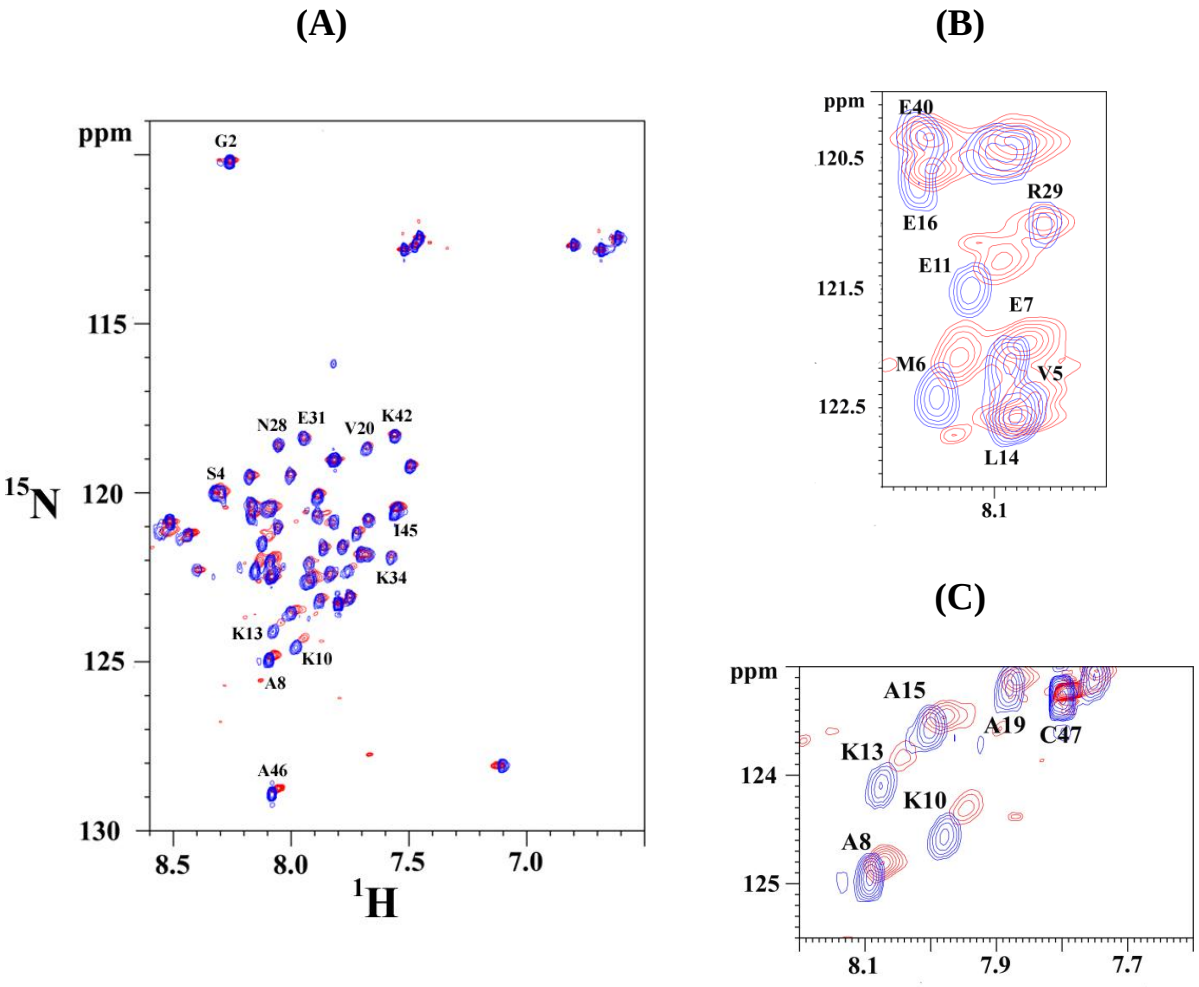


Figure 3.2.19: ^{1}H-^{15}N HSQC spectra of H_{1-47} at 10 °C in 25 mM phosphate buffer, pH 6.5 at highest (red) and lowest (blue) concentrations (BIUKOVIĆ *et al.*, 2009).

The relaxation rates in proteins are very much sensitive to overall rotational diffusion properties as well as shape and size of the protein molecule in solution. Under fast exchange conditions between monomeric and dimeric states, ^{15}N relaxations rates become powerful reporters of the

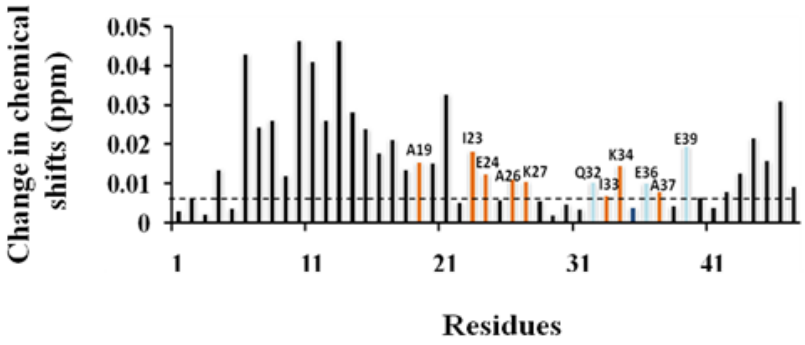


Figure 3.2.20: Combined amide (^{1}H) and nitrogen (^{15}N) chemical shift changes ($[(\Delta^{1}H_{N})^{2} + (\Delta^{15}N_{ppm}/6.51)^{2}]^{0.5}$) for $H_{1\text{-}47}$ at concentrations between 3 and 0.125 mM. Orange colored bars represent the residues forming the dimeric interface of $H_{1\text{-}47}$. Charged residues exhibiting large shifts, which are not directly situated at the dimeric interface, are marked in cyan (BIUKOVIĆ *et al.*, 2009).

average rotational diffusion properties of the contributing oligomeric states. Thus, ^{15}N R_1 and R_2 relaxation rates and ^{1}H-^{15}N NOE were measured in order to provide direct estimate of backbone mobility in monomeric as well as in the oligomeric forms of $H_{1\text{-}47}$. Generalized order parameters (S^2), internal correlation times (τ_e), and conformational exchange-induced line broadening (R_{ex}) were calculated from these data by the Lipari and Szabo model free approach. S^2 order parameters reflect spatial restriction of ^{1}H-^{15}N bonds on the ps to ns time scale, and range from 0 for unconstrained motion to 1 for completely rigid proteins. Figure 3.2.21 shows the obtained ^{15}N R_1 and R_2 relaxation rates and ^{1}H-^{15}N NOE as well as the generalized order parameters obtained in the Lipari-Szabo model free analysis. The corresponding model free statistics are listed in table 3.2.3. With the exception of the terminal residues, the S^2 values for $H_{1\text{-}47}$ at high and low concentrations are rather uniform, ranging between 0.66 and 0.85. The resulted order parameters are shown in figure 3.2.21D. Their magnitude is characteristic of well-folded proteins, demonstrating that fast internal motions are limited for residues 15 to 42. The τ_e values are in the picosecond range, as it is typically found in well-folded proteins. Within the model of the isotropic rotational

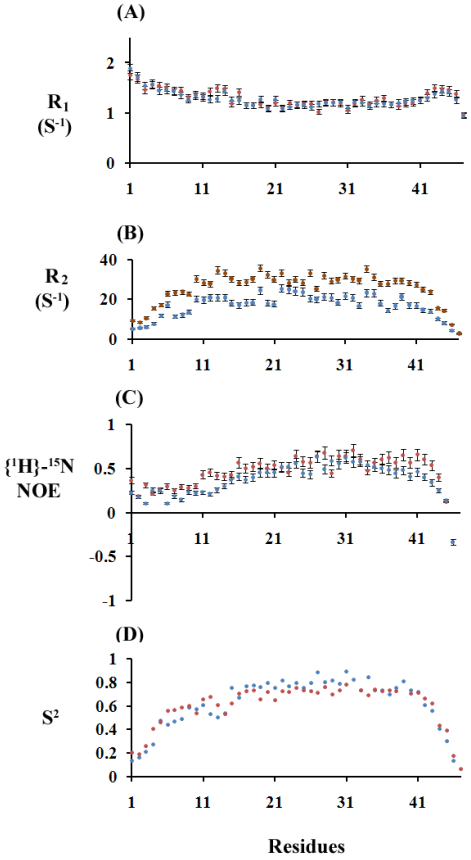


Figure 3.2.21: (A) ^{15}N R_1, (B) R_2 relaxation rates, (C) ^{1}H-^{15}N heteronuclear NOEs for the $H_{1\text{-}47}$ at 10 ºC and (D) Lipari-Szabo order parameter S^2 versus residue number at 3 mM (brown) and 0.125 mM (blue) concentration of $H_{1\text{-}47}$ (BIUKOVIĆ *et al.*, 2009).

diffusion, the rotational diffusion properties are contained in a single characteristic time constant, the rotational correlation time. Almost quantitative doubling of the rotational correlation time from 16 ns to 25 ns for the protein between high and low concentrations unequivocally establishes dimerization process. The latter is accompanied by structural consolidation at the region 4-21 as evidenced by the increase of the corresponding order parameters.

Res	S^2	S^2err	τ_e	$\tau_e$err	R_{ex}	Res	S^2	S^2err	τ_e	$\tau_e$err	R_{ex}
1	0.134	0.026	1303.000	33.406	4.177	**1**	0.169	0.024	1472.200	66.548	1.631
2	0.162	0.029	1223.900	28.266	0.000	**2**	0.155	0.021	1223.300	20.914	2.399
3	0.207	0.029	1112.100	22.296	0.000	**3**	0.216	0.027	1373.100	50.169	0.000
4	0.274	0.038	1260.300	45.214	0.000	**4**	0.343	0.034	1214.400	29.979	0.173
5	0.472	0.053	1141.100	57.016	0.003	**5**	0.397	0.039	1248.600	37.605	1.032
6	0.436	0.102	1001.900	88.053	0.000	**6**	0.496	0.047	1264.700	48.225	0.488
7	0.47	0.048	1060.200	45.824	0.337	**7**	0.504	0.045	1186.500	40.782	0.542
8	0.488	0.054	1001.500	50.128	0.004	**8**	0.526	0.048	1232.500	51.168	0.127
9	0.583	0.052	1007.000	70.816	0.227	**9**	0.532	0.042	1215.800	46.885	0.143
10	0.572	0.091	1011.500	141.310	0.000	**10**	0.601	0.053	1200.600	57.916	3.658
11	0.608	0.084	972.200	147.930	0.000	**11**	0.605	0.049	1443.200	102.780	1.097
12	0.528	0.092	1042.000	103.510	0.000	**12**	0.618	0.045	1478.300	118.440	0.212
13	0.504	0.101	1127.100	142.950	0.000	**13**	0.582	0.06	1425.500	96.525	0.000
14	0.536	0.092	1157.800	133.480	0.000	**14**	0.512	0.072	1448.200	90.142	0.000
15	0.751	0.056	926.890	223.230	0.317	**15**	0.639	0.044	1409.000	106.190	1.253
16	0.666	0.062	1184.400	167.780	1.067	**16**	0.648	0.046	1787.500	173.450	0.019
17	0.767	0.056	896.630	251.050	0.283	**17**	0.667	0.045	1540.400	150.020	0.001
18	0.772	0.053	929.980	224.850	0.236	**18**	0.683	0.043	1587.300	169.450	0.155
19	0.762	0.065	1062.500	206.240	0.000	**19**	0.688	0.046	1774.000	176.600	4.004
20	0.793	0.048	967.600	264.160	0.000	**20**	0.714	0.041	1488.300	158.610	0.586
21	0.751	0.055	1099.000	232.180	0.077	**21**	0.651	0.047	1708.100	177.380	0.635
22	0.816	0.056	1051.100	313.210	0.000	**22**	0.722	0.044	1527.900	158.620	0.844
23	0.766	0.062	1225.900	281.630	0.000	**23**	0.666	0.041	1425.800	100.070	0.000
24	0.796	0.06	1260.200	310.540	0.000	**24**	0.695	0.047	2000.000	159.520	0.119
25	0.752	0.061	1081.400	193.460	0.000	**25**	0.677	0.038	1791.600	173.290	0.063

Table 3.2.3: Model free parameters for $H_{1\text{-}47}$ at 0.125 and 3 mM concentration (BIUKOVIĆ *et al.*, 2009).

26	0.793	0.053	940.080	271.890	1.897	**26**	0.72	0.044	1718.300	195.590	0.843
27	0.882	0.041	1123.500	532.840	0.002	**27**	0.645	0.041	2000.000	140.460	1.682
28	0.798	0.054	1075.200	278.350	2.579	**28**	0.703	0.046	2000.000	156.440	1.131
29	0.817	0.052	1341.500	406.630	1.592	**29**	0.647	0.047	1430.000	108.780	0.388
30	0.786	0.052	1296.000	292.560	0.055	**30**	0.679	0.045	2000.000	146.660	0.186
31	0.891	0.042	1071.300	577.200	0.675	**31**	0.729	0.04	2000.000	167.060	0.089
32	0.82	0.052	1256.400	368.020	1.430	**32**	0.691	0.046	2000.000	161.530	1.481
33	0.735	0.056	1513.500	277.440	0.068	**33**	0.675	0.047	1927.000	160.250	0.015
34	0.843	0.046	1039.500	406.890	3.867	**34**	0.708	0.046	1444.100	135.590	2.871
35	0.735	0.067	1310.400	242.620	0.000	**35**	0.689	0.044	1745.100	178.170	0.391
36	0.734	0.054	1277.900	197.970	0.488	**36**	0.668	0.041	1920.200	151.870	0.156
37	0.696	0.051	1251.600	177.220	2.933	**37**	0.676	0.044	2000.000	140.410	0.215
38	0.753	0.055	1052.600	205.080	0.398	**38**	0.668	0.048	1773.000	184.110	0.087
39	0.805	0.052	1012.700	296.740	2.365	**39**	0.674	0.045	2000.000	147.930	0.284
40	0.729	0.054	1048.700	163.420	0.013	**40**	0.648	0.044	1787.500	172.180	0.019
41	0.72	0.058	1195.900	202.990	0.004	**41**	0.65	0.048	2000.000	151.380	0.481
42	0.609	0.056	1250.600	110.450	0.018	**42**	0.601	0.045	1951.400	140.950	0.486
43	0.56	0.055	1197.500	85.930	0.147	**43**	0.555	0.043	1739.100	152.460	0.288
44	0.407	0.047	1191.200	47.663	1.879	**44**	0.332	0.041	1488.300	88.449	0.000
45	0.299	0.04	1099.200	31.582	0.000	**45**	0.333	0.033	1114.000	17.067	2.376
46	0.134	0.023	829.590	22.395	0.000	**46**	0.148	0.019	827.740	17.820	0.000
47	-	-	-	-	-	**47**	0.052	0.008	556.380	29.833	0.000

Table 3.2.3 (continued): Model free parameters for H_{1-47} at 0.125 and 3 mM concentration (BIUKOVIĆ *et al.*, 2009).

3.2.7.4 Measurement of diffusion coefficients

The ^{15}N relaxation measurements as well as the corresponding model free analysis demonstrated two times increase of the rotational correlation time of H_{1-47} with the increase of its concentration from 0.125 to 3 mM, directly demonstrating its dimeric state at concentrations higher than 3 mM. To estimate the dimerization equilibrium constant we have performed NMR pulsed field gradient (PFG) diffusion experiments in a dilution series, revealing the equilibrium between monomeric and dimeric states of H_{1-47}. The results clearly indicate that H_{1-47} diffuses significantly faster at low concentration and the monomer-dimer equilibrium of H_{1-47} exists at a

protein concentration < 3mM. We have also determined the diffusion coefficients for the constructs, H_{1-61}, H_{1-69} and H_{1-80} at different concentrations (shown in figure 3.2.22A-D). In all of the cases, the decrease in the experimentally determined diffusion coefficient with the increasing concentration suggests the dimer formation of the proteins. The results indicate that H_{1-61}, H_{1-69} and H_{1-80} have a higher tendency to form dimer than H_{1-47}.The diffusion coefficient of H_{1-47}, H_{1-61}, H_{1-69} and H_{1-80} were measured as function of protein concentration. Solution for NMR analysis was prepared by dissolving the proteins in 50 mM Tris, 250 mM NaCl. The dimer dissociation constant for H_{1-47} was estimated by direct fit of translation diffusion velocities. H_{1-61}, H_{1-69} and H_{1-80} show a stronger tendency of dimer formation.

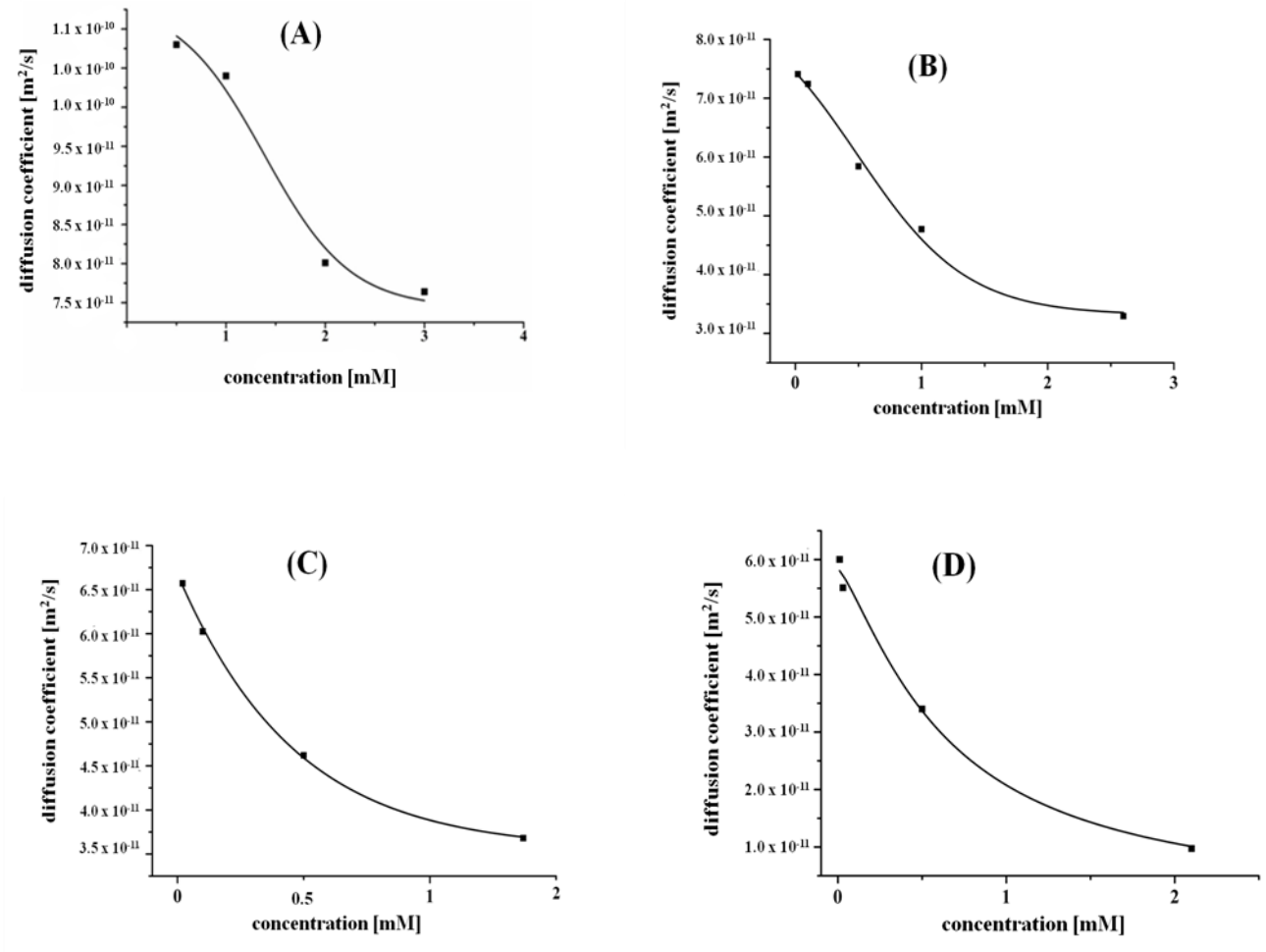
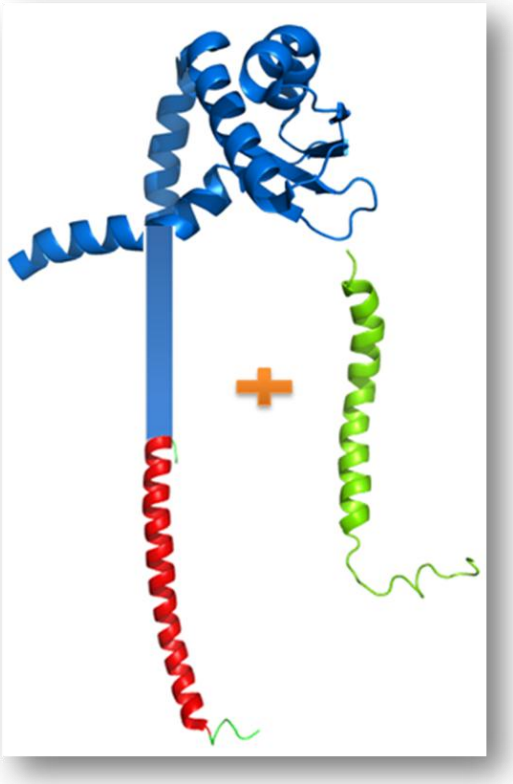


Figure 3.2.22: Change of diffusion coefficient against concentration for H_{1-47} (*A*), H_{1-61} (*B*), H_{1-69} (*C*) and H_{1-80} (*D*), respectively (BIUKOVIĆ *et al.*, 2009).

3.3 Interaction of subunit H with subunit E and the NMR solution structure of the N-terminus E_{1-52} of A_1A_o ATP synthase from *Methanocaldococcus jannaschii*



3.3.1. Introduction

Single particle analysis of the A_1A_O ATP synthase has revealed two peripheral stalks, and a collar-like structure (BERNAL *et al.*, 2004; COSKUN *et al.*, 2004; VONCK *et al.*, 2009). The most likely candidates are the subunits H, E and *a*, respectively (SCHÄFER *et al.*, 2006b; GRÜBER *et al.*, 2008). Both the peripheral stalks and the collar-like structure are involved in constituting the stator region of the enzyme, which is believed to have important roles in enzyme assembly as well as energy coupling between A_1 and A_O (GRÜBER, 2003). Subunit H has a boomerang-like shape, which is divided into two arms of 120 Å and 60 Å in length (BIUKOVIĆ *et al.*, 2007). An NMR solution structure of the 60 Å long N-terminal arm, composed of the N-terminal residues 1-47 (H_{1-47}), has been solved, revealing an α-helical feature (BIUKOVIĆ *et al.*, 2009). The crystallographic structure of the C-terminal part (residues 81 -198) of subunit E of the A-ATP synthase from *Pyrococcus horikoshii* OT3 has been reported, showing that the C-terminal domain of E consists of four antiparallel β-strands and six α-helices (LOKANATH *et al.*, 2007). The goal of this work was to analyze the interaction between H_{1-47} and the N- as well as C-terminal domain of the proposed neighboring subunit E. In the crystal structure the N-terminal domain of subunit E was not observed (LOKANATH *et al.*, 2007). In order to go into details of the N-terminal domain of subunit E, the construct E_{1-52}, containing the N-terminal 52 residues, of the A_1A_O ATP synthase from *M. jannaschii* was designed for NMR structural studies. The solution structure of E_{1-52} has been solved using nuclear magnetic resonance spectroscopy.

3.3.2. NMR titration of H_{1-47} with the N- and C-terminus of E subunit

NMR titration experiments using ^{1}H-^{15}N HSQC spectra were used to characterize the interactions between subunit H_{1-47} and the two segments, E_{1-100} and $E_{101-206}$ from *M. jannaschii*, respectively. The subunit E constructs (E_{1-100} and $E_{101-206}$) have been cloned and purified by Dr. Asha Balakrishna of our laboratory. Two sets of titrations were performed: HSQC spectra of ^{15}N labeled H_{1-47} at a fixed concentration of 100 μM, were recorded in the absence or in the presence of increasing amounts of unlabeled E_{1-100} and $E_{101-206}$ separately. Figure 3.3.1A shows the complete and figure 3.3.1B shows sections from the overlaid 2D ^{1}H-^{15}N HSQC spectrum of the truncated form H_{47} (shown in blue) and the H_{47} - E_{1-100} complex (shown in red), recorded in the presence or absence of E_{1-100}. A number of residues show significant chemical shift perturbations upon binding of ^{15}N labeled H_{1-47} to E_{1-100}. As can be seen, there is strong chemical shift change in the N-terminal region of the primary sequence of H_{47} (residues 1-6, 10-11, 15, 20 and 24). By

comparison, when equimolar amounts of the C-terminal domain of subunit E, $E_{101-206}$, was added to H_{1-47} no significant change in the spectrum could be detected (Figure 3.3.2A).

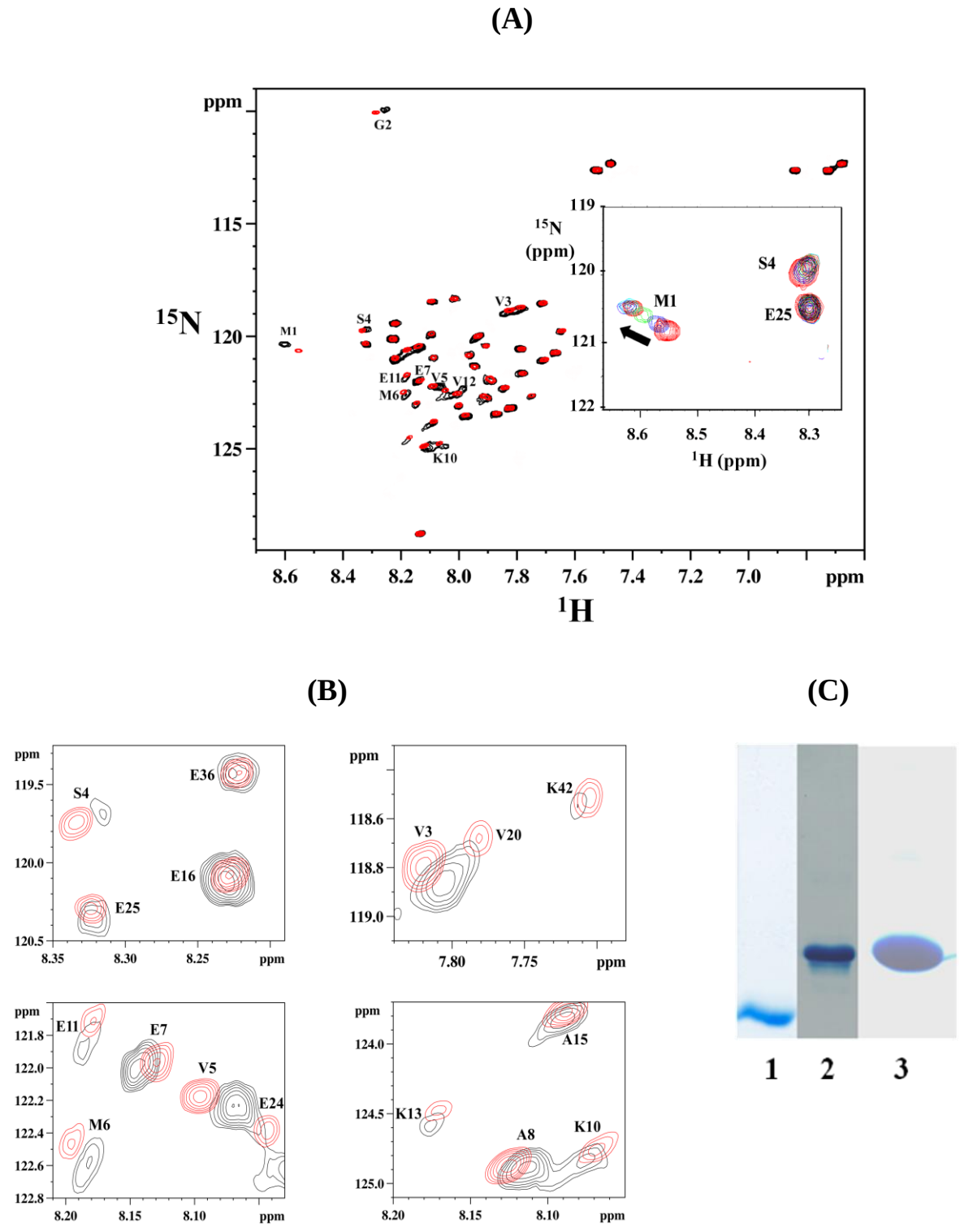


Figure 3.3.1: (A) Overlay of 2D ^{1}H–^{15}N-HSQC spectra of H_{1-47} in the absence and presence of 1.5 equivalents of unlabelled E_{1-100}. (B) Sections from the overlaid spectra of H_{1-47} in the absence (red) and presence (black) of 1.5 equivalents of unlabelled E_{1-100}. (C) Gel picture showing H_{1-47} (lane 1), E_{1-100} (lane 2) and $E_{101-206}$ (lane 3) (GAYEN *et al.*, 2008).

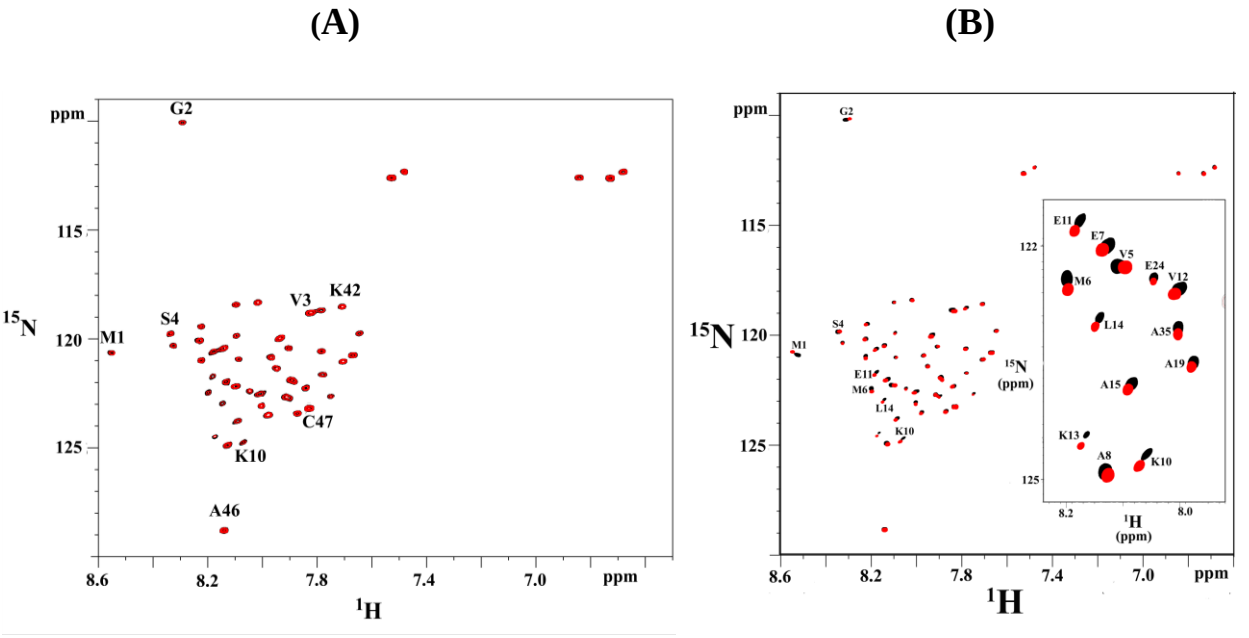


Figure 3.3.2: (A) Overlay of 2D ^{1}H-^{15}N-HSQC spectra of H_{1-47} in the absence (*red*) and presence (*black*) of 1.5 equivalent of $E_{101-206}$. (B) Overlay of 2D ^{1}H-^{15}N-HSQC spectra of H_{1-47} in the absence (*red*) and presence (*black*) of 1.5 equivalent of E_{41-60} peptide (GAYEN *et al.*, 2008).

In order to map the region of E_{1-100} involved in the interaction with H_{1-47}, the latter has been titrated with the peptides $_{1}$MKLMGVDKIK-SKILDDAKAE$_{20}$ (E_{1-20}), $_{21}$ANKIISEAEA-EKAKILEKAK$_{40}$ (E_{21-40}) and $_{41}$EEAEKRKAEI-LKKGEKEAEM$_{60}$ (E_{41-60}), which are synthesized and purified at the Division of Chemical Biology and Biotechnology, School of Biological Sciences, Nanyang Technological University, Singapore. The HSQC spectrum of ^{15}N labeled H_{1-47} in the presence of the peptide E_{41-60} shows chemical shift changes of the N-terminal amino acids (Figure 3.3.2B), similar to the ones observed in the presence of the entire E_{1-100} domain (Figure 3.3.1A). In contrast, no change of the spectra were observed in NMR-experiments in which ^{15}N labeled H_{1-47} was supplemented with the peptides E_{1-20}, E_{21-40}, respectively.

3.3.3. NMR titration of subunit F and the C-terminus $E_{101-206}$

The solution structure of subunit F of the methanogenic A_1A_O ATP synthase using NMR, exhibits a distinct two-domain structure, with a globular N-terminal having 78 residues and the residues 79-101 forming the C-terminal tail (GAYEN *et al.*, 2007). The N-terminal domain of F is proposed to be close to the bottom of the rotary D subunit (GRÜBER *et al.*, 2008), which is indicated to be in close proximity with the C-terminal part of subunit E (GRÜBER *et al.*, 2008). Based on this, we examined the ability of interaction of subunit F and $E_{101-206}$ by NMR titration

experiments. When $E_{101-206}$ was titrated against the labeled subunit F, no significant chemical shift changes could be observed, indicating that there is no interaction between both proteins.

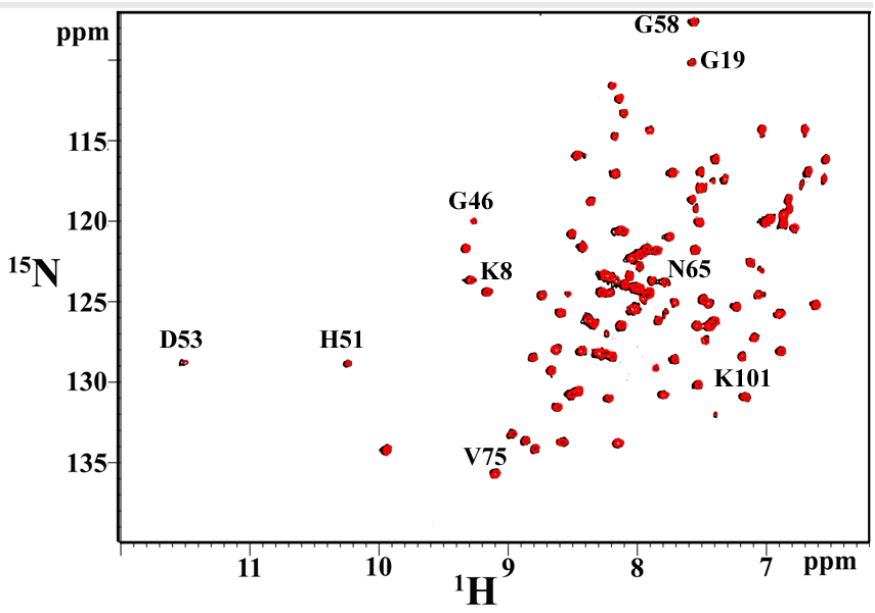


Figure 3.3.3: Overlay of 2D ^{1}H-^{15}N-HSQC spectra of F subunit in the absence (*red*) and presence (*black*) of 1.5 equivalent of unlabeled $E_{101-206}$. The spectrum was collected in a Bruker Avance 600 MHz spectrometer in 25 mM sodium phosphate buffer (pH 6.5) at 288 K (GAYEN *et al.*, 2008).

3.3.4. NMR solution structure of N-terminus of subunit E (E_{1-52}) of A_1A_O ATP synthase from *M. jannaschii*

3.3.4.1 Resonance assignment of E_{1-52}

Assignment for the backbone HN, N, C^{α}, C^{β} and C′ resonances of the E_{1-52} was achieved by a combined analysis of the triple resonance HNCACB, CBCA(CO)NH, and HNCO spectra. Figure 3.3.1 shows the good quality of the assigned HSQC spectrum indicating backbone HN and ^{15}N cross peaks of all residues of the E_{1-52}. Overall, the limited chemical shift dispersion of only ~2.0 ppm for the amide proton resonances (Figure 3.3.4) may indicate predominantly helical conformations of E_{1-52} as shown for other helical proteins (GIRVIN *et al.*, 1998; TAYLOR *et al.*, 2000).

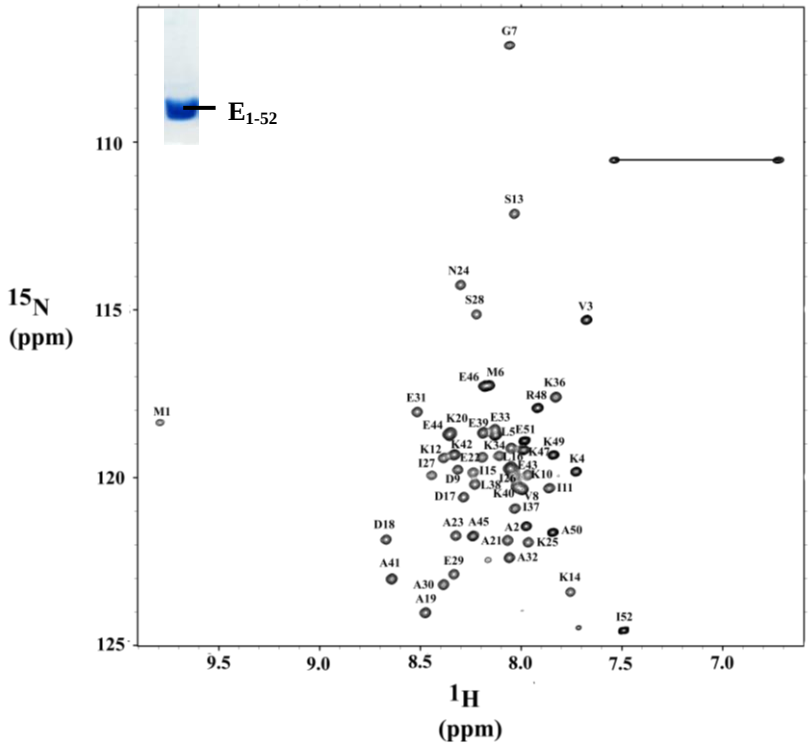


Figure 3.3.4: ^{1}H-^{15}N HSQC spectrum of subunit E_{1-52} of the A_1A_O ATP synthase from *M. jannaschii* in 25 mM sodium phosphate buffer (pH 6.5) at 308 K. Backbone and side chain (sc) amide assignment is shown as line for residue N_{24}. (Insert) SDS gel showing the sample of purified E_{1-52} (GAYEN *et al.*, 2009).

3.3.4.2 NMR derived 3-dimensional structures of E_{1-52}

A total of 545 inter-proton distance restraints (169 intraresidual, 152 sequential and 146 medium range) were used in the structure calculation of E_{1-52}, respectively, by molecular dynamics and simulated annealing. From 100 calculated structures, 20 with the lowest total energy were selected and used for structural analysis. Figure 3.3.5A represents an overlay of the 20 lowest-energy structures of E_{1-52} and the statistics are given in table 3.3.1. A bar diagram summarizing the backbone and the backbone/sidechain NOE connectivities of E_{1-52} is shown in figure 3.3.6. The 20 structures have an overall r.m.s.d. of 0.44 Å for backbone atoms and 1.10 Å for all heavy atoms in the helical regions (8-48) of the three dimensional structures of the protein. Figure 3.3.5B-D shows various representation of the solution structure of E_{1-52}. The protein adopts a regular α-helical conformation extending from residues V8 to R48. The helical conformation is not straight rather its N-terminal part (residues V8 – I26) and C-terminal part (I27 – R48) are slightly bend at opposite directions (Figure 3.3.5B-D). NH secondary chemical shifts of E_{1-52}, showing a 3-4 residue periodicity (Figure 3.3.7), support the existence of helix curvature consistent with many amphipathic and/or curved α-helices (WAGNER *et al.*, 1983; KUNTZ *et al.*, 1991; ZHOU *et al.*, 1992; NIELSEN *et al.*, 1996).

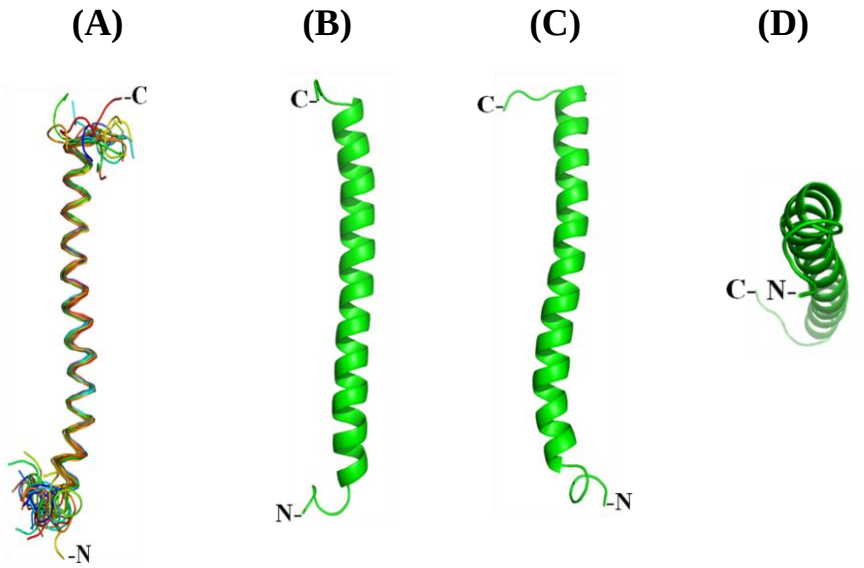


Figure 3.3.5: Ribbon diagram of the NMR structure of E_{1-52}. (*A*) Best fit superimposition of the 20 lowest-energy NMR structures. Side views (*B*), (C) and top view (D) of the average structure of E_{1-52} (GAYEN *et al.*, 2009).

The curvature of helix is caused by the differences in the H-bonding on opposite faces of the helix. Short H-bonds are correlated with low-field shifts, observed for residues K10, K14,

Total number of NMR restraints	545
Intraresidual (|i–j|=0)	169
Sequential (|i–j|=1)	152
Short range (|i–j|≤ 1)	321
Medium-range (2 ≤|i–j|≤ 5)	146
Long-range (|*i*–*j*| > 5)	0
Dihedral angle constraints	78
Total number of restraint violations > 0.3 Å	0
Ramachandran plot statistics (%)	
Residues in most favoured regions	91.4
Residues in additionally allowed regions	8.6
Residues in generously allowed regions	0
Residues in disallowed regions	0
Average RMSD to mean (Å)	
Backbone (residues 8–48)	0.44 ± 0.17 Å
Heavy atoms (residues 8–48)	1.10 ± 0.17 Å

Table 3.3.1: Statistics for the 20 final structural models of E_{1-52} (GAYEN *et al.*, 2009).

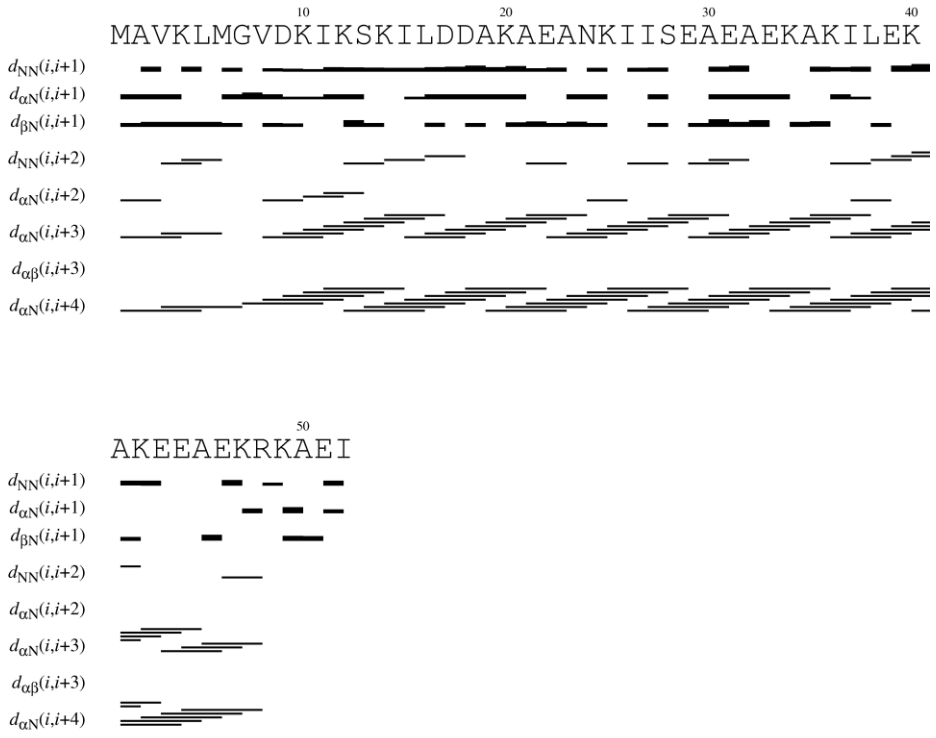


Figure 3.3.6: Summary of structurally important NOE connectivities derived from CYANA (GÜNTERT *et al.*, 1997; GAYEN *et al.*, 2009).

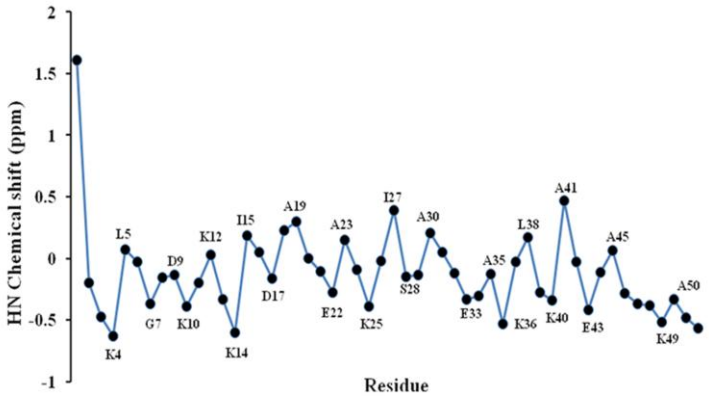


Figure 3.3.7: (A) Secondary chemical shifts for NH protons for E_{1-52} plotted against the residue sequence number for the E_{1-52} (GAYEN *et al.*, 2009).

D17, E22, K25, E29, E33, K36, K40 and K43 and long H-bonds with shifts to higher field are observed for residues D9, K2, I15, A19, A23, I27, A30, A35, L38, A41 and A45. The highest bending angles calculated from local helix axes fitted to CA atoms of ((i-3) - i) and (i - (i+3)) is in the region of K10 - S13 using HELANAL algorithm (KUMAR *et al.*, 1996). This is further supported by the observation of $d_{\alpha N(i,i+2)}$ NOE connectivities (Figure 3.3.6) near the N-terminus of the helix D9 – K12 and N24 – K25, indicating fraying or a turn at this end of the α-helix.

All the structures have energies lower than - 4 kcal·mol^{-1} calculated by GROMOS force field as implemented in SwissPDB viewer (GUEX *et al.*, 1997). There is no NOE violations

greater than 0.5 Å and no dihedral violations greater than 5°. Analysis of the ramachandran plots shows 91.4 % residues in the most favored regions, 8.6 % in the additionally allowed regions.

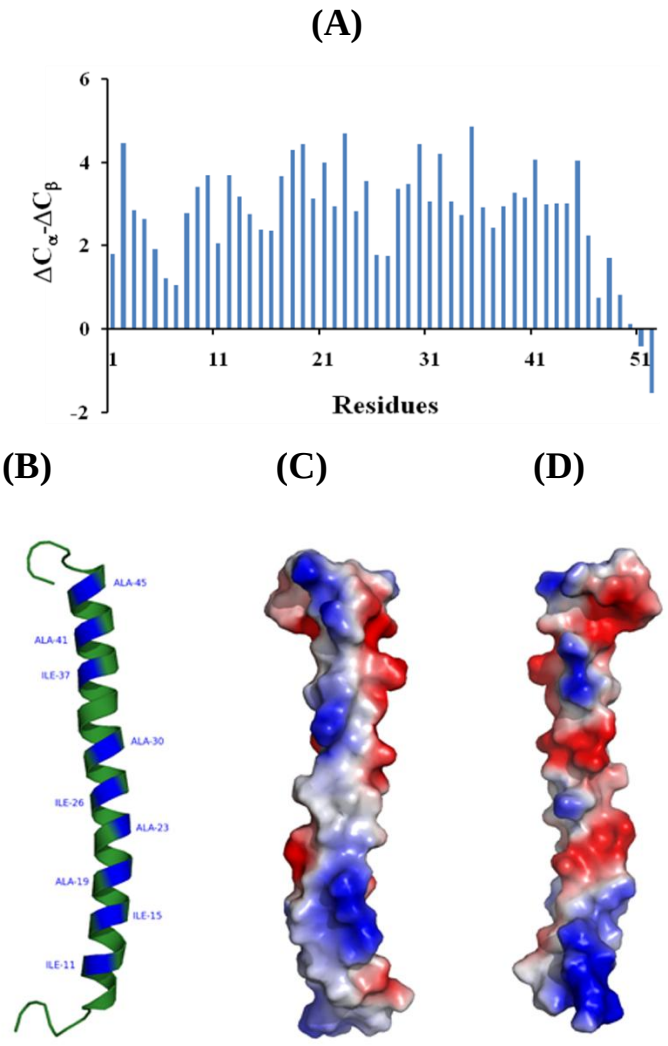


Figure 3.3.8: (A) The amino acid sequence of E_{1-52} and the secondary structure elements based on $^{13}C^{\alpha}$-$^{13}C^{\beta}$ (B) Ribbon representation of E_{1-52} showing the hydrophobic amino acids colored in blue. (C, D) Surface representation of E_{1-52}. Red and blue represent negatively and positively charged areas, respectively, calculated with the program PyMOL (http//www.pymol.org) (GAYEN *et al.*, 2009).

In the solution structure, E_{1-52} has a total length of 60.5 Å. $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts depend on local backbone geometry. Their deviations from random coil values, termed secondary shifts, provide sensitive probes to detect the secondary structural propensities. The combined carbon ($\Delta^{13}C^{\alpha}$–$\Delta^{13}C^{\beta}$) secondary shifts of E_{1-52} as a function of the protein sequence (Figure 3.3.8A) show a continuous stretch of positive combined carbon secondary shifts for almost all the residues. The middle of the protein (residues 8-48) appears to have higher positive values demonstrating the presence of a helical structure in the region. The helical domain of E_{1-52}

presents the structural characteristics of an amphipathic helix (Figure 3.3.8C - D) with the hydrophobic residues I11, I15, A19, A23, I26, A30, I37, A41 and A45 facing one side and the hydrophilic amino acids D9, K10, K14, D17, K20, K25, E29, E31, K36, E39, E43 and K47 making the second face of the helix.

4. Discussion

4.1 Solution structure of subunit F from *M. mazei* Gö1

High resolution crystal structures of catalytic and nucleotide binding subunits A (MAEGAWA *et al.,* 2006) and B (SCHÄFER *et al.,* 2006a) respectively, of the A-ATP synthase provides information on the nucleotide binding events in the A_1 part. However, the dual function of ATP synthesis or ATP hydrolysis in the A_1 domain and the ion transport in the A_O domain is combined with a complicated coupling process. The central stalk subunits C, D and F play the leading roles in the coupling event. These subunits undergo conformational changes and/or rearrangements during enzyme function. Thus, in order to understand the molecular mechanism by which these subunits play their biological role, it is very much important to know their structure at atomic level. The first low resolution shape of subunit F from the archaea *Methanosarcina mazei* Gö1 in solution determined by small angle X-ray scattering has been published in 2006 (SCHÄFER *et al.,* 2006b). The shape of the protein is elongated (shown in figure 4.1.1) (7.5 nm) divided in two parts: a globular part with a length of about 4.5 nm and a hook like domain of about 3.0 nm in length (SCHÄFER *et al.,* 2006b). The structure of subunit F determined by NMR spectroscopy has similar dimensions (7.8 nm) and shape. It also reveals two well defined domains (GAYEN *et al.,* 2007). The N-terminal region forms the globular part and the C-terminus forms the hook like part. The N-terminal globular domain is made up by an alternating arrangement of four β-strands (β1-4) and three α-helices (α1-3). The four β-strands are oriented parallel to each other. The surface representation of subunit F indicates that the four-stranded β-sheet in the N-terminal part forms a hydrophobic surface (Figure 4.1.2). As subunit F of the A-ATP synthase is in close contact with rotating subunit D (COSKUN *et al.,* 2004), the hydrophobic surface might mediate the interaction of subunits F and D. Consequently, the positively and negatively charged surface of the bottom of the N-terminal domain of subunit F would be oriented to the central stalk subunit C and thereby toward the membrane side. Such positioning of subunit F relative to subunit D and subunit C allows the

Figure 4.1.1: Structural model of the non-catalytic B subunit and the coupling subunit F. The atomic model of subunit B from the *M. mazei* Gö1 A_1A_O ATP synthase derived from the X-ray coordinates determined recently (SCHÄFER *et al.,* 2006a) (PDB 2c61). The first low resolution shape of subunit F of A_1A_O ATP synthase from *M. mazei* Gö1 was determined by solution X-ray scattering (SCHÄFER *et al.,* 2006b). The C-terminal region interacting with subunit F is labeled by sticks and balls (SCHÄFER *et al.,* 2006b).

positively charged area in the C-terminus of subunit F to interact with the negatively charged residues of the C-terminus of subunit B. Recently, the interaction of subunit F with the subunit B was proved by intrinsic tryptophan fluorescence experiment indicating a binding constant (K_d) of 332 nM, assuming that subunit B is interacting with subunit F in a 1:1 molecular ratio. Since, both subunits have one native tryptophan residue at the C-terminus, W430 and W100 for subunits B and F, respectively, the tryptophan of subunit F (W100) was substituted with tyrosine and isoleucine residue in order to measure the binding characteristic with subunit B. The fluorescence intensity dropped by 19% upon addition of mutant FW100Y with subunit B, indicating interaction between W430 of subunit B and the Y100 of mutant FW100Y. In comparison, a fluorescence quenching of only 5% was observed in case of mutant FW100I (GAYEN *et al.,* 2007). These results pointed out a π-π interaction of adenine rings of tryptophans at the C-terminus of both subunits.

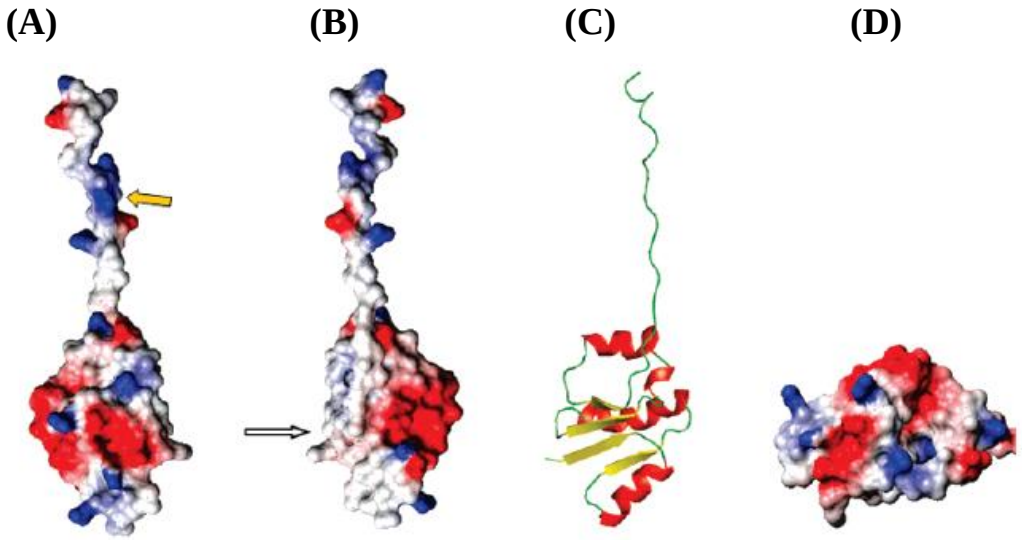


Figure 4.1.2: Charge distributions on the surfaces of subunit F of the methanogenic A_1A_O ATP synthase. Red and blue areas are negatively and positively charged areas, respectively, calculated with the program MOLMOL (KORADI *et al.,* 1996). At the left (A) the arrow shows the positively charged tail, which forms a cross-link with the C-terminus of subunit B (GAYEN *et al.,* 2007). The hydrophic area, which is proposed to interact with the central subunit D, is emphasized by a white arrow (B). A ribbon representation of the F subunit in an orientation similar to that of the surface structure is shown in (C). (D) Surface representation of the N-terminal domain from the membrane side (GAYEN *et al.,* 2007).

The fact that the C-terminus of subunit F is flexible in comparison to the N-terminal domain was revealed by the direct measurements of backbone dynamics including R_1, R_2 and two-dimensional ^1H-^{15}N heteronuclear NOE experiments. The N-terminal domain (residues 1-78) showed uniform NOE, R_1 and R_2 values indicating that N-terminal domain form a rigid body. On the other hand C-terminus (residues 79-101) exhibited completely different and more

flexible dynamical properties as can be seen in smaller NOE, longer R_1 and shorter R_2. This property makes subunit F to undergo rearrangements during energy coupling as they are described for the proposed homologue ε of the *E. coli* F_1F_O ATP synthase (TSUMURAYA *et al.*, 2009). Subunits F and ε have the same overall appearance and domain structure (SCHÄFER *et al.*, 2006b). However, subunit F of *M. mazei* Gö1 shows a sequence identity of 9% to subunit ε of *E. coli* F_1F_O ATP synthase and the secondary and tertiary structures of subunit ε are also different from subunit F. Figure 4.1.3 shows structures of subunit ε from *E. coli* F_1F_O ATP synthase and subunit F. The N-terminal domain of subunit ε forms a β sandwich with six stranded sheets (β1-6) and the C-terminal domain formed by two helices (α1-2) arranged in a anti parallel coiled-coil structure (UHLIN *et al.*, 1997). Studies on isolated subunit ε and in complex with the truncated γ subunit revealed, that the C-terminal domain of subunit ε is also flexible and can undergo the rearrangement between two different conformations – the closed / hairpin-folded and extended states (WILKENS *et al.*, 1995; RODGERS *et al.*, 2000). The conformations of this subunit could reversibly change between two states depending on the ATP concentration (YAGI *et al.*, 2007). The C-terminus of subunit ε extends along the axial helices of subunit γ in the extended state and inhibits the conformational change in subunit β and thereby inhibits the clockwise rotation of subunit γ (YAGI *et al.*, 2007). As a result, both ATP synthesis and hydrolysis are suppressed. On the other hand, when the C-terminus fold back to the hairpin-folded state by an arm like motion, subunit ε does not have any direct interaction with the $\alpha_3\beta_3$ ring (YAGI *et al.*, 2007) and thereby has no effect on the coupling process.

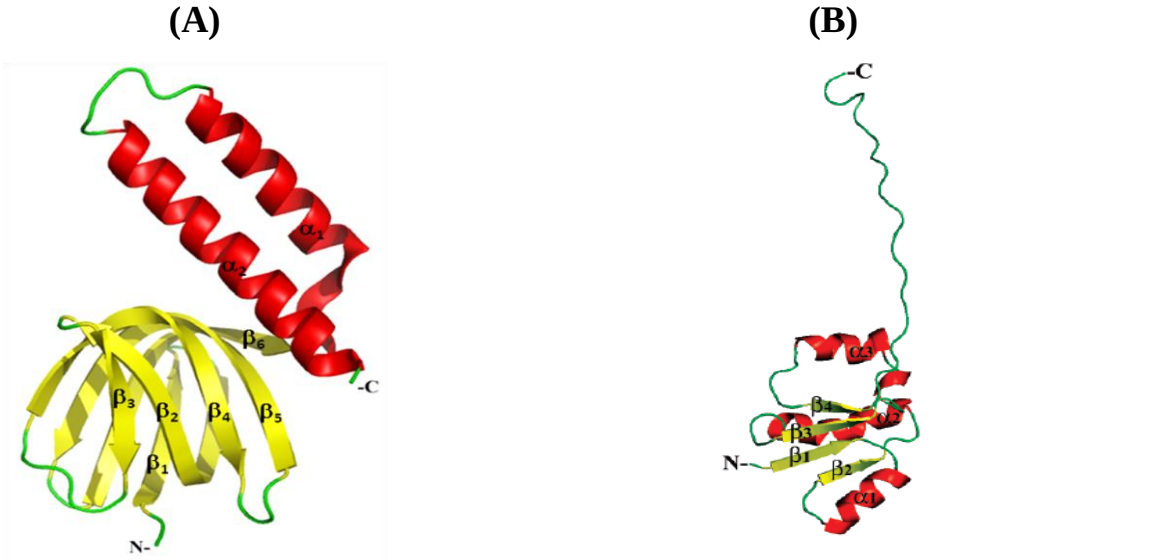


Figure 4.1.3: Ribbon representation of subunit ε of the F_1F_O ATP synthase from *E. coli* (A) (PDB code 1AQT; (UHLIN *et al.*, 1997)) and subunit F of the A_1A_O ATP synthase from *M. mazei* Gö1 (B) (PDB code 2OV6; (GAYEN *et al.*, 2007).

In a similar way, the structure of F_{81-101} displays a closed form compared to the extended form in the whole subunit F. The structure revealed that there is a short flexible N-terminal tail of the amino acids G81 – T85 and a helical structural motif formed by the residues S86 to A94. To investigate the stability of this closed form of the C-terminal end of subunit F and of the elongated form determined previously, 100 ns molecular dynamics (MD) simulation was performed on both structures using the continuum representation of bulk solvent water (RAGHUNATHAN *et al.,* 2009). The results reveal that there is an almost two-fold higher mobility among the residues 86-94 when elongated than when closed and helical. In addition, the solution structure of F_{81-101} reveals two cation-π interactions, one is between conserved R88 and W100 and the second is between K92 and W100. The distance between the pairs R88-W100 and K92-W100 was monitored during the simulations to check for the presence of cation-π interactions. The results demonstrate that the separation between R88-W100 and K92-W100 varies between 5-15 Å for the retracted and elongated states. Together, these observations highlight the intrinsic flexibility of the C-terminus, which in turn enables inter-conversions between the elongated and the closed states with a periodicity of ~30 ns (RAGHUNATHAN *et al.,* 2009). The flexible C-terminal tail enables the subunit F to undergo up and down movements relative to its neighboring nucleotide binding subunit B and functions as a coupling element by connecting the nucleotide-binding subunit with the ion-translocating part. The fact that the conformational change in the C-terminal domains of these proteins is not surprising, since similar kind of conformational change was also observed for Ras-related nuclear (Ran)-GTPase proteins (LEE *et al.*, 2005; HAO *et al.,* 2008) during nuclear protein import (Figure 4.1.4). Its C-terminal extension, consisting of an unstructured linker and a 16 residue α-helix, contacts the core of the protein in Ran-GDP state and is pointing away from the core in Ran-GTP state, thereby playing a crucial role of its binding to the transport factors. Taken together, these facts suggest that this kind of conformational change may be a general form of rarrangements in the family of hydrolases and or synthases.

Based on these observations, subunit F in its elongated form was tried to fit into the A_1 complex. Figure 4.1.5 shows a nice fitting of structure of subunit F inside the complex, and suggests a possible mechanism of the coupling function in the elongated state of subunit F. It is reported previously that subunit F formed a cross-linked product with subunit B inside the methanogenic A_3B_3DF subcomplex (COSKUN *et al.,* 2002), A_1 ATPase (COSKUN *et al.,* 2004), and A_1A_O ATP synthase (SCHÄFER *et al.,* 2006b), depending on whether MgATP or the nonhydrolyzable ATP analogue MgAMP-PNP is bound to the enzyme. The peptides involved in

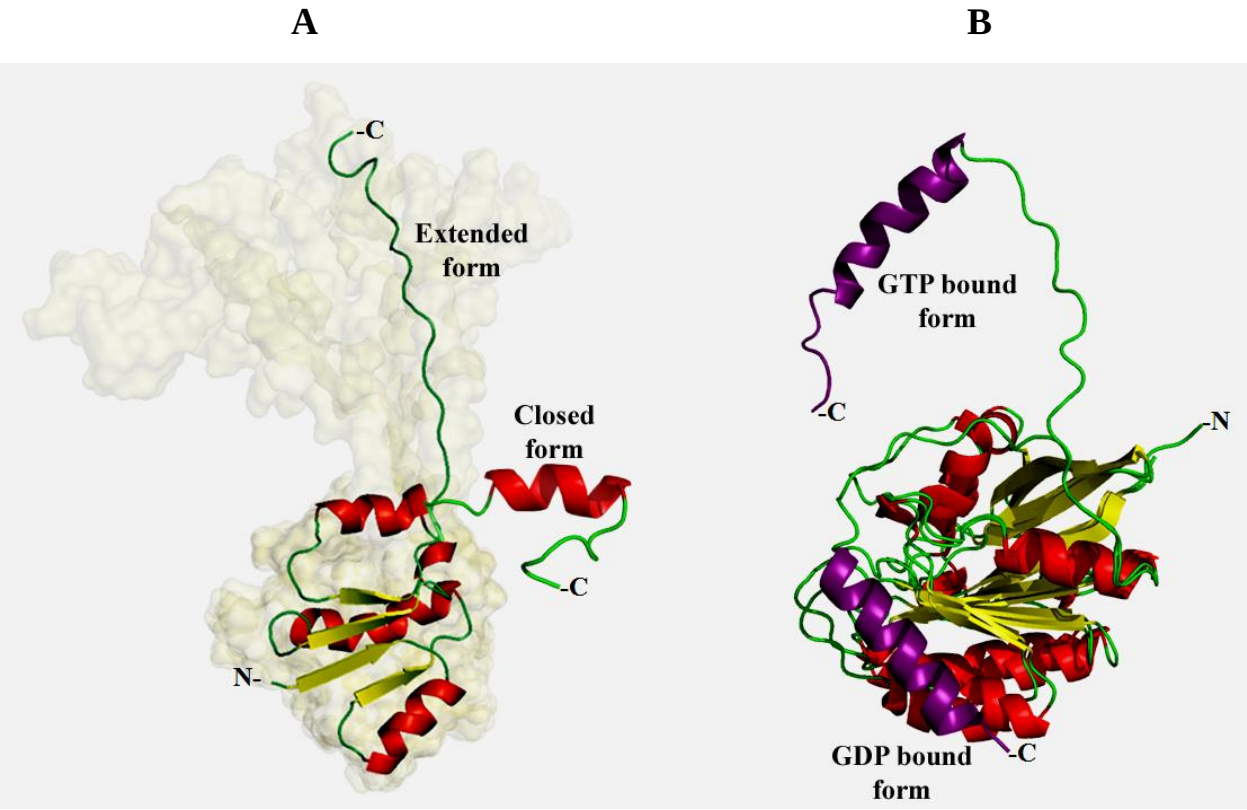


Figure 4.1.4: (A) A model of extended form and closed form of the C-terminus of the NMR structure of subunit F from *M. mazei* Gö1 A-ATP synthase inside an envelope, representing the conformational mobility of the C-terminus of F, as derived from backbone dynamics and two-dimensional ^{1}H-^{15}N heteronuclear NOE experiments. (B) The two structures of Ran proteins from two different organisms showing the conformational change of the C-terminal helix: the first one is the structure of a Q69L mutant of Ran from dog with a bound GDP molecule (PDB id 1byu, (STEWART *et al.*, 1998); the second structure corresponds to Ran from human in complex with human RanBP2 and a non-hydrolysable GTP analogue (PDB id 1rrp, (VETTER *et al.*, 1999).

this EDC (1-ethyl-3-[(dimethylamino) propyl]carbodiimide) induced zero-length cross-link of B and F are $_{390}$EALSERDTK$_{399}$ and $_{88}$REKIK$_{92}$ respectively (SCHÄFER *et al.*, 2006b). Peptide $_{390}$EALSERDTK$_{399}$ is a sequence homolog of the so-called C-terminal DELSEED region of nucleotide-binding subunit α of the F-ATP synthases (GIBBONS *et al.*, 2000) which undergoes a rotational movement of 6° after binding of nucleotides (AGGELER *et al.*, 1998; MENZ *et al.*, 2001; FENIOUK *et al.*, ; VENKATARAMAN *et al.*, 2006; MANIMEKALAI *et al.*, 2009). Therefore, the close proximity of the sequences $_{88}$REKIK$_{92}$ of the C-terminus of subunit F and the C-terminal residues of B (Figure 4.1.5) might induce a conformational change in the lower C-terminal region of B with the subsequent closing of the binding pocket for the nucleotide and thereby may inhibit ATP synthesis/hydrolysis activities. It has been shown that the B-F interaction is absent in the presence of MgADP + Pi or MgADP, due to a relative movement of subunit F (SCHÄFER *et al.*, 2006b).

In summary, the first three-dimensional structure of the central stalk subunit F of the A-ATP synthase in solution provides the structural basis toward a fuller understanding of the mechanistical properties of this subunit inside this class of enzymes. Moreover, it gives more insight into the structural and thereby functional homologies and diversities of the coupling subunits ε and F of F- and A-ATP synthases respectively.

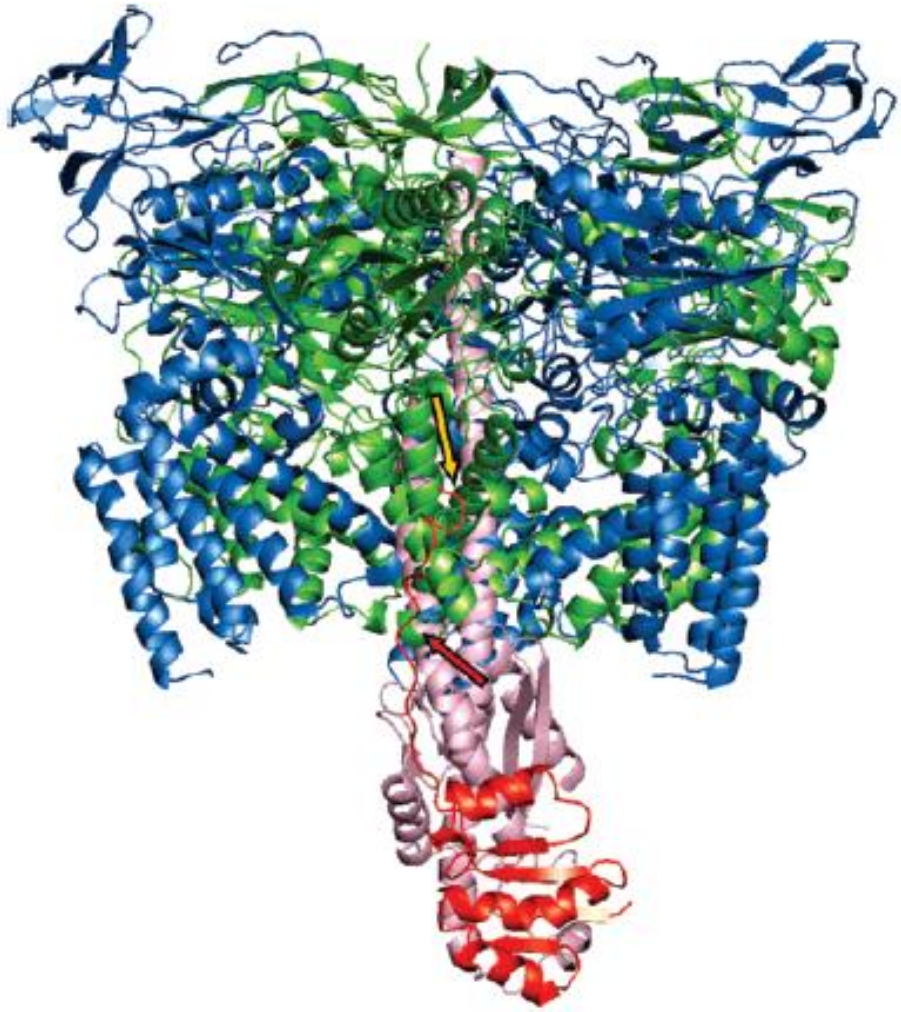


Figure 4.1.5: Proposed binding of subunit F inside the A_3B_3DF complex of the A_1 ATPase. *Pyrococcus horikoshii* A-ATP synthase subunit A (pink, PDB code 1vdz; (MAEGAWA *et al.*, 2006)) and *M. mazei* Gö1 A-ATP synthase subunit B (green, PDB code 2c61; (SCHÄFER *et al.*, 2006a)), as well as the bovine mitochondrial F-ATP synthase *c* subunit (violet, PDB code 1e1q, chain G; (GIBBONS *et al.*, 2000)), which is homologous to subunit D of the A-ATP synthase, were fitted into the electron density map of the A_1 ATPase (COSKUN *et al.*, 2004) and A_1A_O ATP synthase (BERNAL *et al.*, 2004), obtained from single-particle analysis electron micrographs (SCHÄFER *et al.*, 2006b). The solution structure of subunit F (GAYEN *et al.*, 2007) of the *M. mazei* Go¨1 A-ATP synthase is highlighted in red. The yellow and red arrows indicate the intrinsic Trp residue (Trp430) of subunit B and the contact region of subunits B and F, respectively. The figure was prepared using PyMOL (http://www.pymol.org) (GAYEN *et al.*, 2007).

4.2 Structural characterization of subunit H of A_1A_O ATP synthase from *M. jannaschii*

The peripheral stalk subunit H of the A-ATP synthase predicted from its amino acid sequence is α-helical and has a high propensity to form a coiled-coil structure (BIUKOVIĆ *et al.,* 2007). The CD spectrum of subunit H is consistent with the predicted high helical content of subunit H (78%) and also with a high degree of coiling based on the $\Theta_{222}/\Theta_{208}$ ratio of 0.96 (BIUKOVIĆ *et al.,* 2007). Classically, the dimerization interface of a two stranded coiled-coil is made up by a periodicity of nonpolar core residues and solvent exposed polar residues (Cooper *et al.,* 1990), as reflected in the amino acid sequence of subunit H. Physical characterization of this subunit by small angle X-ray scattering supports unequivocally that subunit H is dimeric in solution (25 ± 2 kDa) and its N terminal part (hook like structure) (BIUKOVIĆ *et al.,* 2007) is rigid. This is believed to be involved in interaction with central stalk subunits. From recent studies using chemical cross linker it has been shown that subunit H is in close contact via the C-terminal peptide I74 to K80 with the N-terminal region T106 to R122 of the catalytic subunit A, depending on nucleotide-binding to the catalytic site (SCHÄFER *et al.,* 2006b). Therefore subunit H can be oriented inside the 2D projection with its C-terminus located close to the external surface of the A_1 headpiece (Figure 4.2.1). This peripheral stalk shows a clear connection to the collar and/or the central stalk. The collar consists of at least the subunit E, which is in direct neighborhood to the rotary subunit D via its peptides $_{127}$LDEAAKK$_{134}$ and $_{119}$AYSSKESEELVK$_{130}$ (SCHÄFER *et al.,* 2006b). The nucleotide dependent cross-link formation of the C-terminal region of subunit H and its close proximity via its N-terminus to the central stalk region leads to the speculation whether subunit H might act as a stator and/or as a coupling or regulatory domain in the A-ATP synthase.

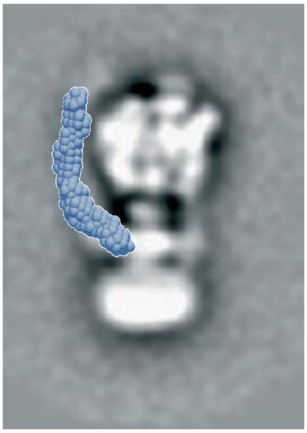


Figure 4.2.1: Composite structure of the methanogenic A_1A_O ATP synthase. The SAXS structure of subunit H (blue) was introduced by eye into the 2D projection of the enzyme complex (COSKUN *et al.,* 2004).

Domain features of subunit H were studied by using a series of deletions and disulfide formations between introduced cysteines (BIUKOVIĆ *et al.,* 2009). The results provide a clearer picture of the architecture of this subunit and helped in establishing that subunit H is separated into two domains. First, the C-terminal part of H, which consists of residues 54–104, is highly α-helical, and the two α-helices in this domain are arranged mostly in a coiled-coil fashion. This is reflected by the high $\Theta_{222}/\Theta_{208}$ ratio of the N-terminal deleted version of subunit H, $H_{47\text{-}105}$, which is higher compared to the C-terminal truncated constructs $H_{1\text{-}67}$, $H_{1\text{-}71}$, and $H_{1\text{-}78}$

(BIUKOVIĆ *et al.,* 2009). It is also in line with the reduced α-helical content of H_{1-91}, H_{1-78}, H_{1-71} and H_{1-67}, respectively, which corresponds with the deletion size of the C-terminal segment. The existence of subunit H as a pair of parallel helices with a coiled-coil arrangement in region 54–104 is also supported by data showing that cysteine mutants introduced at positions 54, 61, 69, and 80 of the truncated proteins had a strong tendency to form disulfides (BIUKOVIĆ *et al.,* 2009). This is further supported by the NMR solution structure of a short length peptide of the very C-terminus of subunit H (residues 84-103) and the cross-linking study with the mutants H_{I92C}, H_{L95C} and H_{L97C}. The results showed that a complete disulfide product between two cysteins at positions 97 and 97' of the first and second helices were generated, whereas a lower dimer product were seen for the cysteines at positions 92 and 95. These results clearly tell that residue 97 is a part of the helix-helix interface and the residues 92 and 95 are in the opposite side of the helix, which is in accordance to the solution structure of H_{85-104}. Thus, the C-terminus of subunit H is very important for the stabilization of the structure of entire subunit H, which is in line to our previous results, where truncation of six amino acids at the C-terminus of subunit H resulted twisting of the upper arm in the low resolution shape determined from SAXS analysis (BIUKOVIĆ *et al.,* 2007).

The second structural domain found is formed by the N-terminal segment (residues 1-47), which is confirmed by the NMR solution structure of H_{1-47}. The $\Theta_{222}/\Theta_{208}$ value of H_{1-47} is 0.6 supporting the predicted absence of coiled-coil assembly in the N-terminal domain (BIUKOVIĆ *et al.,* 2009). This is further supported by the higher dissociation constant of H_{1-47} compared to the constants of the longer constructs H_{1-61}, H_{1-69} and H_{1-80}, respectively. In order to understand further how the N-terminal domain of subunit H is formed, I solved the NMR solution structure of H_{1-47}. The structure of H_{1-47} exhibits an α-helix formed between amino acids 15–42. Fitting the structure over residues 15-33 (r.m.s.d. 0.68) and 33-42 (r.m.s.d. 0.61) shows, that each region forms an α-helix and the 10 structures agree well with each other over these stretches. By comparison, fitting over the length of the protein, including residue 33, gives less agreement (r.m.s.d. 0.95) because of a small kink in the centre of the molecule. In the solution structure, H_{1-47} has a total length of 61.9 Å, including the α-helical structure of 42.8 Å in length. The NMR structure of H_{1-47} reveals a strip of hydrophobic alanine residues made by A19, A26 and A37 with a slight twist to each other (Figure 4.2.2). The proximity of alanine residues of the first and second α-helix of the dimeric subunit H, A19, A26 and A37, respectively, have been analyzed recently in our laboratory by replacing alanine residue with a cysteine residue, generating the subunit H mutants H_{A19C}, H_{A26C} and H_{A37C} (BIUKOVIĆ *et al.,* 2009). When the mutants were

incubated in the absence of an oxidizing reagent a small amount of dimer formation were observed for mutant H_{A19C}, which was decreased in the case of the mutant H_{A26C} to H_{A37C}. In the case of mutant H_{A19C} about 50% of the protein formed a dimer, when the protein was supplemented with the oxidizing agent. By comparison, the amount of cross-link products was dropped when the mutant protein H_{A26C} or H_{A37C} have been used. In contrast, when the alanine residue 35, located at the opposite site of A37, was exchanged by a cysteine, no disulfide formation could be detected under the conditions used. These results indicated a closer proximity of residues 19 and 19' of both helices and supports the small twist the three alanine residues (19, 26, 37) have relative to each other in the α-helical structure (Figure 4.2.2). By contrast, the absence of any dimeric product of the H_{A35C} mutant is in line with the opposite location of this amino acid relative to the alanine-strip shown in the structure, and implies that this residue is located at the outside of the helix-helix assembly. Taken together, the low coiled-coil content seen in the CD spectra as well as the low dimeric formation indicate that the major part of the two N-terminal segments (H_{1-47} and H_{1-47}') might run in a more adjacent arrangement to each other and coming closer at the end of H_{1-47} and H_{1-47}'.

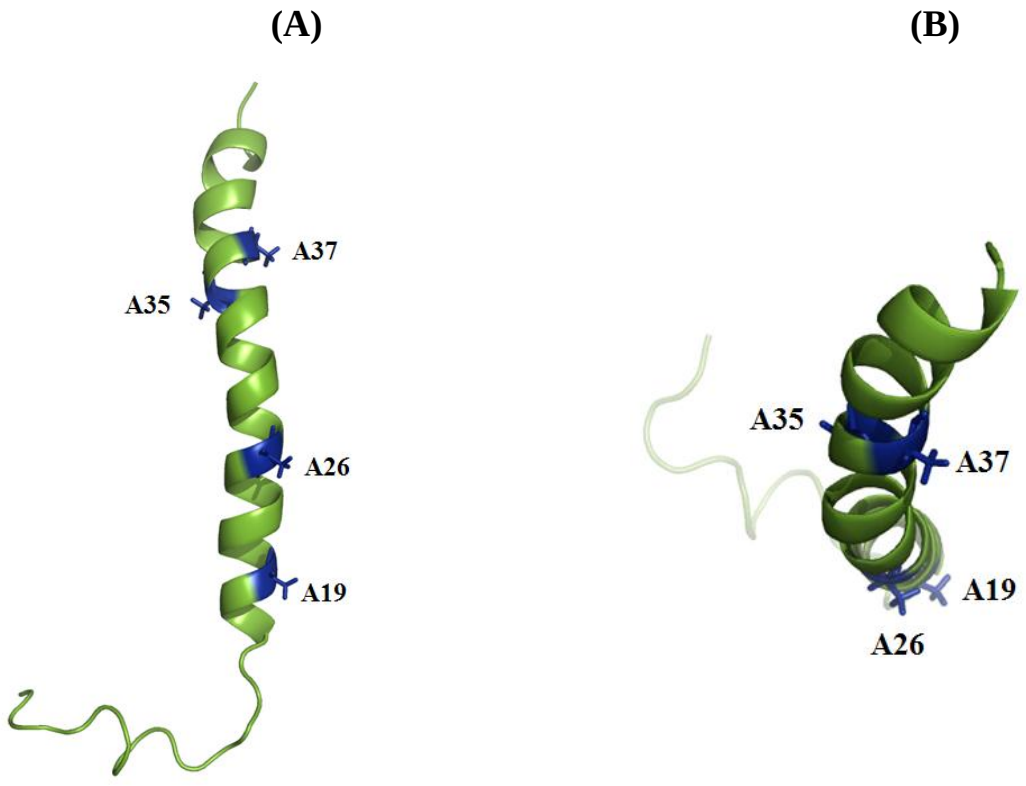


Figure 4.2.2: Ribbon diagram of the NMR solution structure of H_{1-47}. Side (A) and top (B) view of the average structure of H_{1-47} illustrating the relative positions of alanine residues 19, 26 35 and 37 inside the α-helix, which are marked in blue color (BIUKOVIĆ *et al.*, 2009).

This feature is consistent with the disulfide product seen in H_{1-47} which became stronger in the seven residue longer construct H_{1-54} (BIUKOVIĆ *et al.,* 2009). The structural switch from a more adjacent into a coiled-coil assembly of both helices inside the dimeric H subunit becomes facilitated by the small kink at residue 33. This kink is characterized by a break in the chemical shift index (CSI) pattern for the helix at this position and reflects also part of the bend in the low resolution solution structure of the boomerang-like shape of subunit H (Figure 4.2.3). The overall length (61.9 Å) and structure of H_{1-47} is remarkably similar with the shorter N-terminal arm of the boomerang shape of H as shown in figure 4.2.3. The shape volume enables both helices, H_{1-47} and H_{1-47} to fit in (BIUKOVIĆ *et al.,* 2009).

The oligomerization state of H_{1-47} has also been investigated based on overall correlation time from NMR relaxation analysis and on translational diffusion coefficients at different concentrations. The ^{15}N relaxation measurements as well as the corresponding model free analysis demonstrated two times increase of the rotational correlation time of H_{1-47} with the increase of its concentration from 0.125 mM to 3 mM, directly demonstrating its dimeric state at concentrations higher than 2 mM. NMR pulsed field gradient (PFG) diffusion experiments in a dilution series, reveal the equilibrium between monomeric and dimeric states of H_{1-47}. The results clearly indicate that H_{1-47} diffuses significantly faster at low concentration and the monomer-dimer equilibrium of H_{1-47} exists at a protein concentration < 3mM. Stimulated by the data of H_{1-47}, the diffusion coefficients for the constructs, H_{1-61}, H_{1-69} and H_{1-80} at different concentrations were also measured. In all of the cases, the decrease in the experimentally determined diffusion coefficient with the increasing concentration suggests the dimer formation of the proteins. The results indicate that H_{1-61}, H_{1-69} and H_{1-80} have a higher tendency to form dimer than H_{1-47}.

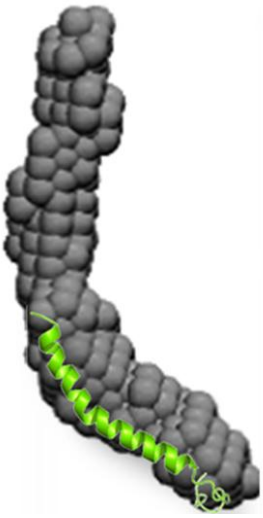


Figure 4.2.3: The H_{1-47} structure is superimposed on the low resolution structure of the dimeric H subunit derived from solution SAXS data (BIUKOVIĆ *et al.,* 2007).

Based on the cross-link data as well as the change in chemical shift data, I modeled the dimeric structure for H_{1-47}, shown in figure 4.2.4. NMR measurements with H_{1-47} as function of concentration point out that H_{1-47} forms a dimer at concentrations above ~2 mM. The ^{1}H and ^{15}N chemical shift perturbations indicate that the residues A19, I23, E24, A26, K27, I33,

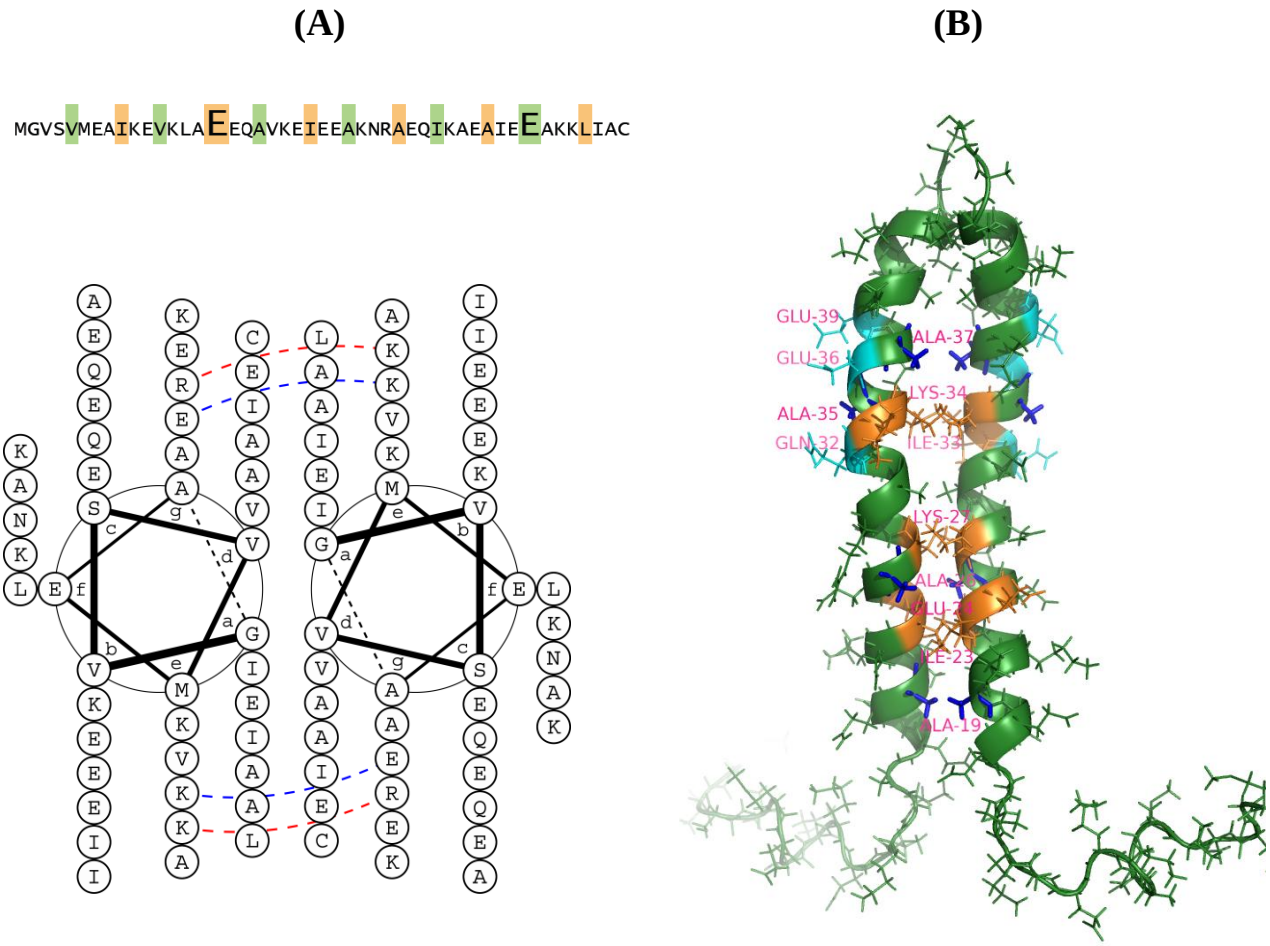


Figure 4.2.4: (A) Location of the interacting residues at the proposed dimeric interface of $H_{1\text{-}47}$ in a helical wheel representation, which was prepared using the program Drawcoil (http://www.gevorggrigoryan.com/drawcoil). residues in 'a' and 'd' positions are marked in beige and green color, respectively. Destabilizing residues Glu16, Glu40 for the dimeric $H_{1\text{-}47}$ are highlighted by larger fonts (BIUKOVIĆ *et al.*, 2009). (B) Dimeric model of $H_{1\text{-}47}$ based on cross linking data and NMR data. Residues in orange color exhibiting significant shifts represent the residues forming the dimeric interface of $H_{1\text{-}47}$. Charged residues exhibiting large shifts which are not directly situated at the dimeric interface are marked in cyan color.

K34 and A37 might be involved in dimer formation. Based on our structural model the charged residues Q32, E36 and E39 exhibiting comparably large shifts, might not be directly situated at the putative dimeric interface but rather their side-chain pK_a might be affected via proxy charged residues from the other monomer. The maximum chemical shift change has been observed for the residues 4 - 21 and residues 43 - 47. NMR secondary chemical shifts as well as NMR relaxation data have been shown that the protein has a flexible N- (residues 1 - 15) and a C-terminal region (residues 43 - 47). These chemical shifts are explained by structural consolidation at these stretches of amino acids as the direct consequence of dimerization. This follows from increased order parameters for the backbone HN vectors derived from ^{15}N relaxation data. In the relaxation data, the decreased values of R_1 as well as increased values of

R_2 suggest the dimer formation of H_{1-47} at higher concentrations ~3 mM. Extracted order parameters S^2 reflect spatial restriction of ^1H-^{15}N bonds on the ps to ns time scale, and range from 0 for unconstrained motion to 1 for completely rigid proteins. With the exception of the terminal residues, the S^2 values for H_{1-47} at high and low concentrations are rather uniform, ranging between 0.66 and 0.85. Their magnitude is characteristic of well-folded proteins, demonstrating that fast internal motions are limited for residues 15 to 42. The τ_e values are also in the picosecond range, as is typical for a well-folded protein. Almost qauntitative doubling of the rotational correlation time from 16 ns to 25 ns for the protein between high and low concentrations unequivocally establishes dimerization process. The latter is accompanying by structure accrition at the region 4 - 21 as evident from an increase of the corresponding order parameters.

4.3 Interaction of subunit H with subunit E and the NMR solution structure of the N-terminus E_{1-52} of A_1A_O ATP synthase from *M. jannaschii*

Comparison of the shape of H and the C-terminal truncated form H_{1-98}, derived from SAXS data, allowed the assignment of subunit H in the position of the peripheral stalk in the 2D projection of the A-ATP synthase (BIUKOVIĆ *et al.*, 2007). This assignment was further confirmed by intrinsic tryptophan fluorescence experiment with subunits H and A (BIUKOVIĆ *et al.*, 2007). The sequence of subunit H lacks any aromatic amino acid and subunit A contains nine native tryptophan; seven of them lie buried inside subunit A, whereby residues W118 and W192 are located on the outside of the molecule. W118 is inside the N-terminal peptide involved in the A-H cross-link formation (SCHÄFER *et al.*, 2006b). The fluorescence intensity dropped by 25% and 20% after addition of subunit H and the N-terminal truncated H_{8-104} protein, respectively. In contrast, in the presence of the C-terminal truncated H_{1-98}, no significant quenching of the fluorescent signal could be detected. These results showed that the C-terminus of subunit H is interacting with the subunit A with a binding constant of 206 nM (BIUKOVIĆ *et al.*, 2007). Connected via its C-terminal arm to the catalytic A subunit, H exceeds the total distance of the A_1 headpiece and the central stalk (GRÜBER *et al.*, 2001; COSKUN *et al.*, 2004) and becomes oriented with its N-terminal arm close to the collar-like structure of the enzyme complex, predicted to be formed by subunit E (Figure 4.3.2) (SCHÄFER *et al.*, 2006b; BIUKOVIĆ *et al.*, 2007). Recently, an E-H complex has been described, using electrophoresis and mass spectrometry (LOKANATH *et al.*, 2007; KISH-TRIER *et al.*, 2009). In the NMR titration experiment presented, it is shown that it is the N-terminal domain (E_{1-100}) of subunit E

which specifically binds to the very N-terminus of H_{1-47}. The high α-helical degree of E_{1-100} determined (71%) might indicate that the amino acid region E_{41-60} ($_{41}$EEAEKRKAEILKKGEKEAEM$_{60}$) of the E_{1-100} domain joins with the N-terminal residues 1-6, 10-11, 15, 20 and 24 of H_{1-47} (Figure 4.3.1) via a helix-helix interaction. The H_{1-47}/E_{41-60}-binding was also proved by fluorescence correlation spectroscopy (FCS) (GAYEN *et al.*, 2008). A binding constant of 3.0 μM of bound Atto488-E_{41-60} to H_{1-47} was calculated. In contrast, no binding could be observed with the C-terminal form, $E_{101-206.}$ As observed by circular dichroism spectroscopy, Trier and Wilkens (KISH-TRIER *et al.*, 2009) discussed that truncated forms of subunit E and H, E_{2-83} and H_{2-91}, from *Thermoplasma acidophilum* A_1A_O ATP synthase form a coiled-coil interaction, reflected by molar ellipticity values at 208 nm and at 222 nm in a ratio close to 0.99. The solved NMR structure of E_{1-52} with its hydrophobic strip in the α helical segment between residues 8-48 forms a perfect surface for an hydrophobic and helix-helix interaction between subunit E and H and the periodicity of Ala and/or Ile residues may indicate a coiled-coil interaction of both proteins. There is biochemical evidence that subunit F of the A-ATP synthase is in close contact to subunit D (COSKUN *et al.*, 2002; GAYEN *et al.*, 2007). As demonstrated by the solution structure, the four-stranded β-sheet in the N-terminal part of subunit F forms a hydrophobic surface, which is suggested to mediate the interaction of both subunits (GAYEN *et al.*, 2007). Such positioning of subunit D relative to F might bring the latter in close proximity to the C-terminal domain of subunit E. However, the data presented show no obvious interaction of F and $E_{101-206,}$ indicating that subunit E is mainly interacting with subunit H and D via its N-and C-terminal part, respectively.

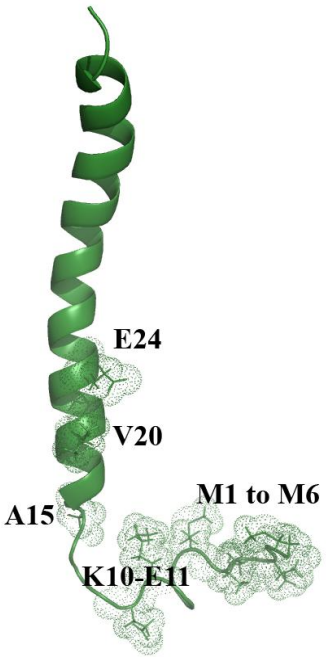


Figure 4.3.1: The structural model of H_{1-47} enables to assign the residues involved in subunit E-H formation. Previously the residues M_1 to M_6, K_{10} to E_{11}, A_{15}, V_{20} and E_{24} of $H_{1-47,}$ marked in stick representation encircled with dotted spheres, have been shown to interact with the N-terminal tail E_{41-60} of subunit E (BIUKOVIĆ *et al.*, 2009).

In summary, the presented data support the view that the α-helical subunit H forms one of the two peripheral stalks of the enzyme, with its C-terminus connected to the N-terminal part of the catalytic A subunit and its N-terminus in close contact to the N-terminus of subunit E, with the latter being in close connection via its C-terminus to subunit D (Figure 4.3.2). The nucleotide

dependent cross-link formation of A-H, its close proximity via its N-terminus to subunit E, and the D-E neighborhood leads to the speculation whether both subunit H and E might be involved in coupling and/or regulatory events in the A-ATP synthase.

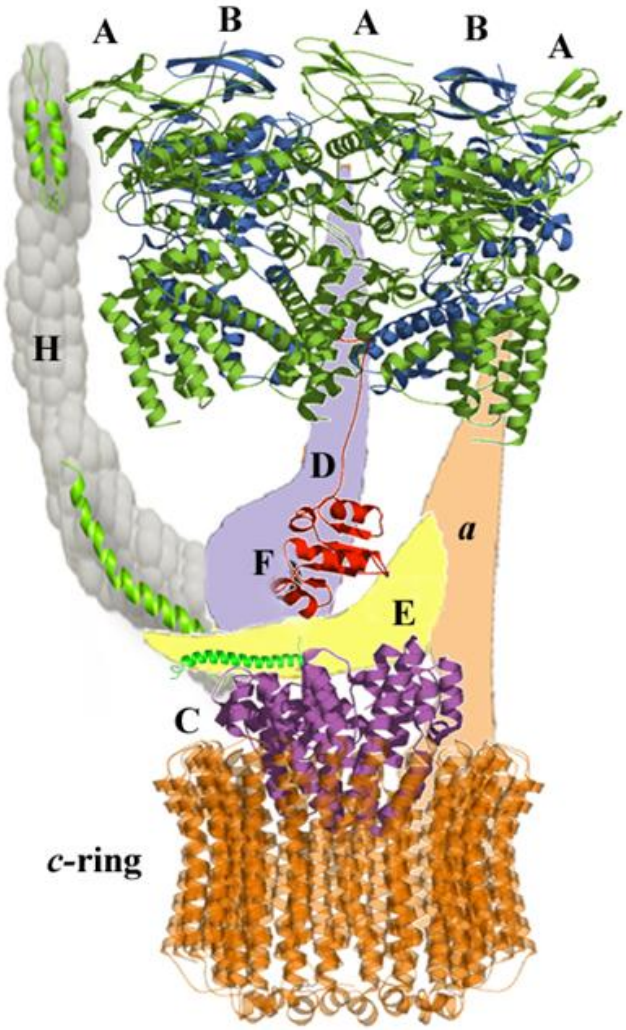


Figure 4.3.2: The arrangemtns of the subunits in the model A_1A_O ATP synthase. One B subunit has been removed from the A_1 section to reveal the D subunit within the A_3B_3 hexamer. P. horikoshii A-ATP synthase subunit A (green, pdb 1vdz, *M. mazei* Gö1 A-ATP synthase subunit B (blue, pdb 2c61), and *M. mazei* Gö1 A-ATP synthase subunit F (olive red, pdb 20V6) were fitted into the model of the A_1 ATPase and A_1A_O ATP synthase, obtained from single particle analysis electron micrographs. The figure was prepared using PyMOL (http://www.pymol.org). (B) The A subunit is attached to subunit N-terminal domain by a peripheral stalk composed of subunit H, whereby the C-terminal part of E is in close proximity to the coupling subunit D.

5. Conclusion

The A_1A_O adenosine triphosphate (ATP) synthase is a fascinating enzyme in cellular bioenergetics responsible for the production of life molecule ATP from ADP and phosphate in archaea. A_1A_O is a complex of two rotary motors A_1 and A_O and the ATP synthesis/hydrolysis reaction that is reversibly catalyzed by A_1 is coupled with proton transport across membrane-embedded A_O. Subunit F is an important subunit in the central stalk of the enzyme. The conformational change in the stalk region is important for both the catalytic events (ATP synthesis or hydrolysis) as well as ion translocation. The solution structure of subunit F is separated into a well structured N-terminal domain (78 residues) and an extended C-terminal tail (22 residue). The N-terminal domain is composed of a four-stranded parallel β-sheet structure and three α-helices and has a dimension of 24.2 Å x 28.5 Å x38.7 Å. The length of the C-terminal tail is 40 Å. The entire subunit F tumbles in two different entities as the two domains are loose in contact. This is confirmed by NMR dynamics studies. To get a deeper insight, I have determined also the solution structure of the peptide from the C-terminus (residues 81-101). The solution structures together with the molecular dynamics study explain the possible role of the C-terminus of subunit F in the catalysis process.

The H subunit of the A_1A_O ATP synthase is a component of one of the stator domain connecting the A_1 and A_O sector and has an important role in the assembly and function of the enzyme. In order to get details about the stalk part of the enzyme subunit H was efficiently expressed and purified. CD spectroscopy of subunit H showed a 78% α-helix and a coiled-coil arrangement. Cross linking studies using a series of deletions and disulfide formation between introduced cysteines indicate that this peripheral stalk is divided into two structural parts. The solution structure of the peptide comprising the last twenty amino acids of subunit H, H_{85-104}, together with the mutagenesis, cross linking analysis at the C-terminus of entire H subunit indicate that the C-terminus is important for the structure formation of subunit H. In order to understand, how the N-terminal domain of subunit H is formed, the first solution NMR structure of H_{1-47} has been determined. The structure consists of an α-helix between residues 15 to 42 and a flexible N-terminal stretch (1-14). The α-helix includes a kink, which would bring the two helices of the C-terminus in the coiled-coil arrangement. The structure of H_{1-47} reveals a strip of hydrophobic alanine residues. NMR measurements with H_{1-47} as function of concentration clearly point out that H_{1-47} forms a dimer at concentrations above ~3 mM. The ^{1}H and ^{15}N chemical shift perturbation experiment indicates that the residues A19, I23, E24, A26, K27, I33, K34 and A37 are involved in dimer formation. Based on our structural model the charged residues Q32, E36 and E39 exhibiting comparably large shifts, might not be directly situated at

the putative dimeric interface but rather their side-chain pK_a might be affected via proxy charged residues from the other monomer. However, the maximum chemical shift change has been observed for the residues 6-21 and residues 43-47. NMR secondary chemical shifts as well as ^{15}N relaxation data reveal that these regions are less structured as compared to the continuous central α-helix. Some structural consolidation at the N-terminus is seen as the direct consequence of dimerization evidenced by the increased order parameters.

Using the purified N- (E_{1-100}) and C-terminal segment ($E_{101-206}$) of subunit E, NMR titration experiments revealed that the N-terminal residues Met_1 to Met_6, Lys_{10} to Glu_{11}, Ala_{15}, Val_{20} and Glu_{24} of H_{1-47} interact specifically with the N-terminal peptide, E_{1-100} of subunit E. A more detailed picture regarding the residues of E_{1-100} involved in the interaction was obtained by titration experiments using the N-terminal peptides E_{1-20}, E_{21-40} and E_{41-60}. The data indicate that the N-terminal tail E_{41-60} interacts with the N-terminal amino acids of H_{1-47} and has been confirmed by fluorescence correlation spectroscopy results. In order to understand further the E-H assembly the N-terminal structure of E_{1-52} of *M. jannaschii* was solved by NMR spectroscopy, showing an α-helix between the residues 8-48. The molecule is amphipathic with a strip of hydrophobic residues, which may associate with the neighboring subunit H via a possible helix-helix interaction.

6. References

Abrahams, J. P., A. G. W. Leslie, R. Lutter and J. E. Walker (1994). *Structure at 2.8 Å resolution of F_1-ATPase from bovine heart mitochondria*. Nature **370**: 621-628.

Aggeler, R., G. Grüber and R. A. Capaldi (1998). *Trapping of conformations of the Escherichia coli F_1 ATPase by disulfide bond formation: A state of the enzyme with all three catalytic sites of equal and low affinity for nucleotides*. FEBS Lett. **426**: 37-40.

Bernal, R. A. and D. Stock (2004). *Three-Dimensional Structure of the Intact Thermus thermophilus H+-ATPase/Synthase by Electron Microscopy*. Structure **12**: 1789-1798.

Bhatt, D., S. P. Cole, T. B. Grabar, S. B. Claggett and B. D. Cain (2005). *Manipulating the Length of the b Subunit F_1 Binding Domain in F_1F_O ATP Synthase from Escherichia coli*. J. Bioenerg. Biomembr. **37**: 67-74.

Biuković, G., S. Gayen, K. Pervushin and G. Grüber (2009). *The domain features of the peripheral stalk subunit H of the methanogenic A_1A_O ATP synthase and the NMR solution structure of H_{1-47}*. Biophys. J. **97**: 286-294.

Biuković, G., M. Rossle, S. Gayen, Y. Mu and G. Grüber (2007). *Small-Angle X-ray Scattering Reveals the Solution Structure of the Peripheral Stalk Subunit H of the A_1A_O ATP Synthase from Methanocaldococcus jannaschii and Its Binding to the Catalytic A Subunit*. Biochemistry **46**: 2070-2078.

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

Böttcher, B., I. Bertsche, R. Reuter and P. Gräber (2000). *Direct visualisation of conformational changes in EF_OF_1 by electron microscopy*. J. Mol. Biol. **296**: 449-457.

Capaldi, R. A. and R. Aggeler (2002). *Mechanism of the F_1F_O-type ATP synthase, a biological rotary motor*. Trends Biochem. Sci. **27**: 154-160.

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

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

Coskun, Ü., Y. L. Chaban, A. Lingl, V. Müller, W. Keegstra, E. J. Boekema and G. Grüber (2004). *Structure and Subunit Arrangement of the A-type ATP Synthase Complex from the Archaeon Methanococcus jannaschii Visualized by Electron Microscopy*. J. Biol. Chem. **279**: 38644-38648.

Coskun, Ü., G. Grüber, M. H. J. Koch, J. Godovac-Zimmermann, T. Lemker and V. Müller (2002). *Cross-talk in the A_1-ATPase from Methanosarcina mazei Gö1 Due to Nucleotide Binding*. J. Biol. Chem. **277**: 17327-17333.

Coskun, Ü., M. Radermacher, V. Müller, T. Ruiz and G. Grüber (2004). *Three-dimensional Organization of the Archaeal A_1-ATPase from Methanosarcina mazei Gö1*. J. Biol. Chem. **279**: 22759-22764.

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

Del Rizzo, P. A., Y. Bi and S. D. Dunn (2006). *ATP Synthase b Subunit Dimerization Domain: A Right-Handed Coiled-coil with Offset Helices*. J. Mol. Biol. **364**: 735-746.

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

Deppenmeier, U. (2002). *Redox-driven proton translocation in methanogenic Archaea*. Cell. Mol. Life Sci. **59**: 1513-1533.

Deppenmeier, U. (2002). The unique biochemistry of methanogenesis. Prog. Nucl. Acid Re. **71**: 223-283.

Dunn, S. D., M. Revington, D. J. Cipriano and B. H. Shilton (2000). *The b Subunit of Escherichia coli ATP Synthase*. J. Bioenerg. Biomembr. **32**: 347-355.

Eghbalnia, H. R., L. Wang, A. Bahrami, A. Assadi and J. Markley (2005). *Protein energetic conformational analysis from NMR chemical shifts (PECAN) and its use in determining secondary structural elements*. J. Biomol. NMR **32**: 71-81.

Feniouk, B. A., T. Suzuki and M. Yoshida (2006). *The role of subunit epsilon in the catalysis and regulation of F_OF_1-ATP synthase*. Biochim. Biophys. Acta-Bioenergetics **1757**: 326-338.

Gayen, S., A. M. Balakrishna and G. Grüber (2009). *NMR solution structure of the N–terminal domain of subunit E (E_{1-52}) of A_1A_O ATP synthase from Methanocaldococcus jannaschii*. J. Bioenerg. Biomembr. **41**: 342-348.

Gayen, S., A. M. Balakrishna, G. Biuković, W. Yulei, C. Hunke and G. Grüber (2008). *Identification of critical residues of subunit H in its interaction with subunit e of the A-ATP synthase from Methanocaldococcus jannaschii*. FEBS J. **275**: 1803-1812.

Gayen, S., S. Vivekanandan, G. Biuković, G. Grüber and H. S. Yoon (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.

Gibbons, C., M. G. Montgomery, A. G. W. Leslie and J. E. Walker (2000). *The structure of the central stalk in bovine F_1-ATPase at 2.4 Å resolution*. Nat. Struct. Mol. Biol. **7**: 1055-1061.

Girvin, M. E., V. K. Rastogi, F. Abildgaard, J. L. Markley and R. H. Fillingame (1998). *Solution Structure of the Transmembrane H+-Transporting Subunit c of the F_1F_O ATP Synthase.* Biochemistry **37**: 8817-8824.

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

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

Grüber, G. (2003). *Introduction: A Close Look at the Vacuolar ATPase*. J. Bioenerg. Biomembr. **35**: 277-280.

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

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

Grüber, G. and V. Marshansky (2008). *New insights into structure-function relationships between archeal ATP synthase and vacuolar type ATPase*. BioEssays **30**: 1096-1109.

Grüber, G., D. I. Svergun, Ü. Coskun, T. Lemker, M. H. J. Koch, H. Schagger and V. Müller (2001). *Structural Insights into the A_1 ATPase from the Archaeon, Methanosarcina mazei Gö1* Biochemistry **40**: 1890-1896.

Grüber, G., D. I. Svergun, Ü. Coskun, T. Lemker, M. H. J. Koch, H. Schagger and V. Müller (2001). *Structural Insights into the A_1 ATPase from the Archaeon, Methanosarcina mazei Gö1*. Biochemistry **40**: 1890-1896.

Guex, N. and M. Peitsch (1997). *SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling*. Electrophoresis **18**: 2714-2723.

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

Hao, Y. and I. G. Macara (2008). *Regulation of chromatin binding by a conformational switch in the tail of the Ran exchange factor RCC1*. J. Cell Biol. **182**: 827-836.

Hausrath, A. C., R. A. Capaldi and B. W. Matthews (2001). *The Conformation of the Epsilon and Gamma Subunits within the Escherichia coli F_1 ATPase*. J. Biol. Chem. **276**: 47227-47232.

Hausrath, A. C., G. Grüber, B. W. Matthews and R. A. Capaldi (1999). *Structural features of the γ subunit of the Escherichia coli F_1 ATPase revealed by a 4.4-Å resolution map obtained by x-ray crystallography*. P. Natl. Acad. Sci. USA **96**: 13697-13702.

Ihara, K., S. Watanabe, K. Sugimura, I. Katagiri and Y. Mukohata (1997). *Identification of Proteolipid from an Extremely Halophilic ArchaeonHalobacterium salinarumas an N,N'-Dicyclohexyl-carbodiimide Binding Subunit of ATP Synthase*. Arch. Biochem. Biophys. **341**: 267-272.

Iino, R., R. Hasegawa, K. V. Tabata and H. Noji (2009). *Mechanism of inhibition by C-terminal α-helices of the ε subunit of Escherichia coli F_OF_1-ATP synthase*. J. Biol. Chem. **284**: 17457-64.

Iwata, M., H. Imamura, E. Stambouli, C. Ikeda, M. Tamakoshi, K. Nagata, H. Makyio, B. Hankamer, J. Barber, M. Yoshida, K. Yokoyama and S. Iwata (2004). *Crystal structure of a central stalk subunit C and reversible association/dissociation of vacuole-type ATPase*. P. Natl. Acad. Sci. USA **101**: 59-64.

Karrasch, S. and J. E. Walker (1999). *Novel features in the structure of bovine ATP synthase*. J. Mol. Biol. **290**: 379-384.

Kato, S., M. Yoshida and Y. Kato-Yamada (2007). *Role of the ε Subunit of Thermophilic F_1-ATPase as a Sensor for ATP*. J. Biol. Chem. **282**: 37618-37623.

Kish-Trier, E., L.-A. K. Briere, S. D. Dunn and S. Wilkens (2008). *The Stator Complex of the A_1A_O-ATP Synthase--Structural Characterization of the E and H Subunits*. J. Mol. Biol. **375**: 673-685.

Kish-Trier, E. and S. Wilkens (2009). *Domain Architecture of the Stator Complex of the A_1A_O-ATP Synthase from Thermoplasma acidophilum*. J. Biol. Chem. **284**: 12031-12040.

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

Koradi, R., M. Billeter and K. Wüthrich (1996). *MOLMOL: a program for display and analysis of macromolecular structures*. J. Mol. Graph. **14**: 51-55.

Kumar, A., M. S. S. Manimekalai, A. M. Balakrishna, C. Hunke, S. Weigelt, N. Sewald and G. Grüber (2009). *Spectroscopic and crystallographic studies of the mutant R416W give insight into the nucleotide binding traits of subunit B of the A_1Ao ATP synthase*. Proteins: Struct. Funct. Bioinf. **75**: 807-819.

Kumar, S. and M. Bansal (1996). *Structural and sequence characteristics of long alpha helices in globular proteins*. Biophys. J. **71**: 1574-1586.

Kuntz, I. D., P. A. Kosen and E. C. Craig (1991). *Amide chemical shifts in many helices in peptides and proteins are periodic*. J. Am. Chem. Soc. **113**: 1406-1408.

Lee, S. J., Y. Matsuura, S. M. Liu and M. Stewart (2005). *Structural basis for nuclear import complex dissociation by RanGTP*. Nature **435**: 693-696.

Lewalter, K. and V. Müller (2006). *Bioenergetics of archaea: Ancient energy conserving mechanisms developed in the early history of life*. Biochim. Biophys. Acta (BBA) - Bioenergetics **1757**: 437-445.

Lokanath, N. K., Y. Matsuura, C. Kuroishi, N. Takahashi and N. Kunishima (2007). *Dimeric Core Structure of Modular Stator Subunit E of Archaeal H+-ATPase*. J. Mol. Biol. **366**: 933-944.

Maegawa, Y., H. Morita, D. Iyaguchi, M. Yao, N. Watanabe and I. Tanaka (2006). *Structure of the catalytic nucleotide-binding subunit A of A-type ATP synthase from Pyrococcus horikoshii reveals a novel domain related to the peripheral stalk*. Acta Crystallogr. D **62**: 483-488.

Manimekalai, M. S. S., A. Kumar, A. M. Balakrishna and G. Grüber (2009). *A second transient position of ATP on its trail to the nucleotide-binding site of subunit B of the motor protein A_1A_O ATP synthase*. J. Struc. Biol. **166**: 38-45.

McGuffin, L., K. Bryson and D. Jones (2000). *The PSIPRED protein structure prediction server*. Bioinformatics **16**: 404-405.

McLachlin, D. T., A. M. Coveny, S. M. Clark and S. D. Dunn (2000). *Site-directed Cross-linking of b to the alpha , beta , and a Subunits of the Escherichia coli ATP Synthase*. J. Biol. Chem.. **275**: 17571-17577.

Menz, R. I., J. E. Walker and A. G. W. Leslie (2001). *Structure of Bovine Mitochondrial F1-ATPase with Nucleotide Bound to All Three Catalytic Sites: Implications for the Mechanism of Rotary Catalysis*. Cell **106**: 331-341.

Müller, V. and G. Grüber (2003). *ATP synthases: structure, function and evolution of unique energy converters*. Cell. Mol. Life Sci. **60**: 474-494.

Müller, V., A. Lingl, K. Lewalter and M. Fritz (2005). *ATP Synthases With Novel Rotor Subunits: New Insights into Structure, Function and Evolution of ATPases*. J. Bioenerg. Biomembr. **37**: 455-460.

Murata, T., I. Yamato, Y. Kakinuma, A. G. W. Leslie and J. E. Walker (2005). *Structure of the Rotor of the V-Type Na+-ATPase from Enterococcus hirae*. Science **308**: 654-659.

Nelson, H., S. Mandiyan and N. Nelson (1989). *A conserved gene encoding the 57-kD a subunit of the yeast vacuolar H+- ATPase*. J. Biol. Chem. **264**: 1775-1778.

Nielsen, K. J., J. M. Hill, M. A. Anderson and D. J. Craik (1996). *Synthesis and Structure Determination by NMR of a Putative Vacuolar Targeting Peptide and Model of a Proteinase Inhibitor from Nicotiana alata*. Biochemistry **35**: 369-378.

Numoto, N., A. Kita and K. Miki (2004). *Structure of the C subunit of V-type ATPase from Thermus thermophilus at 1.85 Å resolution*. Acta Crystallogr. D **D60**: 810-815.

Pervushin, K., G. Wider and K. Wuthrich (1997). *Deuterium Relaxation in a Uniformly ^{15}N-Labeled Homeodomain and Its DNA Complex 1*. J. Am. Chem. Soc. **119**: 3842-3843.

Priya, R., G. Biuković, S. Gayen, S. Vivekanandan and G. Grüber (2009). *NMR solution structure of the b30-82 domain of subunit b of Escherichia coli F_1F_O ATP synthase*. J. Bacteriol. **191**: in press.

Priya, R., V. Tadwal, M. Rössle, S. Gayen, C. Hunke, W. Peng, J. Torres and G. Grüber (2008). *Low resolution structure of subunit b (b 22–156) of Escherichia coli F_1F_O ATP synthase in solution and the b –δ assembly*. J. Bioenerg. Biomembr. **40**: 245-255.

Qi, J. and M. Forgac (2007). *Cellular Environment Is Important in Controlling V-ATPase Dissociation and Its Dependence on Activity*. J. Biol. Chem. **282**: 24743-24751.

Raghunathan, D., S. Gayen, C. S. Verma and G. Grüber (2009). *Cross-talk along the stalk: Dynamics of the interaction of B and F in A_1A_O ATP Synthase of Methanosarcina mazei Gö1*. Biochemistry. submitted.

Rodgers, A. J. W. and M. C. J. Wilce (2000). *Structure of the γ-ε complex of ATP synthase*. Nat. Struc. Mol. Biol. **7**: 1051-1054.

Rodgers, A. J. W., S. Wilkens, R. Aggeler, M. B. Morris, S. M. Howitt and R. A. Capaldi (1997). *The Subunit delta - Subunit b Domain of the Escherichia coli F_1F_O ATPase. The b subunits interact with F_1 as a dimer and through the δ subunit* J. Biol. Chem.. **272**: 31058-31064.

Rudolf, K. T., K. Anne-Kristin, S. Henning, B. Wolfgang and H. Reiner (2008). *Methanogenic archaea: ecologically relevant differences in energy conservation*. Nat. Rev. Microbiol. **6**: 579-591.

Ruppert, C., H. Kavermann, S. Wimmers, R. Schmid, J. Kellermann, F. Lottspeich, H. Huber, K. O. Stetter and V. Muller (1999). *The Proteolipid of the A_1A_O ATP Synthase from Methanococcus jannaschii Has Six Predicted Transmembrane Helices but Only Two Proton-translocating Carboxyl Groups*. J. Biol. Chem. **274**: 25281-25284.

Saraste, M., P. R. Sibbald and A. Wittinghofer (1990). *The P-loop-a common motif in ATP- and GTP binding proteins*. Trends Biochem. Sci. **15**: 430–434.

Schäfer, G., M. Engelhard and V. Müller (1999). *Bioenergetics of the Archaea*. Microbiol. Mol. Biol. Rev. **63**: 570-620.

Schäfer, I. B., S. M. Bailer, M. G. Düser, M. Börsch, R. A. Bernal, D. Stock and G. Grüber (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.

Schäfer, I., M. Rössle, G. Biuković, V. Müller and G. Grüber (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.

Senior, A. E., A. Muharemagic and S. Wilke-Mounts (2006). *Assembly of the Stator in Escherichia coli ATP Synthase. Complexation of α Subunit with Other F_1 Subunits Is Prerequisite for δ Subunit Binding to the N-Terminal Region of α.* Biochemistry **45**: 15893-15902.

Shaka, A., P. Barker and R. Freeman (1985). *Computeroptimized decoupling scheme for wideband applications and low level operation*. J. Magn. Reson. **64**: 547-552.

Steinert, K., V. Wagner, P. G. Kroth-Pancic and S. Bickel-Sandkotter (1997). *Characterization and Subunit Structure of the ATP Synthase of the Halophilic Archaeon Haloferax volcanii and Organization of the ATP Synthase Genes*. J. Biol. Chem. **272**: 6261-6269.

Stewart, M., H. M. Kent and A. J. McCoy (1998). *The structure of the Q69L mutant of GDP-ran shows a major conformational change in the switch II loop that accounts for its failure to bind nuclear transport factor 2 (NTF2)*. J. Mol. Biol. **284**: 1517-1527.

Stock, D., C. Gibbons, I. Arechaga, A. G. W. Leslie and J. E. Walker (2000). *The rotary mechanism of ATP synthase*. Curr. Opin. Struc. Biol. **10**: 672-679.

Stock, D., A. G. Leslie, nbsp, W and J. E. Walker (1999). *Molecular Architecture of the Rotary Motor in ATP Synthase*. Science **286**: 1700-1705.

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

Taylor, R. M., S. D. Zakharov, J. B. Heymann, M. E. Girvin and W. A. Cramer (2000). *Folded State of the Integral Membrane Colicin E1 Immunity Protein in Solvents of Mixed Polarity*. Biochemistry **39**: 12131-12139.

Tsumuraya, M., S. Furuike, K. Adachi, K. Kinosita Jr and M. Yoshida (2009). *Effect of ε subunit on the rotation of thermophilic Bacillus F_1-ATPase*. FEBS Lett. **583**: 1121-1126.

Uhlin, U., G. B. Cox and J. M. Guss (1997). *Crystal structure of the ε subunit of the proton-translocating ATP synthase from Escherichia coli*. Structure **5**: 1219-1230.

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

Venkataraman, K., P. Neeti, E. W. John, G. W. L. Andrew and M. M. David (2006). *Novel features of the rotary catalytic mechanism revealed in the structure of yeast F_1 ATPase*. EMBO J. **25**: 5433-5442.

Vetter, I. R., C. Nowak, T. Nishimoto, J. Kuhlmann and A. Wittinghofer (1999). *Structure of a Ran-binding domain complexed with Ran bound to a GTP analogue: implications for nuclear transport*. Nature **398**: 39-46.

Vonck, J., K. Y. Pisa, N. Morgner, B. Brutschy and V. Müller (2009). *Three-dimensional Structure of A1A0 ATP Synthase from the Hyperthermophilic Archaeon Pyrococcus furiosus by Electron Microscopy*. J. Biol. Chem. **284**: 10110-10119.

Vonck, J., T. K. von Nidda, T. Meier, U. Matthey, D. J. Mills, W. Kühlbrandt and P. Dimroth (2002). *Molecular Architecture of the Undecameric Rotor of a Bacterial Na+-ATP Synthase*. J. Mol. Biol. **321**: 307-316.

Wagner, G., A. Pardi and K. Wüthrich (1983). *Hydrogen bond length and proton NMR chemical shifts in proteins*. J. Am. Chem. Soc. **105**: 5948-5949.

Walker, J. E. and V. K. Dickson (2006). *The peripheral stalk of the mitochondrial ATP synthase*. Biochim. Biophys. Acta (BBA) - Bioenergetics **1757**: 286-296.

Walker, J. E., M. Saraste, M. J. Runswick, N. J. Gay and (1982). *Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold*. EMBO J. **1**: 945–951.

Weber, J. (2006). *ATP synthase: Subunit-subunit interactions in the stator stalk*. Biochim. Biophys. Acta (BBA) - Bioenergetics **1757**: 1162-1170.

Wilkens, S., D. Borchardt, J. Weber and A. E. Senior (2005). *Structural Characterization of the Interaction of the δ and α Subunits of the Escherichia coli F_1F_O-ATP Synthase by NMR Spectroscopy*. Biochemistry **44**: 11786-11794.

Wilkens, S. and R. A. Capaldi (1998). *Electron microscopic evidence of two stalks linking the F_1 and F_O parts of the Escherichia coli ATP synthase*. Biochim. Biophys. Acta (BBA) - Bioenergetics **1365**: 93-97.

Wilkens, S., F. W. Dahlquist, L. P. McIntosh, L. W. Donaldson and R. A. Capaldi (1995). *Structural features of the ε subunit of the Escherichia coli ATP synthase determined by NMR spectroscopy*. Nat. Struc. Mol. Biol. **2**: 961-967.

Wilkens, S., S. Dunn, J. Chandler, F. Dahlquist and R. Capaldi (1997). *Solution structure of the N-terminal domain of the delta subunit of the E. coli ATPsynthase*. Nat. Struc. Biol. **4**: 198-201.

Wilms, R., C. Freiberg, E. Wegerle, I. Meier, F. Mayer and V. Müller (1996). *Subunit Structure and Organization of the Genes of the A_1A_O ATPase from the Archaeon Methanosarcina mazei Go1*. J. Biol. Chem. **271**: 18843-18852.

Wood, K. S. and S. D. Dunn (2007). *Role of the Asymmetry of the Homodimeric b2 Stator Stalk in the Interaction with the F1 Sector of Escherichia coli ATP Synthase*. J. Biol. Chem.. **282**: 31920-31927.

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

Yagi, H., N. Kajiwara, H. Tanaka, T. Tsukihara, Y. Kato-Yamada, M. Yoshida and H. Akutsu (2007). *Structures of the thermophilic F_1-ATPase ε subunit suggesting ATP-regulated arm motion of its C-terminal domain in F_1*. P. Natl. Acad. Sci. USA **104**: 11233-11238.

Zhou, N. E., B. Y. Zhu, B. D. Sykes and R. S. Hodges (1992). *Relationship between amide proton chemical shifts and hydrogen bonding in amphipathic .alpha.-helical peptides*. J. Am. Chem. Soc. **114**: 4320-4326.

Appendix A

Chemical shift assignments of subunit F of A_1A_O ATP synthase from *M. mazei* Gö1

The chemical shift assignment are listed in the NMR-STAR format (Columns are as follows: Atom shift assign ID, Residue author sequence code, Residue sequence code, Residue label, Atom name, Atom type, Chemical shift value, Chemical shift value error, Chemical shift ambiguity code

1	1	9	PRO	HA	H	4.193	0.020	1
2	1	9	PRO	HB2	H	2.092	0.020	2
3	1	9	PRO	HB3	H	1.842	0.020	2
4	1	9	PRO	HG2	H	1.762	0.020	1
5	1	9	PRO	HG3	H	1.762	0.020	1
6	1	9	PRO	HD2	H	3.493	0.020	2
7	1	9	PRO	HD3	H	3.248	0.020	2
8	1	9	PRO	C	C	174.100	0.400	1
9	1	9	PRO	CA	C	60.950	0.400	1
10	1	9	PRO	CB	C	29.320	0.400	1
11	1	9	PRO	CG	C	24.540	0.400	1
12	1	9	PRO	CD	C	47.880	0.400	1
13	2	10	MET	H	H	8.610	0.020	1
14	2	10	MET	HA	H	4.559	0.020	1
15	2	10	MET	HB2	H	1.895	0.020	2
16	2	10	MET	HB3	H	2.049	0.020	2
17	2	10	MET	HG2	H	2.399	0.020	2
18	2	10	MET	HG3	H	2.471	0.020	2
19	2	10	MET	C	C	173.000	0.400	1
20	2	10	MET	CA	C	52.240	0.400	1
21	2	10	MET	CB	C	29.840	0.400	1
22	2	10	MET	CG	C	29.420	0.400	1
23	2	10	MET	N	N	119.600	0.400	1
24	3	11	GLU	H	H	8.242	0.020	1
25	3	11	GLU	HA	H	4.525	0.020	1
26	3	11	GLU	HB2	H	1.867	0.020	2
27	3	11	GLU	HB3	H	1.765	0.020	2
28	3	11	GLU	HG2	H	2.165	0.020	2
29	3	11	GLU	HG3	H	2.212	0.020	2
30	3	11	GLU	C	C	173.100	0.400	1
31	3	11	GLU	CA	C	53.600	0.400	1
32	3	11	GLU	CB	C	28.530	0.400	1
33	3	11	GLU	CG	C	33.620	0.400	1
34	3	11	GLU	N	N	118.700	0.400	1
35	4	12	LEU	H	H	8.406	0.020	1
36	4	12	LEU	HA	H	5.148	0.020	1
37	4	12	LEU	HB2	H	1.452	0.020	1
38	4	12	LEU	HB3	H	1.452	0.020	1
39	4	12	LEU	HG	H	1.702	0.020	1
40	4	12	LEU	HD1	H	0.631	0.020	2
41	4	12	LEU	HD2	H	0.704	0.020	2
42	4	12	LEU	C	C	171.800	0.400	1
43	4	12	LEU	CA	C	51.690	0.400	1
44	4	12	LEU	CB	C	42.340	0.400	1
45	4	12	LEU	CG	C	29.650	0.400	1
46	4	12	LEU	CD1	C	23.490	0.400	1
47	4	12	LEU	CD2	C	22.810	0.400	1
48	4	12	LEU	N	N	124.900	0.400	1
49	5	13	ALA	H	H	8.704	0.020	1
50	5	13	ALA	HA	H	5.439	0.020	1
51	5	13	ALA	HB	H	1.299	0.020	1
52	5	13	ALA	C	C	172.600	0.400	1
53	5	13	ALA	CA	C	47.370	0.400	1
54	5	13	ALA	CB	C	22.090	0.400	1
55	5	13	ALA	N	N	123.000	0.400	1
56	6	14	VAL	H	H	8.319	0.020	1
57	6	14	VAL	HA	H	5.021	0.020	1
58	6	14	VAL	HB	H	1.351	0.020	1
59	6	14	VAL	HG1	H	0.147	0.020	2
60	6	14	VAL	HG2	H	0.356	0.020	2
61	6	14	VAL	C	C	170.600	0.400	1
62	6	14	VAL	CA	C	57.040	0.400	1
63	6	14	VAL	CB	C	32.800	0.400	1
64	6	14	VAL	CG1	C	18.530	0.400	1
65	6	14	VAL	CG2	C	17.910	0.400	1
66	6	14	VAL	N	N	118.400	0.400	1
542	50	58	VAL	CB	C	30.110	0.400	1
543	50	58	VAL	CG1	C	18.840	0.400	1
544	50	58	VAL	CG2	C	18.730	0.400	1
545	50	58	VAL	N	N	128.100	0.400	1
546	51	59	MET	H	H	8.408	0.020	1
547	51	59	MET	HA	H	4.912	0.020	1
548	51	59	MET	HB2	H	2.133	0.020	1
549	51	59	MET	HB3	H	2.133	0.020	1
550	51	59	MET	HG2	H	2.327	0.020	2
551	51	59	MET	HG3	H	2.571	0.020	2
552	51	59	MET	C	C	170.200	0.400	1
553	51	59	MET	CA	C	51.840	0.400	1
554	51	59	MET	CB	C	37.420	0.400	1
555	51	59	MET	CG	C	30.340	0.400	1
556	51	59	MET	N	N	123.100	0.400	1
557	52	60	HIS	H	H	10.310	0.020	1
558	52	60	HIS	HA	H	4.370	0.020	1
559	52	60	HIS	HB2	H	1.377	0.020	2
560	52	60	HIS	HB3	H	2.484	0.020	2
561	52	60	HIS	HD2	H	7.273	0.020	1
562	52	60	HIS	HE1	H	7.191	0.020	1
563	52	60	HIS	C	C	175.100	0.400	1
564	52	60	HIS	CA	C	53.730	0.400	1
565	52	60	HIS	CB	C	29.290	0.400	1
566	52	60	HIS	N	N	123.800	0.400	1
567	53	61	ASN	H	H	8.728	0.020	1
568	53	61	ASN	HA	H	4.110	0.020	1
569	53	61	ASN	HB2	H	2.461	0.020	2
570	53	61	ASN	HB3	H	2.217	0.020	2
571	53	61	ASN	HD21	H	6.632	0.020	2
572	53	61	ASN	HD22	H	7.400	0.020	2
573	53	61	ASN	C	C	174.900	0.400	1
574	53	61	ASN	CA	C	54.340	0.400	1
575	53	61	ASN	CB	C	36.970	0.400	1
576	53	61	ASN	N	N	126.600	0.400	1
577	53	61	ASN	ND2	N	112.500	0.400	1
578	54	62	ASP	H	H	11.520	0.020	1
579	54	62	ASP	HA	H	4.172	0.020	1
580	54	62	ASP	HB2	H	2.180	0.020	2
581	54	62	ASP	HB3	H	2.486	0.020	2
582	54	62	ASP	C	C	175.800	0.400	1
583	54	62	ASP	CA	C	54.590	0.400	1
584	54	62	ASP	CB	C	37.050	0.400	1
585	54	62	ASP	N	N	123.600	0.400	1
586	55	63	ASP	H	H	7.239	0.020	1
587	55	63	ASP	HA	H	4.435	0.020	1
588	55	63	ASP	HB2	H	2.280	0.020	2
589	55	63	ASP	HB3	H	2.876	0.020	2
590	55	63	ASP	C	C	176.400	0.400	1
591	55	63	ASP	CA	C	53.640	0.400	1
592	55	63	ASP	CB	C	37.630	0.400	1
593	55	63	ASP	N	N	117.600	0.400	1
594	56	64	ILE	H	H	7.530	0.020	1
595	56	64	ILE	HA	H	3.306	0.020	1
596	56	64	ILE	HB	H	1.693	0.020	1
597	56	64	ILE	HG13	H	1.523	0.020	2
598	56	64	ILE	HG2	H	0.714	0.020	1
599	56	64	ILE	HD1	H	0.673	0.020	1
600	56	64	ILE	C	C	175.700	0.400	1
601	56	64	ILE	CA	C	62.780	0.400	1
602	56	64	ILE	CB	C	35.240	0.400	1
603	56	64	ILE	CG1	C	26.290	0.400	1
604	56	64	ILE	CG2	C	14.430	0.400	1
605	56	64	ILE	CD1	C	10.780	0.400	1
606	56	64	ILE	N	N	120.000	0.400	1
607	57	65	GLY	H	H	7.650	0.020	1

```
  67   7  15 ILE H    H   7.884 0.020 1
  68   7  15 ILE HA   H   5.190 0.020 1
  69   7  15 ILE HB   H   1.661 0.020 1
  70   7  15 ILE HG12 H   1.461 0.020 2
  71   7  15 ILE HG13 H   1.473 0.020 2
  72   7  15 ILE HG2  H   0.908 0.020 1
  73   7  15 ILE HD1  H   0.669 0.020 1
  74   7  15 ILE C    C 172.400 0.400 1
  75   7  15 ILE CA   C  56.630 0.400 1
  76   7  15 ILE CB   C  40.980 0.400 1
  77   7  15 ILE CG1  C  23.900 0.400 1
  78   7  15 ILE CG2  C  14.920 0.400 1
  79   7  15 ILE CD1  C  11.900 0.400 1
  80   7  15 ILE N    N 125.700 0.400 1
  81   8  16 GLY H    H   8.284 0.020 1
  82   8  16 GLY HA2  H   3.843 0.020 2
  83   8  16 GLY HA3  H   4.588 0.020 2
  84   8  16 GLY C    C 169.000 0.400 1
  85   8  16 GLY CA   C  42.470 0.400 1
  86   8  16 GLY N    N 112.000 0.400 1
  87   9  17 LYS H    H   9.168 0.020 1
  88   9  17 LYS HA   H   4.324 0.020 1
  89   9  17 LYS HB2  H   1.739 0.020 2
  90   9  17 LYS HB3  H   2.042 0.020 2
  91   9  17 LYS HG2  H   1.560 0.020 2
  92   9  17 LYS HG3  H   1.546 0.020 2
  93   9  17 LYS HD2  H   1.697 0.020 2
  94   9  17 LYS HD3  H   1.658 0.020 2
  95   9  17 LYS HE2  H   2.895 0.020 2
  96   9  17 LYS HE3  H   3.275 0.020 2
  97   9  17 LYS HZ   H   7.279 0.020 1
  98   9  17 LYS C    C 176.300 0.400 1
  99   9  17 LYS CA   C  53.530 0.400 1
 100   9  17 LYS CB   C  31.110 0.400 1
 101   9  17 LYS CG   C  22.520 0.400 1
 102   9  17 LYS CD   C  25.740 0.400 1
 103   9  17 LYS CE   C  39.160 0.400 1
 104   9  17 LYS N    N 119.300 0.400 1
 105  10  18 SER H    H   9.347 0.020 1
 106  10  18 SER HA   H   3.827 0.020 1
 107  10  18 SER HB2  H   3.771 0.020 1
 108  10  18 SER HB3  H   3.771 0.020 1
 109  10  18 SER C    C 174.500 0.400 1
 110  10  18 SER CA   C  60.200 0.400 1
 111  10  18 SER CB   C  62.110 0.400 1
 112  10  18 SER N    N 116.700 0.400 1
 113  11  19 GLU H    H   9.453 0.020 1
 114  11  19 GLU HA   H   4.010 0.020 1
 115  11  19 GLU HB2  H   1.958 0.020 1
 116  11  19 GLU HB3  H   1.958 0.020 1
 117  11  19 GLU HG2  H   2.254 0.020 1
 118  11  19 GLU HG3  H   2.254 0.020 1
 119  11  19 GLU C    C 174.200 0.400 1
 120  11  19 GLU CA   C  56.880 0.400 1
 121  11  19 GLU N    N 118.900 0.400 1
 122  12  20 PHE H    H   7.585 0.020 1
 123  12  20 PHE HA   H   4.651 0.020 1
 124  12  20 PHE HB2  H   3.169 0.020 2
 125  12  20 PHE HB3  H   3.288 0.020 2
 126  12  20 PHE HD1  H   7.273 0.020 1
 127  12  20 PHE HD2  H   7.273 0.020 1
 128  12  20 PHE HE1  H   7.583 0.020 1
 129  12  20 PHE HE2  H   7.583 0.020 1
 130  12  20 PHE C    C 174.300 0.400 1
 131  12  20 PHE CA   C  57.000 0.400 1
 132  12  20 PHE N    N 119.700 0.400 1
 133  13  21 VAL H    H   8.154 0.020 1
 134  13  21 VAL HA   H   3.946 0.020 1
 135  13  21 VAL HB   H   2.054 0.020 1
 136  13  21 VAL HG1  H   0.744 0.020 2
 137  13  21 VAL HG2  H   0.744 0.020 2
 138  13  21 VAL C    C 175.900 0.400 1
 139  13  21 VAL CA   C  61.000 0.400 1
 140  13  21 VAL CB   C  29.190 0.400 1
 141  13  21 VAL CG1  C  16.500 0.400 1
 608  57  65 GLY HA2  H   3.702 0.020 2
 609  57  65 GLY HA3  H   3.763 0.020 2
 610  57  65 GLY C    C 171.300 0.400 1
 611  57  65 GLY CA   C  43.800 0.400 1
 612  57  65 GLY N    N 102.700 0.400 1
 613  58  66 ASN H    H   6.995 0.020 1
 614  58  66 ASN HA   H   4.795 0.020 1
 615  58  66 ASN HB2  H   2.696 0.020 2
 616  58  66 ASN HB3  H   2.881 0.020 2
 617  58  66 ASN HD21 H   7.582 0.020 2
 618  58  66 ASN HD22 H   6.870 0.020 2
 619  58  66 ASN C    C 172.500 0.400 1
 620  58  66 ASN CA   C  50.180 0.400 1
 621  58  66 ASN CB   C  37.140 0.400 1
 622  58  66 ASN N    N 115.200 0.400 1
 623  58  66 ASN ND2  N 113.800 0.400 1
 624  59  67 LEU H    H   7.320 0.020 1
 625  59  67 LEU HA   H   4.356 0.020 1
 626  59  67 LEU HB2  H   2.109 0.020 2
 627  59  67 LEU HB3  H   1.821 0.020 2
 628  59  67 LEU HG   H   1.496 0.020 1
 629  59  67 LEU HD1  H   1.141 0.020 2
 630  59  67 LEU HD2  H   0.673 0.020 2
 631  59  67 LEU C    C 171.100 0.400 1
 632  59  67 LEU CA   C  50.600 0.400 1
 633  59  67 LEU CB   C  38.470 0.400 1
 634  59  67 LEU CG   C  24.920 0.400 1
 635  59  67 LEU CD1  C  22.950 0.400 1
 636  59  67 LEU N    N 120.200 0.400 1
 637  60  68 PRO HA   H   4.373 0.020 1
 638  60  68 PRO HB2  H   2.387 0.020 2
 639  60  68 PRO HB3  H   2.008 0.020 2
 640  60  68 PRO HG2  H   1.747 0.020 1
 641  60  68 PRO HG3  H   1.747 0.020 1
 642  60  68 PRO HD2  H   3.273 0.020 2
 643  60  68 PRO HD3  H   3.762 0.020 2
 644  60  68 PRO C    C 174.500 0.400 1
 645  60  68 PRO CA   C  59.990 0.400 1
 646  60  68 PRO CB   C  28.600 0.400 1
 647  60  68 PRO CG   C  24.850 0.400 1
 648  60  68 PRO CD   C  47.440 0.400 1
 649  61  69 GLU H    H   8.736 0.020 1
 650  61  69 GLU HA   H   3.655 0.020 1
 651  61  69 GLU HB2  H   1.871 0.020 2
 652  61  69 GLU HB3  H   1.918 0.020 2
 653  61  69 GLU HG2  H   2.160 0.020 1
 654  61  69 GLU HG3  H   2.160 0.020 1
 655  61  69 GLU C    C 175.200 0.400 1
 656  61  69 GLU CA   C  58.120 0.400 1
 657  61  69 GLU CB   C  27.080 0.400 1
 658  61  69 GLU CG   C  33.260 0.400 1
 659  61  69 GLU N    N 124.400 0.400 1
 660  62  70 VAL H    H   8.448 0.020 1
 661  62  70 VAL HA   H   3.687 0.020 1
 662  62  70 VAL HB   H   1.944 0.020 1
 663  62  70 VAL HG1  H   0.830 0.020 2
 664  62  70 VAL HG2  H   0.880 0.020 2
 665  62  70 VAL C    C 174.200 0.400 1
 666  62  70 VAL CA   C  62.920 0.400 1
 667  62  70 VAL CB   C  28.290 0.400 1
 668  62  70 VAL CG1  C  19.050 0.400 1
 669  62  70 VAL CG2  C  17.730 0.400 1
 670  62  70 VAL N    N 116.300 0.400 1
 671  63  71 LEU H    H   6.729 0.020 1
 672  63  71 LEU HA   H   4.175 0.020 1
 673  63  71 LEU HB2  H   1.510 0.020 2
 674  63  71 LEU HB3  H   1.760 0.020 2
 675  63  71 LEU HG   H   1.400 0.020 1
 676  63  71 LEU HD1  H   0.830 0.020 2
 677  63  71 LEU HD2  H   0.620 0.020 2
 678  63  71 LEU C    C 175.700 0.400 1
 679  63  71 LEU CA   C  53.690 0.400 1
 680  63  71 LEU CB   C  39.770 0.400 1
 681  63  71 LEU CG   C  27.120 0.400 1
 682  63  71 LEU CD1  C  22.680 0.400 1
```

```
142  13  21 VAL CG2  C  16.500 0.400 1
143  13  21 VAL N    N 108.300 0.400 1
144  14  22 THR H    H   7.938 0.020 1
145  14  22 THR HA   H   4.069 0.020 1
146  14  22 THR HB   H   3.702 0.020 1
147  14  22 THR HG2  H   1.075 0.020 1
148  14  22 THR C    C 173.200 0.400 1
149  14  22 THR CA   C  65.210 0.400 1
150  14  22 THR CB   C  66.110 0.400 1
151  14  22 THR CG2  C  18.260 0.400 1
152  14  22 THR N    N 118.600 0.400 1
153  15  23 GLY H    H   8.301 0.020 1
154  15  23 GLY HA2  H   3.708 0.020 1
155  15  23 GLY HA3  H   3.708 0.020 1
156  15  23 GLY C    C 175.500 0.400 1
157  15  23 GLY CA   C  41.890 0.400 1
158  15  23 GLY N    N 107.300 0.400 1
159  16  24 PHE H    H   7.114 0.020 1
160  16  24 PHE HA   H   3.974 0.020 1
161  16  24 PHE HB2  H   2.632 0.020 2
162  16  24 PHE HB3  H   2.831 0.020 2
163  16  24 PHE HD1  H   6.488 0.020 1
164  16  24 PHE HD2  H   6.488 0.020 1
165  16  24 PHE HE1  H   6.880 0.020 1
166  16  24 PHE HE2  H   6.880 0.020 1
167  16  24 PHE C    C 174.600 0.400 1
168  16  24 PHE CA   C  58.320 0.400 1
169  16  24 PHE CB   C  35.840 0.400 1
170  17  25 ARG H    H   7.828 0.020 1
171  17  25 ARG HA   H   4.157 0.020 1
172  17  25 ARG HB2  H   1.832 0.020 2
173  17  25 ARG HB3  H   1.847 0.020 2
174  17  25 ARG HG2  H   1.627 0.020 2
175  17  25 ARG HG3  H   1.486 0.020 2
176  17  25 ARG HD2  H   3.127 0.020 1
177  17  25 ARG HD3  H   3.127 0.020 1
178  17  25 ARG HE   H   8.081 0.020 1
179  17  25 ARG HH21 H   7.242 0.020 1
180  17  25 ARG HH22 H   7.242 0.020 1
181  17  25 ARG C    C 177.800 0.400 1
182  17  25 ARG CA   C  56.650 0.400 1
183  17  25 ARG CB   C  27.170 0.400 1
184  17  25 ARG CG   C  25.62  0.400 1
185  17  25 ARG N    N 120.200 0.400 1
186  17  25 ARG NH2  N  83.28  0.400 1
187  18  26 LEU H    H   7.999 0.020 1
188  18  26 LEU HA   H   3.987 0.020 1
189  18  26 LEU HB2  H   1.455 0.020 2
190  18  26 LEU HB3  H   1.620 0.020 2
191  18  26 LEU HG   H   1.474 0.020 1
192  18  26 LEU HD1  H   0.718 0.020 2
193  18  26 LEU HD2  H   0.718 0.020 2
194  18  26 LEU C    C 175.000 0.400 1
195  18  26 LEU CA   C  54.230 0.400 1
196  18  26 LEU CB   C  38.770 0.400 1
197  18  26 LEU CG   C  24.030 0.400 1
198  18  26 LEU CD1  C  22.410 0.400 1
199  18  26 LEU CD2  C  22.250 0.400 1
200  18  26 LEU N    N 119.700 0.400 1
201  19  27 ALA H    H   7.140 0.020 1
202  19  27 ALA HA   H   4.175 0.020 1
203  19  27 ALA HB   H   1.228 0.020 1
204  19  27 ALA C    C 174.400 0.400 1
205  19  27 ALA CA   C  49.530 0.400 1
206  19  27 ALA CB   C  15.960 0.400 1
207  19  27 ALA N    N 119.400 0.400 1
208  20  28 GLY H    H   7.629 0.020 1
209  20  28 GLY HA2  H   3.592 0.020 2
210  20  28 GLY HA3  H   4.167 0.020 2
211  20  28 GLY C    C 171.700 0.400 1
212  20  28 GLY CA   C  42.700 0.400 1
213  20  28 GLY N    N 104.700 0.400 1
214  21  29 ILE H    H   7.449 0.020 1
215  21  29 ILE HA   H   3.974 0.020 1
216  21  29 ILE HB   H   1.558 0.020 1
683  63  71 LEU CD2  C  21.280 0.400 1
684  63  71 LEU N    N 120.100 0.400 1
685  64  72 ARG H    H   8.076 0.020 1
686  64  72 ARG HA   H   3.624 0.020 1
687  64  72 ARG HB2  H   1.729 0.020 2
688  64  72 ARG HB3  H   1.700 0.020 2
689  64  72 ARG HG2  H   1.516 0.020 2
690  64  72 ARG HG3  H   1.386 0.020 2
691  64  72 ARG HD2  H   3.102 0.020 2
692  64  72 ARG HD3  H   3.283 0.020 2
693  64  72 ARG HH11 H   6.524 0.020 1
694  64  72 ARG HH12 H   6.524 0.020 1
695  64  72 ARG HH21 H   6.970 0.020 1
696  64  72 ARG HH22 H   6.970 0.020 1
697  64  72 ARG C    C 175.500 0.400 1
698  64  72 ARG CA   C  57.770 0.400 1
699  64  72 ARG CB   C  27.540 0.400 1
700  64  72 ARG CG   C  25.440 0.400 1
701  64  72 ARG CD   C  41.260 0.400 1
702  64  72 ARG N    N 119.100 0.400 1
703  64  72 ARG NH2  N  83.04  0.400 1
704  65  73 LYS H    H   8.100 0.020 1
705  65  73 LYS HA   H   3.897 0.020 1
706  65  73 LYS HB2  H   1.700 0.020 2
707  65  73 LYS HB3  H   1.756 0.020 2
708  65  73 LYS HG2  H   1.198 0.020 2
709  65  73 LYS HG3  H   1.350 0.020 2
710  65  73 LYS HD2  H   1.540 0.020 2
711  65  73 LYS HE2  H   2.736 0.020 2
712  65  73 LYS HE3  H   2.814 0.020 2
713  65  73 LYS HZ   H   7.443 0.020 1
714  65  73 LYS C    C 175.500 0.400 1
715  65  73 LYS CA   C  57.190 0.400 1
716  65  73 LYS CB   C  29.310 0.400 1
717  65  73 LYS CG   C  22.050 0.400 1
718  65  73 LYS CD   C  26.700 0.400 1
719  65  73 LYS CE   C  39.510 0.400 1
720  65  73 LYS N    N 119.100 0.400 1
721  66  74 ASN H    H   7.614 0.020 1
722  66  74 ASN HA   H   4.323 0.020 1
723  66  74 ASN HB2  H   2.706 0.020 2
724  66  74 ASN HB3  H   2.761 0.020 2
725  66  74 ASN HD21 H   6.752 0.020 2
726  66  74 ASN HD22 H   7.085 0.020 2
727  66  74 ASN C    C 175.800 0.400 1
728  66  74 ASN CA   C  53.340 0.400 1
729  66  74 ASN CB   C  35.730 0.400 1
730  66  74 ASN N    N 116.800 0.400 1
731  66  74 ASN ND2  N 109.100 0.400 1
732  67  75 LEU H    H   8.317 0.020 1
733  67  75 LEU HA   H   3.863 0.020 1
734  67  75 LEU HB2  H   1.681 0.020 2
735  67  75 LEU HB3  H   1.752 0.020 2
736  67  75 LEU HG   H   1.385 0.020 1
737  67  75 LEU HD1  H   0.657 0.020 2
738  67  75 LEU HD2  H   0.840 0.020 2
739  67  75 LEU C    C 175.400 0.400 1
740  67  75 LEU CA   C  55.230 0.400 1
741  67  75 LEU CB   C  38.970 0.400 1
742  67  75 LEU CG   C  23.820 0.400 1
743  67  75 LEU CD1  C  21.310 0.400 1
744  67  75 LEU CD2  C  20.980 0.400 1
745  67  75 LEU N    N 119.200 0.400 1
746  68  76 ASN H    H   7.976 0.020 1
747  68  76 ASN HA   H   4.364 0.020 1
748  68  76 ASN HB2  H   2.724 0.020 2
749  68  76 ASN HB3  H   2.825 0.020 2
750  68  76 ASN HD21 H   7.436 0.020 2
751  68  76 ASN HD22 H   6.595 0.020 2
752  68  76 ASN C    C 173.900 0.400 1
753  68  76 ASN CA   C  52.920 0.400 1
754  68  76 ASN CB   C  36.180 0.400 1
755  68  76 ASN N    N 116.700 0.400 1
756  68  76 ASN ND2  N 110.800 0.400 1
757  69  77 GLU H    H   7.590 0.020 1
```

```
217  21  29 ILE HG12 H   1.363 0.020 2
218  21  29 ILE HG13 H   1.557 0.020 2
219  21  29 ILE HG2  H   1.259 0.020 1
220  21  29 ILE HD1  H   0.631 0.020 1
221  21  29 ILE C    C 172.800 0.400 1
222  21  29 ILE CA   C  57.760 0.400 1
223  21  29 ILE CB   C  33.700 0.400 1
224  21  29 ILE CG1  C  30.690 0.400 1
225  21  29 ILE CG2  C  15.330 0.400 1
226  21  29 ILE CD1  C   9.573 0.400 1
227  21  29 ILE N    N 121.200 0.400 1
228  22  30 SER H    H   8.182 0.020 1
229  22  30 SER HA   H   4.352 0.020 1
230  22  30 SER HB2  H   3.780 0.020 2
231  22  30 SER HB3  H   3.726 0.020 2
232  22  30 SER C    C 173.400 0.400 1
233  22  30 SER CA   C  57.410 0.400 1
234  22  30 SER CB   C  61.400 0.400 1
235  22  30 SER N    N 121.500 0.400 1
236  23  31 LYS H    H   8.266 0.020 1
237  23  31 LYS HA   H   4.127 0.020 1
238  23  31 LYS HB2  H   1.758 0.020 2
239  23  31 LYS HB3  H   2.050 0.020 2
240  23  31 LYS HG2  H   1.110 0.020 1
241  23  31 LYS HG3  H   1.110 0.020 1
242  23  31 LYS HD2  H   1.286 0.020 1
243  23  31 LYS HD3  H   1.286 0.020 1
244  23  31 LYS HE2  H   2.840 0.020 2
245  23  31 LYS HE3  H   3.107 0.020 2
246  23  31 LYS HZ   H   7.601 0.020 1
247  23  31 LYS C    C 170.200 0.400 1
248  23  31 LYS CA   C  53.470 0.400 1
249  23  31 LYS CB   C  27.810 0.400 1
250  23  31 LYS CG   C  22.18  0.400 1
251  23  31 LYS CD   C  28.33  0.400 1
252  23  31 LYS CE   C  39.380 0.400 1
253  23  31 LYS N    N 125.900 0.400 1
254  24  32 VAL H    H   6.997 0.020 1
255  24  32 VAL HA   H   4.379 0.020 1
256  24  32 VAL HB   H   1.629 0.020 1
257  24  32 VAL HG1  H   0.504 0.020 2
258  24  32 VAL HG2  H   0.550 0.020 2
259  24  32 VAL C    C 172.200 0.400 1
260  24  32 VAL CA   C  57.340 0.400 1
261  24  32 VAL CB   C  31.070 0.400 1
262  24  32 VAL CG1  C  18.910 0.400 1
263  24  32 VAL CG2  C  18.690 0.400 1
264  24  32 VAL N    N 120.500 0.400 1
265  25  33 TYR H    H   8.934 0.020 1
266  25  33 TYR HA   H   4.582 0.020 1
267  25  33 TYR HB2  H   2.930 0.020 2
268  25  33 TYR HB3  H   2.740 0.020 2
269  25  33 TYR HD1  H   6.530 0.020 1
270  25  33 TYR HD2  H   6.530 0.020 1
271  25  33 TYR HE1  H   6.930 0.020 1
272  25  33 TYR HE2  H   6.930 0.020 1
273  25  33 TYR C    C 170.400 0.400 1
274  25  33 TYR CA   C  53.730 0.400 1
275  25  33 TYR CB   C  35.880 0.400 1
276  25  33 TYR N    N 128.400 0.400 1
277  26  34 GLU H    H   8.485 0.020 1
278  26  34 GLU HA   H   5.111 0.020 1
279  26  34 GLU HB2  H   1.700 0.020 1
280  26  34 GLU HB3  H   1.700 0.020 1
281  26  34 GLU HG2  H   1.942 0.020 2
282  26  34 GLU HG3  H   1.981 0.020 2
283  26  34 GLU C    C 174.500 0.400 1
284  26  34 GLU CA   C  52.270 0.400 1
285  26  34 GLU CB   C  27.810 0.400 1
286  26  34 GLU CG   C  33.730 0.400 1
287  26  34 GLU N    N 125.500 0.400 1
288  27  35 THR H    H   8.791 0.020 1
289  27  35 THR HA   H   4.603 0.020 1
290  27  35 THR HB   H   3.955 0.020 1
291  27  35 THR HG2  H   0.798 0.020 1
758  69  77 GLU HA   H   4.210 0.020 1
759  69  77 GLU HB2  H   1.840 0.020 2
760  69  77 GLU HB3  H   2.023 0.020 2
761  69  77 GLU HG2  H   2.180 0.020 2
762  69  77 GLU HG3  H   2.367 0.020 2
763  69  77 GLU C    C 173.700 0.400 1
764  69  77 GLU CA   C  53.160 0.400 1
765  69  77 GLU CB   C  27.730 0.400 1
766  69  77 GLU CG   C  33.710 0.400 1
767  69  77 GLU N    N 115.100 0.400 1
768  70  78 SER H    H   7.110 0.020 1
769  70  78 SER HB2  H   3.880 0.020 1
770  70  78 SER HB3  H   3.880 0.020 1
771  70  78 SER C    C 172.700 0.400 1
772  70  78 SER CA   C  57.020 0.400 1
773  70  78 SER CB   C  60.660 0.400 1
774  70  78 SER N    N 114.900 0.400 1
775  71  79 VAL H    H   8.515 0.020 1
776  71  79 VAL HA   H   4.150 0.020 1
777  71  79 VAL HB   H   1.970 0.020 1
778  71  79 VAL HG1  H   0.840 0.020 2
779  71  79 VAL HG2  H   0.840 0.020 2
780  71  79 VAL C    C 173.100 0.400 1
781  71  79 VAL CA   C  59.950 0.400 1
782  71  79 VAL CB   C  29.910 0.400 1
783  71  79 VAL CG1  C  18.490 0.400 1
784  71  79 VAL CG2  C  17.730 0.400 1
785  71  79 VAL N    N 125.300 0.400 1
786  72  80 GLN H    H   7.486 0.020 1
787  72  80 GLN HA   H   4.505 0.020 1
788  72  80 GLN HB2  H   1.860 0.020 2
789  72  80 GLN HB3  H   1.972 0.020 2
790  72  80 GLN HG2  H   2.162 0.020 2
791  72  80 GLN HG3  H   2.107 0.020 2
792  72  80 GLN HE21 H   7.582 0.020 2
793  72  80 GLN HE22 H   6.835 0.020 2
794  72  80 GLN C    C 169.300 0.400 1
795  72  80 GLN CA   C  50.750 0.400 1
796  72  80 GLN CB   C  27.080 0.400 1
797  72  80 GLN CG   C  31.180 0.400 1
798  72  80 GLN N    N 121.200 0.400 1
799  72  80 GLN NE2  N 113.200 0.400 1
800  73  81 PRO HA   H   4.535 0.020 1
801  73  81 PRO HB2  H   1.395 0.020 2
802  73  81 PRO HB3  H   1.318 0.020 2
803  73  81 PRO HG2  H   1.924 0.020 2
804  73  81 PRO HG3  H   1.458 0.020 2
805  73  81 PRO HD2  H   3.461 0.020 2
806  73  81 PRO HD3  H   3.561 0.020 2
807  73  81 PRO C    C 173.300 0.400 1
808  73  81 PRO CA   C  60.050 0.400 1
809  73  81 PRO CB   C  31.530 0.400 1
810  73  81 PRO CG   C  26.59  0.400 1
811  73  81 PRO CD   C  47.280 0.400 1
812  74  82 THR H    H   7.944 0.020 1
813  74  82 THR HA   H   4.191 0.020 1
814  74  82 THR HB   H   3.924 0.020 1
815  74  82 THR HG2  H   1.051 0.020 1
816  74  82 THR C    C 170.400 0.400 1
817  74  82 THR CA   C  60.240 0.400 1
818  74  82 THR CB   C  68.140 0.400 1
819  74  82 THR CG2  C  18.870 0.400 1
820  74  82 THR N    N 116.900 0.400 1
821  75  83 VAL H    H   8.812 0.020 1
822  75  83 VAL HA   H   4.755 0.020 1
823  75  83 VAL HB   H   1.845 0.020 1
824  75  83 VAL HG1  H   0.661 0.020 2
825  75  83 VAL HG2  H   0.818 0.020 2
826  75  83 VAL C    C 172.600 0.400 1
827  75  83 VAL CA   C  58.970 0.400 1
828  75  83 VAL CB   C  29.480 0.400 1
829  75  83 VAL CG1  C  20.890 0.400 1
830  75  83 VAL CG2  C  17.460 0.400 1
831  75  83 VAL N    N 129.100 0.400 1
832  76  84 VAL H    H   9.166 0.020 1
```

```
292  27  35 THR C    C 171.900 0.400 1
293  27  35 THR CA   C  62.750 0.400 1
294  27  35 THR CB   C  67.010 0.400 1
295  27  35 THR CG2  C  24.570 0.400 1
296  27  35 THR N    N 119.100 0.400 1
297  28  36 PRO HA   H   4.408 0.020 1
298  28  36 PRO HB2  H   1.975 0.020 2
299  28  36 PRO HB3  H   2.306 0.020 2
300  28  36 PRO HG2  H   1.871 0.020 1
301  28  36 PRO HG3  H   1.871 0.020 1
302  28  36 PRO HD2  H   3.537 0.020 2
303  28  36 PRO HD3  H   3.614 0.020 2
304  28  36 PRO C    C 172.400 0.400 1
305  28  36 PRO CA   C  61.070 0.400 1
306  28  36 PRO CB   C  29.830 0.400 1
307  28  36 PRO CG   C  23.910 0.400 1
308  28  36 PRO CD   C  48.670 0.400 1
309  29  37 ASP H    H   7.086 0.020 1
310  29  37 ASP HA   H   4.679 0.020 1
311  29  37 ASP HB2  H   2.715 0.020 2
312  29  37 ASP HB3  H   2.979 0.020 2
313  29  37 ASP C    C 173.400 0.400 1
314  29  37 ASP CA   C  49.210 0.400 1
315  29  37 ASP CB   C  39.950 0.400 1
316  29  37 ASP N    N 114.800 0.400 1
317  30  38 ILE H    H   8.643 0.020 1
318  30  38 ILE HA   H   3.799 0.020 1
319  30  38 ILE HB   H   2.017 0.020 1
320  30  38 ILE HG12 H   0.854 0.020 2
321  30  38 ILE HG13 H   1.102 0.020 2
322  30  38 ILE HG2  H   1.600 0.020 1
323  30  38 ILE HD1  H   0.790 0.020 1
324  30  38 ILE C    C 172.800 0.400 1
325  30  38 ILE CA   C  63.710 0.400 1
326  30  38 ILE CB   C  32.630 0.400 1
327  30  38 ILE CG1  C  26.620 0.400 1
328  30  38 ILE CG2  C  14.800 0.400 1
329  30  38 ILE CD1  C   9.635 0.400 1
330  30  38 ILE N    N 120.600 0.400 1
331  31  39 PRO HA   H   4.324 0.020 1
332  31  39 PRO HB2  H   2.040 0.020 2
333  31  39 PRO HB3  H   2.219 0.020 2
334  31  39 PRO HG2  H   1.760 0.020 2
335  31  39 PRO HG3  H   1.821 0.020 2
336  31  39 PRO HD2  H   3.652 0.020 2
337  31  39 PRO HD3  H   3.670 0.020 2
338  31  39 PRO C    C 177.100 0.400 1
339  31  39 PRO CA   C  63.690 0.400 1
340  31  39 PRO CB   C  27.770 0.400 1
341  31  39 PRO CG   C  25.170 0.400 1
342  31  39 PRO CD   C  52.870 0.400 1
343  32  40 ALA H    H   8.182 0.020 1
344  32  40 ALA HA   H   4.160 0.020 1
345  32  40 ALA HB   H   1.485 0.020 1
346  32  40 ALA C    C 178.000 0.400 1
347  32  40 ALA CA   C  52.170 0.400 1
348  32  40 ALA CB   C  15.820 0.400 1
349  32  40 ALA N    N 118.900 0.400 1
350  33  41 THR H    H   8.169 0.020 1
351  33  41 THR HA   H   4.342 0.020 1
352  33  41 THR HB   H   3.747 0.020 1
353  33  41 THR HG2  H   0.996 0.020 1
354  33  41 THR C    C 172.500 0.400 1
355  33  41 THR CA   C  64.710 0.400 1
356  33  41 THR CB   C  65.240 0.400 1
357  33  41 THR CG2  C  18.770 0.400 1
358  33  41 THR N    N 118.400 0.400 1
359  34  42 GLU H    H   8.467 0.020 1
360  34  42 GLU HA   H   3.462 0.020 1
361  34  42 GLU HB2  H   2.130 0.020 1
362  34  42 GLU HB3  H   2.130 0.020 1
363  34  42 GLU HG2  H   1.966 0.020 1
364  34  42 GLU HG3  H   1.966 0.020 1
365  34  42 GLU C    C 175.200 0.400 1
366  34  42 GLU CA   C  57.750 0.400 1
833  76  84 VAL HA   H   4.023 0.020 1
834  76  84 VAL HB   H   1.802 0.020 1
835  76  84 VAL HG1  H   0.700 0.020 2
836  76  84 VAL HG2  H   0.710 0.020 2
837  76  84 VAL C    C 169.600 0.400 1
838  76  84 VAL CA   C  58.750 0.400 1
839  76  84 VAL CB   C  30.790 0.400 1
840  76  84 VAL CG1  C  19.870 0.400 1
841  76  84 VAL CG2  C  18.590 0.400 1
842  76  84 VAL N    N 130.500 0.400 1
843  77  85 ALA H    H   8.203 0.020 1
844  77  85 ALA HA   H   4.978 0.020 1
845  77  85 ALA HB   H   1.029 0.020 1
846  77  85 ALA C    C 174.500 0.400 1
847  77  85 ALA CA   C  46.810 0.400 1
848  77  85 ALA CB   C  17.720 0.400 1
849  77  85 ALA N    N 128.700 0.400 1
850  78  86 LEU H    H   8.877 0.020 1
851  78  86 LEU HA   H   4.583 0.020 1
852  78  86 LEU HB2  H   1.691 0.020 2
853  78  86 LEU HB3  H   1.689 0.020 2
854  78  86 LEU HD1  H   0.790 0.020 2
855  78  86 LEU HD2  H   0.958 0.020 2
856  78  86 LEU C    C 173.500 0.400 1
857  78  86 LEU CA   C  50.780 0.400 1
858  78  86 LEU CB   C  38.950 0.400 1
859  78  86 LEU CD1  C  23.650 0.400 1
860  78  86 LEU CD2  C  20.890 0.400 1
861  78  86 LEU N    N 123.400 0.400 1
862  79  87 GLY H    H   8.250 0.020 1
863  79  87 GLY HA2  H   3.650 0.020 2
864  79  87 GLY HA3  H   3.790 0.020 2
865  79  87 GLY C    C 172.000 0.400 1
866  79  87 GLY CA   C  41.750 0.400 1
867  79  87 GLY N    N 106.700 0.400 1
868  80  88 GLY H    H   8.227 0.020 1
869  80  88 GLY HA2  H   3.994 0.020 2
870  80  88 GLY HA3  H   3.990 0.020 2
871  80  88 GLY C    C 171.400 0.400 1
872  80  88 GLY CA   C  42.220 0.400 1
873  80  88 GLY N    N 107.600 0.400 1
874  81  89 SER H    H   8.460 0.020 1
875  81  89 SER HA   H   4.370 0.020 1
876  81  89 SER HB2  H   3.995 0.020 2
877  81  89 SER HB3  H   3.907 0.020 2
878  81  89 SER C    C 172.600 0.400 1
879  81  89 SER CA   C  55.880 0.400 1
880  81  89 SER CB   C  61.190 0.400 1
881  81  89 SER N    N 115.400 0.400 1
882  82  90 GLY H    H   8.515 0.020 1
883  82  90 GLY HA2  H   4.680 0.020 1
884  82  90 GLY HA3  H   4.680 0.020 1
885  82  90 GLY C    C 171.600 0.400 1
886  82  90 GLY CA   C  42.800 0.400 1
887  82  90 GLY N    N 110.900 0.400 1
888  83  91 SER H    H   8.161 0.020 1
889  83  91 SER HA   H   4.670 0.020 1
890  83  91 SER HB2  H   3.850 0.020 1
891  83  91 SER HB3  H   3.850 0.020 1
892  83  91 SER CA   C  55.930 0.400 1
893  83  91 SER CB   C  60.920 0.400 1
894  83  91 SER N    N 115.500 0.400 1
895  85  93 SER HA   H   4.318 0.020 1
896  85  93 SER HB2  H   3.937 0.020 1
897  85  93 SER HB3  H   3.937 0.020 1
898  85  93 SER CA   C  55.880 0.400 1
899  85  93 SER CB   C  61.140 0.400 1
900  86  94 THR H    H   8.172 0.020 1
901  86  94 THR HA   H   4.027 0.020 1
902  86  94 THR HB   H   3.807 0.020 1
903  86  94 THR CA   C  59.470 0.400 1
904  86  94 THR CB   C  66.820 0.400 1
905  86  94 THR N    N 115.500 0.400 1
906  87  95 SER H    H   8.172 0.020 1
907  87  95 SER HA   H   4.320 0.020 1
```

```
367  34  42 GLU CB   C  26.780 0.400 1
368  34  42 GLU CG   C  33.970 0.400 1
369  34  42 GLU N    N 121.400 0.400 1
370  35  43 SER H    H   7.806 0.020 1
371  35  43 SER HA   H   3.991 0.020 1
372  35  43 SER HB2  H   3.831 0.020 1
373  35  43 SER HB3  H   3.831 0.020 1
374  35  43 SER C    C 174.200 0.400 1
375  35  43 SER CA   C  57.050 0.400 1
376  35  43 SER CB   C  60.580 0.400 1
377  35  43 SER N    N 112.100 0.400 1
378  36  44 ALA H    H   7.814 0.020 1
379  36  44 ALA HA   H   3.807 0.020 1
380  36  44 ALA HB   H   1.182 0.020 1
381  36  44 ALA C    C 176.000 0.400 1
382  36  44 ALA CA   C  52.390 0.400 1
383  36  44 ALA CB   C  16.460 0.400 1
384  36  44 ALA N    N 123.700 0.400 1
385  37  45 VAL H    H   8.314 0.020 1
386  37  45 VAL HA   H   3.078 0.020 1
387  37  45 VAL HB   H   2.065 0.020 1
388  37  45 VAL HG1  H   0.693 0.020 2
389  37  45 VAL HG2  H   0.979 0.020 2
390  37  45 VAL C    C 174.400 0.400 1
391  37  45 VAL CA   C  64.170 0.400 1
392  37  45 VAL CB   C  28.950 0.400 1
393  37  45 VAL CG1  C  21.780 0.400 1
394  37  45 VAL CG2  C  20.040 0.400 1
395  37  45 VAL N    N 118.400 0.400 1
396  38  46 ARG H    H   7.843 0.020 1
397  38  46 ARG HA   H   3.683 0.020 1
398  38  46 ARG HB2  H   1.764 0.020 2
399  38  46 ARG HB3  H   1.755 0.020 2
400  38  46 ARG HG2  H   1.456 0.020 2
401  38  46 ARG HG3  H   1.450 0.020 2
402  38  46 ARG HD2  H   3.084 0.020 2
403  38  46 ARG HD3  H   3.092 0.020 2
404  38  46 ARG HH11 H   7.609 0.020 1
405  38  46 ARG HH12 H   7.609 0.020 1
406  38  46 ARG C    C 176.200 0.400 1
407  38  46 ARG CA   C  57.790 0.400 1
408  38  46 ARG CB   C  26.640 0.400 1
409  38  46 ARG CG   C  25.990 0.400 1
410  38  46 ARG CD   C  40.390 0.400 1
411  38  46 ARG N    N 116.100 0.400 1
412  38  46 ARG NH2  N  85.28  0.400 1
413  39  47 SER H    H   7.607 0.020 1
414  39  47 SER HA   H   4.046 0.020 1
415  39  47 SER HB2  H   3.767 0.020 2
416  39  47 SER HB3  H   3.760 0.020 2
417  39  47 SER C    C 174.500 0.400 1
418  39  47 SER CA   C  59.020 0.400 1
419  39  47 SER CB   C  60.310 0.400 1
420  39  47 SER N    N 112.900 0.400 1
421  40  48 VAL H    H   8.244 0.020 1
422  40  48 VAL HA   H   3.660 0.020 1
423  40  48 VAL HB   H   1.808 0.020 1
424  40  48 VAL HG1  H   0.806 0.020 2
425  40  48 VAL HG2  H   0.718 0.020 2
426  40  48 VAL C    C 175.300 0.400 1
427  40  48 VAL CA   C  63.510 0.400 1
428  40  48 VAL CB   C  28.150 0.400 1
429  40  48 VAL CG1  C  21.720 0.400 1
430  40  48 VAL CG2  C  20.840 0.400 1
431  40  48 VAL N    N 123.100 0.400 1
432  41  49 LEU H    H   8.071 0.020 1
433  41  49 LEU HA   H   3.785 0.020 1
434  41  49 LEU HB2  H   1.732 0.020 2
435  41  49 LEU HB3  H   1.695 0.020 2
436  41  49 LEU HG   H   1.378 0.020 1
437  41  49 LEU HD1  H   0.614 0.020 2
438  41  49 LEU HD2  H   0.614 0.020 2
439  41  49 LEU C    C 175.700 0.400 1
440  41  49 LEU CA   C  54.630 0.400 1
441  41  49 LEU CB   C  38.300 0.400 1
908  87  95 SER HB2  H   3.930 0.020 1
909  87  95 SER HB3  H   3.930 0.020 1
910  87  95 SER C    C 171.700 0.400 1
911  87  95 SER CA   C  55.890 0.400 1
912  87  95 SER CB   C  61.000 0.400 1
913  87  95 SER N    N 123.600 0.400 1
914  88  96 LEU H    H   8.135 0.020 1
915  88  96 LEU HA   H   4.660 0.020 1
916  88  96 LEU HB2  H   1.520 0.020 1
917  88  96 LEU HB3  H   1.520 0.020 1
918  88  96 LEU C    C 174.100 0.400 1
919  88  96 LEU CA   C  52.030 0.400 1
920  88  96 LEU CB   C  39.360 0.400 1
921  88  96 LEU CG   C  23.990 0.400 1
922  88  96 LEU CD1  C  22.220 0.400 1
923  88  96 LEU CD2  C  20.580 0.400 1
924  88  96 LEU N    N 123.700 0.400 1
925  89  97 ARG H    H   7.920 0.020 1
926  89  97 ARG C    C 172.200 0.400 1
927  89  97 ARG CA   C  53.860 0.400 1
928  89  97 ARG CB   C  26.540 0.400 1
929  89  97 ARG N    N 120.700 0.400 1
930  90  98 GLU HA   H   4.089 0.020 1
931  90  98 GLU HB2  H   1.915 0.020 2
932  90  98 GLU HB3  H   1.910 0.020 2
933  90  98 GLU HG2  H   2.128 0.020 2
934  90  98 GLU HG3  H   1.968 0.020 2
935  90  98 GLU CA   C  56.450 0.400 1
936  90  98 GLU CB   C  27.280 0.400 1
937  90  98 GLU CG   C  33.320 0.400 1
938  91  99 LYS H    H   7.848 0.020 1
939  91  99 LYS HA   H   4.040 0.020 1
940  91  99 LYS HB2  H   1.740 0.020 1
941  91  99 LYS HB3  H   1.740 0.020 1
942  91  99 LYS HE2  H   3.040 0.020 2
943  91  99 LYS HE3  H   3.048 0.020 2
944  91  99 LYS C    C 176.300 0.400 1
945  91  99 LYS CA   C  54.240 0.400 1
946  91  99 LYS CB   C  29.730 0.400 1
947  91  99 LYS N    N 120.400 0.400 1
948  92 100 ILE H    H   7.660 0.020 1
949  92 100 ILE HA   H   3.724 0.020 1
950  92 100 ILE HB   H   1.829 0.020 1
951  92 100 ILE HG2  H   1.380 0.020 1
952  92 100 ILE HD1  H   0.690 0.020 1
953  92 100 ILE C    C 173.600 0.400 1
954  92 100 ILE CA   C  59.600 0.400 1
955  92 100 ILE CB   C  39.520 0.400 1
956  92 100 ILE N    N 113.600 0.400 1
957  93 101 LYS H    H   8.048 0.020 1
958  93 101 LYS HA   H   3.713 0.020 1
959  93 101 LYS HE2  H   2.980 0.020 1
960  93 101 LYS HE3  H   2.980 0.020 1
961  93 101 LYS C    C 175.100 0.400 1
962  93 101 LYS CA   C  58.43  0.400 1
963  93 101 LYS CB   C  27.740 0.400 1
964  93 101 LYS N    N 122.700 0.400 1
965  94 102 GLN H    H   8.330 0.020 1
966  94 102 GLN HA   H   3.890 0.020 1
967  94 102 GLN HG2  H   2.136 0.020 2
968  94 102 GLN HG3  H   2.448 0.020 2
969  94 102 GLN HE21 H   7.500 0.020 2
970  94 102 GLN HE22 H   6.720 0.020 2
971  94 102 GLN CA   C  55.640 0.400 1
972  94 102 GLN CB   C  25.950 0.400 1
973  94 102 GLN N    N 119.300 0.400 1
974  94 102 GLN NE2  N 111.600 0.400 1
975  95 103 ALA H    H   7.886 0.020 1
976  95 103 ALA HA   H   4.172 0.020 1
977  95 103 ALA CA   C  52.040 0.400 1
978  95 103 ALA CB   C  16.970 0.400 1
979  95 103 ALA N    N 119.800 0.400 1
980  96 104 VAL HA   H   4.271 0.020 1
981  96 104 VAL HB   H   2.215 0.020 1
982  96 104 VAL HG1  H   0.853 0.020 2
```

```
 442  41  49 LEU CG   C  24.110 0.400 1
 443  41  49 LEU CD1  C  22.560 0.400 1
 444  41  49 LEU CD2  C  19.100 0.400 1
 445  41  49 LEU N    N 117.200 0.400 1
 446  42  50 GLU H    H   6.986 0.020 1
 447  42  50 GLU HA   H   4.213 0.020 1
 448  42  50 GLU HB2  H   1.843 0.020 2
 449  42  50 GLU HB3  H   1.840 0.020 2
 450  42  50 GLU HG2  H   2.318 0.020 2
 451  42  50 GLU HG3  H   2.159 0.020 2
 452  42  50 GLU C    C 172.600 0.400 1
 453  42  50 GLU CA   C  53.240 0.400 1
 454  42  50 GLU CB   C  27.590 0.400 1
 455  42  50 GLU CG   C  33.420 0.400 1
 456  42  50 GLU N    N 114.700 0.400 1
 490  46  54 VAL H    H   7.214 0.020 1
 491  46  54 VAL HA   H   3.755 0.020 1
 492  46  54 VAL HB   H   1.989 0.020 1
 493  46  54 VAL HG1  H   0.655 0.020 2
 494  46  54 VAL HG2  H   0.740 0.020 2
 495  46  54 VAL C    C 172.500 0.400 1
 496  46  54 VAL CA   C  60.450 0.400 1
 497  46  54 VAL CB   C  29.730 0.400 1
 498  46  54 VAL CG1  C  19.900 0.400 1
 499  46  54 VAL CG2  C  19.640 0.400 1
 500  46  54 VAL N    N 122.100 0.400 1
 501  47  55 GLY H    H   9.313 0.020 1
 502  47  55 GLY HA2  H   3.767 0.020 2
 503  47  55 GLY HA3  H   4.330 0.020 2
 504  47  55 GLY C    C 171.500 0.400 1
 505  47  55 GLY CA   C  42.850 0.400 1
 506  47  55 GLY N    N 114.800 0.400 1
 507  48  56 ILE H    H   7.564 0.020 1
 508  48  56 ILE HA   H   4.871 0.020 1
 509  48  56 ILE HB   H   1.496 0.020 1
 510  48  56 ILE HG12 H   1.334 0.020 1
 511  48  56 ILE HG13 H   1.334 0.020 1
 512  48  56 ILE HG2  H   0.912 0.020 1
 513  48  56 ILE HD1  H   0.665 0.020 1
 514  48  56 ILE C    C 175.7   0.400 1
 515  48  56 ILE CA   C  56.590 0.400 1
 516  48  56 ILE CB   C  38.570 0.400 1
 517  48  56 ILE CG1  C  25.160 0.400 1
 518  48  56 ILE CG2  C  15.070 0.400 1
 519  48  56 ILE CD1  C  10.830 0.400 1
 520  48  56 ILE N    N 122.100 0.400 1
 521  49  57 LEU H    H   8.706 0.020 1
 522  49  57 LEU HA   H   5.030 0.020 1
 523  49  57 LEU HB2  H   1.462 0.020 2
 524  49  57 LEU HB3  H   1.265 0.020 2
 525  49  57 LEU HG   H   0.827 0.020 1
 526  49  57 LEU HD1  H   0.679 0.020 2
 527  49  57 LEU HD2  H   0.645 0.020 2
 528  49  57 LEU C    C 170.400 0.400 1
 529  49  57 LEU CA   C  49.500 0.400 1
 530  49  57 LEU CB   C  42.130 0.400 1
 531  49  57 LEU CG   C  25.410 0.400 1
 532  49  57 LEU CD1  C  24.610 0.400 1
 533  49  57 LEU CD2  C  21.030 0.400 1
 534  49  57 LEU N    N 128.800 0.400 1
 535  50  58 VAL H    H   9.018 0.020 1
 536  50  58 VAL HA   H   4.906 0.020 1
 537  50  58 VAL HB   H   1.676 0.020 1
 538  50  58 VAL HG1  H   0.695 0.020 2
 539  50  58 VAL HG2  H   0.806 0.020 2
 540  50  58 VAL C    C 171.500 0.400 1
 541  50  58 VAL CA   C  58.380 0.400 1
 983  96 104 VAL HG2  H   0.796 0.020 2
 984  96 104 VAL C    C 174.000 0.400 1
 985  96 104 VAL CA   C  59.020 0.400 1
 986  96 104 VAL CB   C  30.650 0.400 1
 987  96 104 VAL CG1  C  18.820 0.400 1
 988  96 104 VAL CG2  C  18.420 0.400 1
 989  96 104 VAL N    N 110.300 0.400 1
 990  97 105 GLY H    H   7.864 0.020 1
 991  97 105 GLY HA2  H   3.717 0.020 2
 992  97 105 GLY HA3  H   4.048 0.020 2
 993  97 105 GLY C    C 171.200 0.400 1
 994  97 105 GLY CA   C  42.490 0.400 1
 995  97 105 GLY N    N 108.900 0.400 1
 996  98 106 VAL H    H   7.966 0.020 1
 997  98 106 VAL HA   H   4.170 0.020 1
 998  98 106 VAL C    C 172.800 0.400 1
 999  98 106 VAL CA   C  59.360 0.400 1
1000  98 106 VAL CB   C  29.830 0.400 1
1001  98 106 VAL N    N 118.500 0.400 1
1002  99 107 ASP H    H   7.714 0.020 1
1003  99 107 ASP HA   H   4.649 0.020 1
1004  99 107 ASP HB2  H   2.304 0.020 2
1005  99 107 ASP HB3  H   2.614 0.020 2
1006  99 107 ASP C    C 173.100 0.400 1
1007  99 107 ASP CA   C  51.630 0.400 1
1008  99 107 ASP CB   C  38.600 0.400 1
1009  99 107 ASP N    N 122.600 0.400 1
1010 100 108 LEU H    H   7.832 0.020 1
1011 100 108 LEU HA   H   4.022 0.020 1
1012 100 108 LEU HB2  H   0.994 0.020 2
1013 100 108 LEU HB3  H   1.282 0.020 2
1014 100 108 LEU HG   H   0.874 0.020 1
1015 100 108 LEU HD1  H   0.503 0.020 2
1016 100 108 LEU HD2  H   1.097 0.020 2
1017 100 108 LEU C    C 173.400 0.400 1
1018 100 108 LEU CA   C  52.680 0.400 1
1019 100 108 LEU CB   C  39.240 0.400 1
1020 100 108 LEU CG   C  28.650 0.400 1
1021 100 108 LEU CD1  C  23.080 0.400 1
1022 100 108 LEU CD2  C  21.470 0.400 1
1023 100 108 LEU N    N 122.600 0.400 1
1024 101 109 TRP H    H   7.924 0.020 1
1025 101 109 TRP HA   H   4.534 0.020 1
1026 101 109 TRP HB2  H   2.991 0.020 2
1027 101 109 TRP HB3  H   3.266 0.020 2
1028 101 109 TRP HE1  H   9.985 0.020 1
1029 101 109 TRP HE3  H   7.497 0.020 1
1030 101 109 TRP HZ2  H   7.540 0.020 1
1031 101 109 TRP HZ3  H   7.612 0.020 1
1032 101 109 TRP C    C 171.600 0.400 1
1033 101 109 TRP CA   C  55.370 0.400 1
1034 101 109 TRP CB   C  26.860 0.400 1
1035 101 109 TRP N    N 119.700 0.400 1
1036 101 109 TRP NE1  N 129.100 0.400 1
1037 102 110 LYS H    H   7.295 0.020 1
1038 102 110 LYS HA   H   3.987 0.020 1
1039 102 110 LYS HB2  H   1.645 0.020 2
1040 102 110 LYS HB3  H   1.569 0.020 2
1041 102 110 LYS HG2  H   1.210 0.020 1
1042 102 110 LYS HG3  H   1.210 0.020 1
1043 102 110 LYS C    C 178.200 0.400 1
1044 102 110 LYS CA   C  55.250 0.400 1
1045 102 110 LYS CB   C  30.750 0.400 1
1046 102 110 LYS CE   C  39.570 0.400 1
1047 102110LYS N    N 126.200 0.400 1
```

Appendix B
Chemical shift assignments of H_{1-47} of A_1A_O ATP synthase from *M. jannaschii*

The chemical shift assignment are listed in the NMR-STAR format (Columns are as follows: Atom shift assign ID, Residue author sequence code, Residue sequence code, Residue label, Atom name, Atom type, Chemical shift value, Chemical shift value error, Chemical shift ambiguity code

1	1	1	MET	H	H	8.769	0.020	1
2	1	1	MET	HA	H	4.510	0.400	1
3	1	1	MET	HB2	H	2.075	0.020	2
4	1	1	MET	HB3	H	2.075	0.020	2
5	1	1	MET	HG2	H	2.612	0.020	2
6	1	1	MET	HG3	H	2.612	0.020	2
7	1	1	MET	CA	C	55.680	0.400	1
8	1	1	MET	CB	C	33.310	0.400	1
9	1	1	MET	CG	C	32.450	0.400	1
10	1	1	MET	N	N	120.926	0.400	1
11	2	2	GLY	H	H	8.502	0.020	1
12	2	2	GLY	HA2	H	3.979	0.020	2
13	2	2	GLY	HA3	H	3.979	0.020	2
14	2	2	GLY	CA	C	45.600	0.020	1
15	2	2	GLY	N	N	110.343	0.400	1
16	3	3	VAL	H	H	8.040	0.020	1
17	3	3	VAL	HA	H	4.168	0.020	1
18	3	3	VAL	HB	H	2.076	0.020	1
19	3	3	VAL	HG1	H	0.918	0.020	2
20	3	3	VAL	HG2	H	0.918	0.020	2
21	3	3	VAL	CA	C	62.380	0.400	1
22	3	3	VAL	CB	C	33.210	0.400	1
23	3	3	VAL	CG1	C	21.390	0.400	2
24	3	3	VAL	CG2	C	21.390	0.400	2
25	3	3	VAL	N	N	119.136	0.400	1
26	4	4	SER	H	H	8.549	0.020	1
27	4	4	SER	HA	H	4.482	0.020	1
28	4	4	SER	HB2	H	3.878	0.020	2
29	4	4	SER	HB3	H	3.878	0.020	2
30	4	4	SER	CA	C	58.470	0.400	1
31	4	4	SER	CB	C	64.880	0.400	1
32	4	4	SER	N	N	120.034	0.400	1
33	5	5	VAL	H	H	8.304	0.020	1
34	5	5	VAL	HA	H	4.070	0.020	1
35	5	5	VAL	HB	H	2.090	0.020	1
36	5	5	VAL	HG1	H	0.943	0.020	2
37	5	5	VAL	HG2	H	0.943	0.020	2
38	5	5	VAL	CA	C	63.230	0.400	1
39	5	5	VAL	CB	C	32.770	0.400	1
40	5	5	VAL	CG1	C	21.370	0.400	2
41	5	5	VAL	CG2	C	21.370	0.400	2
42	5	5	VAL	N	N	122.497	0.400	1
43	6	6	MET	H	H	8.407	0.020	1
44	6	6	MET	HA	H	4.428	0.020	1
45	6	6	MET	HB2	H	2.052	0.020	2
46	6	6	MET	HB3	H	2.052	0.020	2
47	6	6	MET	HG2	H	2.555	0.020	2
48	6	6	MET	HG3	H	2.555	0.020	2
49	6	6	MET	CA	C	56.170	0.400	1
50	6	6	MET	CB	C	33.270	0.400	1
51	6	6	MET	CG	C	32.750	0.400	1
52	6	6	MET	N	N	122.782	0.400	1
53	7	7	GLU	H	H	8.350	0.020	1
54	7	7	GLU	HA	H	4.204	0.020	1
55	7	7	GLU	HB2	H	2.001	0.020	2
56	7	7	GLU	HB3	H	2.001	0.020	2
57	7	7	GLU	HG2	H	2.288	0.020	2
58	7	7	GLU	HG3	H	2.288	0.020	2
59	7	7	GLU	CA	C	57.280	0.400	1
60	7	7	GLU	CB	C	30.730	0.400	1
61	7	7	GLU	CG	C	36.570	0.400	1
62	7	7	GLU	N	N	122.264	0.400	1
63	8	8	ALA	H	H	8.334	0.020	1
64	8	8	ALA	HA	H	4.267	0.020	1
65	8	8	ALA	HB	H	1.397	0.020	1
66	8	8	ALA	CA	C	53.180	0.400	1
248	25	25	GLU	H	H	8.537	0.020	1
249	25	25	GLU	HA	H	4.140	0.020	1
250	25	25	GLU	HB2	H	2.094	0.020	2
251	25	25	GLU	HB3	H	2.094	0.020	2
252	25	25	GLU	HG2	H	2.320	0.020	2
253	25	25	GLU	HG3	H	2.320	0.020	2
254	25	25	GLU	CA	C	58.780	0.400	1
255	25	25	GLU	CB	C	29.970	0.400	1
256	25	25	GLU	CG	C	36.620	0.400	1
257	25	25	GLU	N	N	120.605	0.400	1
258	26	26	ALA	H	H	8.129	0.020	1
259	26	26	ALA	HA	H	4.166	0.020	1
260	26	26	ALA	HB	H	1.516	0.020	1
261	26	26	ALA	CA	C	55.230	0.400	1
262	26	26	ALA	CB	C	18.470	0.400	1
263	26	26	ALA	N	N	122.943	0.400	1
264	27	27	LYS	H	H	8.120	0.020	1
265	27	27	LYS	HA	H	4.175	0.020	1
266	27	27	LYS	HB2	H	1.926	0.020	2
267	27	27	LYS	HB3	H	1.926	0.020	2
268	27	27	LYS	HD2	H	1.540	0.020	2
269	27	27	LYS	HD3	H	1.540	0.020	2
270	27	27	LYS	HE2	H	2.962	0.020	2
271	27	27	LYS	HE3	H	2.962	0.020	2
272	27	27	LYS	HG2	H	1.417	0.020	2
273	27	27	LYS	HG3	H	1.417	0.020	2
274	27	27	LYS	CA	C	58.830	0.400	1
275	27	27	LYS	CB	C	32.610	0.400	1
276	27	27	LYS	CD	C	29.650	0.400	1
277	27	27	LYS	CE	C	42.310	0.400	1
278	27	27	LYS	CG	C	25.200	0.400	1
279	27	27	LYS	N	N	120.716	0.400	1
280	28	28	ASN	H	H	8.310	0.020	1
281	28	28	ASN	HA	H	4.575	0.020	1
282	28	28	ASN	HB2	H	2.887	0.020	2
283	28	28	ASN	HB3	H	2.887	0.020	2
284	28	28	ASN	HD21	H	7.736	0.020	2
285	28	28	ASN	HD22	H	6.953	0.020	2
286	28	28	ASN	CA	C	55.350	0.400	1
287	28	28	ASN	CB	C	38.420	0.400	1
288	28	28	ASN	N	N	118.724	0.400	1
289	28	28	ASN	ND2	N	112.900	0.400	1
290	29	29	ARG	H	H	8.301	0.020	1
291	29	29	ARG	HA	H	4.191	0.020	1
292	29	29	ARG	HB2	H	1.927	0.020	2
293	29	29	ARG	HB3	H	1.927	0.020	2
294	29	29	ARG	HD2	H	3.243	0.020	2
295	29	29	ARG	HD3	H	3.243	0.020	2
296	29	29	ARG	HG2	H	1.754	0.020	2
297	29	29	ARG	HG3	H	1.703	0.020	2
298	29	29	ARG	CA	C	58.390	0.400	1
299	29	29	ARG	CB	C	30.510	0.400	1
300	29	29	ARG	CD	C	43.420	0.400	1
301	29	29	ARG	CG	C	27.530	0.400	1
302	30	30	ALA	H	H	8.110	0.020	1
303	30	30	ALA	HA	H	4.175	0.020	1
304	30	30	ALA	HB	H	1.522	0.020	1
305	30	30	ALA	CA	C	55.230	0.400	1
306	30	30	ALA	CB	C	18.560	0.400	1
307	30	30	ALA	N	N	122.173	0.400	1
308	31	31	GLU	H	H	8.230	0.020	1
309	31	31	GLU	HA	H	4.160	0.020	1
310	31	31	GLU	HB2	H	2.113	0.020	2
311	31	31	GLU	HB3	H	2.113	0.020	2
312	31	31	GLU	HG2	H	2.375	0.020	2
313	31	31	GLU	HG3	H	2.375	0.020	2

```
 67   8   8 ALA CB   C  19.19  0.400 1
 68   8   8 ALA N    N 125.161 0.400 1
 69   9   9 ILE H    H   8.149 0.020 1
 70   9   9 ILE HA   H   4.048 0.020 1
 71   9   9 ILE HB   H   1.873 0.020 1
 72   9   9 ILE HD1  H   0.902 0.020 1
 73   9   9 ILE HG12 H   1.539 0.020 2
 74   9   9 ILE HG13 H   1.539 0.020 2
 75   9   9 ILE HG2  H   1.222 0.020 1
 76   9   9 ILE CA   C  62.030 0.400 1
 77   9   9 ILE CB   C  38.690 0.400 1
 78   9   9 ILE CD1  C  13.200 0.400 1
 79   9   9 ILE CG1  C  28.240 0.400 1
 80   9   9 ILE CG2  C  17.830 0.400 1
 81   9   9 ILE N    N 120.286 0.400 1
 82  10  10 LYS H    H   8.282 0.020 1
 83  10  10 LYS HA   H   4.233 0.020 1
 84  10  10 LYS HB2  H   1.829 0.020 2
 85  10  10 LYS HB3  H   1.829 0.020 2
 86  10  10 LYS HD2  H   1.506 0.020 2
 87  10  10 LYS HD3  H   1.506 0.020 2
 88  10  10 LYS HE2  H   2.992 0.020 2
 89  10  10 LYS HE3  H   2.992 0.020 2
 90  10  10 LYS HG2  H   1.416 0.020 2
 91  10  10 LYS HG3  H   1.416 0.020 2
 92  10  10 LYS CA   C  57.250 0.400 1
 93  10  10 LYS CB   C  33.240 0.400 1
 94  10  10 LYS CD   C  29.470 0.400 1
 95  10  10 LYS CE   C  42.330 0.400 1
 96  10  10 LYS CG   C  25.070 0.400 1
 97  10  10 LYS N    N 125.057 0.400 1
 98  11  11 GLU H    H   8.394 0.020 1
 99  11  11 GLU HA   H   4.217 0.020 1
100  11  11 GLU HB2  H   2.002 0.020 2
101  11  11 GLU HB3  H   2.002 0.020 2
102  11  11 GLU HG2  H   2.282 0.020 2
103  11  11 GLU HG3  H   2.282 0.020 2
104  11  11 GLU CA   C  57.540 0.400 1
105  11  11 GLU CB   C  30.700 0.400 1
106  11  11 GLU CG   C  36.590 0.400 1
107  11  11 GLU N    N 122.003 0.400 1
108  12  12 VAL H    H   8.220 0.020 1
109  12  12 VAL HA   H   3.988 0.020 1
110  12  12 VAL HB   H   2.081 0.020 1
111  12  12 VAL HG1  H   0.956 0.020 2
112  12  12 VAL HG2  H   0.956 0.020 2
113  12  12 VAL CA   C  63.720 0.400 1
114  12  12 VAL CB   C  32.680 0.400 1
115  12  12 VAL CG1  C  21.660 0.400 2
116  12  12 VAL CG2  C  21.660 0.400 2
117  12  12 VAL N    N 122.828 0.400 1
118  13  13 LYS H    H   8.385 0.020 1
119  13  13 LYS HA   H   4.316 0.020 1
120  13  13 LYS HB2  H   1.815 0.020 2
121  13  13 LYS HB3  H   1.815 0.020 2
122  13  13 LYS HD2  H   1.670 0.020 2
123  13  13 LYS HD3  H   1.670 0.020 2
124  13  13 LYS HG2  H   1.455 0.020 2
125  13  13 LYS HG3  H   1.455 0.020 2
126  13  13 LYS CA   C  56.990 0.400 1
127  13  13 LYS CB   C  33.030 0.400 1
128  13  13 LYS CD   C  29.300 0.400 1
129  13  13 LYS CE   C  42.280 0.400 1
130  13  13 LYS CG   C  25.060 0.400 1
131  13  13 LYS N    N 124.784 0.400 1
132  14  14 LEU H    H   8.360 0.020 1
133  14  14 LEU HA   H   4.272 0.020 1
134  14  14 LEU HB2  H   1.672 0.020 2
135  14  14 LEU HB3  H   1.672 0.020 2
136  14  14 LEU HD1  H   0.906 0.020 2
137  14  14 LEU HD2  H   0.906 0.020 2
138  14  14 LEU CA   C  56.450 0.400 1
139  14  14 LEU CB   C  42.510 0.400 1
140  14  14 LEU CD1  C  25.200 0.400 2
141  14  14 LEU CD2  C  23.680 0.400 2
314  31  31 GLU CA   C  58.380 0.400 1
315  31  31 GLU CB   C  29.830 0.400 1
316  31  31 GLU CG   C  36.610 0.400 1
317  31  31 GLU N    N 118.620 0.400 1
318  32  32 GLN H    H   8.160 0.020 1
319  32  32 GLN HA   H   4.214 0.020 1
320  32  32 GLN HB2  H   2.163 0.020 2
321  32  32 GLN HB3  H   2.163 0.020 2
322  32  32 GLN HE21 H   7.694 0.020 2
323  32  32 GLN HE22 H   6.895 0.020 2
324  32  32 GLN HG2  H   2.386 0.020 2
325  32  32 GLN HG3  H   2.467 0.020 2
326  32  32 GLN CA   C  58.100 0.400 1
327  32  32 GLN CB   C  28.880 0.400 1
328  32  32 GLN CG   C  33.820 0.400 1
329  32  32 GLN N    N 121.642 0.400 1
330  32  32 GLN NE2  N 112.600 0.400 1
331  33  33 ILE H    H   8.180 0.020 1
332  33  33 ILE HA   H   3.927 0.020 1
333  33  33 ILE HB   H   1.891 0.020 1
334  33  33 ILE HD1  H   0.921 0.020 1
335  33  33 ILE HG12 H   1.638 0.020 2
336  33  33 ILE HG13 H   1.638 0.020 2
337  33  33 ILE HG2  H   1.206 0.020 1
338  33  33 ILE CA   C  63.620 0.400 1
339  33  33 ILE CB   C  38.530 0.400 1
340  33  33 ILE CD1  C  13.310 0.400 1
341  33  33 ILE CG1  C  28.370 0.400 1
342  33  33 ILE CG2  C  17.640 0.400 1
343  33  33 ILE N    N 121.132 0.400 1
344  34  34 LYS H    H   7.964 0.020 1
345  34  34 LYS HA   H   4.186 0.020 1
346  34  34 LYS HB2  H   1.892 0.020 2
347  34  34 LYS HB3  H   1.892 0.020 2
348  34  34 LYS HD2  H   1.537 0.020 2
349  34  34 LYS HD3  H   1.537 0.020 2
350  34  34 LYS HE2  H   2.978 0.020 2
351  34  34 LYS HE3  H   2.978 0.020 2
352  34  34 LYS HG2  H   1.450 0.020 2
353  34  34 LYS HG3  H   1.450 0.020 2
354  34  34 LYS CA   C  58.380 0.400 1
355  34  34 LYS CB   C  33.120 0.400 1
356  34  34 LYS CD   C  29.650 0.400 1
357  34  34 LYS CE   C  42.280 0.400 1
358  34  34 LYS CG   C  25.250 0.400 1
359  34  34 LYS N    N 122.920 0.400 1
360  35  35 ALA H    H   8.215 0.020 1
361  35  35 ALA HA   H   4.178 0.020 1
362  35  35 ALA HB   H   1.509 0.020 1
363  35  35 ALA CA   C  54.890 0.400 1
364  35  35 ALA CB   C  18.710 0.400 1
365  35  35 ALA N    N 123.352 0.400 1
366  36  36 GLU H    H   8.434 0.020 1
367  36  36 GLU HA   H   4.145 0.020 1
368  36  36 GLU HB2  H   2.075 0.020 2
369  36  36 GLU HB3  H   2.075 0.020 2
370  36  36 GLU HG2  H   2.303 0.020 2
371  36  36 GLU HG3  H   2.466 0.020 2
372  36  36 GLU CA   C  58.360 0.400 1
373  36  36 GLU CB   C  30.010 0.400 1
374  36  36 GLU CG   C  36.800 0.400 1
375  36  36 GLU N    N 119.725 0.400 1
376  37  37 ALA H    H   8.084 0.020 1
377  37  37 ALA HA   H   4.285 0.020 1
378  37  37 ALA HB   H   1.507 0.020 1
379  37  37 ALA CA   C  54.440 0.400 1
380  37  37 ALA CB   C  18.640 0.400 1
381  37  37 ALA N    N 123.710 0.400 1
382  38  38 ILE H    H   7.997 0.020 1
383  38  38 ILE HA   H   3.953 0.020 1
384  38  38 ILE HB   H   1.969 0.020 1
385  38  38 ILE HD1  H   0.914 0.020 1
386  38  38 ILE HG12 H   1.658 0.020 2
387  38  38 ILE HG13 H   1.658 0.020 2
388  38  38 ILE HG2  H   1.203 0.020 1
```

```
142 14 14 LEU CG   C  27.470 0.400 1
143 14 14 LEU N    N 123.256 0.400 1
144 15 15 ALA H    H   8.301 0.020 1
145 15 15 ALA HA   H   4.261 0.020 1
146 15 15 ALA HB   H   1.450 0.020 1
147 15 15 ALA CA   C  53.810 0.400 1
148 15 15 ALA CB   C  18.960 0.400 1
149 15 15 ALA N    N 124.067 0.400 1
150 16 16 GLU H    H   8.441 0.020 1
151 16 16 GLU HA   H   4.195 0.020 1
152 16 16 GLU HB2  H   2.077 0.020 2
153 16 16 GLU HB3  H   2.077 0.020 2
154 16 16 GLU HG2  H   2.261 0.020 2
155 16 16 GLU HG3  H   2.390 0.020 2
156 16 16 GLU CA   C  58.190 0.400 1
157 16 16 GLU CB   C  30.360 0.400 1
158 16 16 GLU CG   C  36.730 0.400 1
159 16 16 GLU N    N 120.401 0.400 1
160 17 17 GLU H    H   8.393 0.020 1
161 17 17 GLU HA   H   4.168 0.020 1
162 17 17 GLU HB2  H   2.080 0.020 2
163 17 17 GLU HB3  H   2.080 0.020 2
164 17 17 GLU HG2  H   2.275 0.020 2
165 17 17 GLU HG3  H   2.406 0.020 2
166 17 17 GLU CA   C  58.270 0.400 1
167 17 17 GLU CB   C  30.340 0.400 1
168 17 17 GLU CG   C  36.820 0.400 1
169 17 17 GLU N    N 120.884 0.400 1
170 18 18 GLN H    H   8.353 0.020 1
171 18 18 GLN HA   H   4.170 0.020 1
172 18 18 GLN HB2  H   2.100 0.020 2
173 18 18 GLN HB3  H   2.100 0.020 2
174 18 18 GLN HE21 H   7.747 0.020 2
175 18 18 GLN HE22 H   6.944 0.020 2
176 18 18 GLN HG2  H   2.414 0.020 2
177 18 18 GLN HG3  H   2.414 0.020 2
178 18 18 GLN CA   C  57.410 0.400 1
179 18 18 GLN CB   C  29.080 0.400 1
180 18 18 GLN CG   C  33.900 0.400 1
181 18 18 GLN N    N 120.718 0.400 1
182 18 18 GLN NE2  N 112.900 0.400 1
183 19 19 ALA H    H   8.191 0.020 1
184 19 19 ALA HA   H   4.272 0.020 1
185 19 19 ALA HB   H   1.481 0.020 1
186 19 19 ALA CA   C  54.310 0.400 1
187 19 19 ALA CB   C  18.820 0.400 1
188 19 19 ALA N    N 123.787 0.400 1
189 20 20 VAL H    H   8.000 0.020 1
190 20 20 VAL HA   H   3.857 0.020 1
191 20 20 VAL HB   H   2.128 0.020 1
192 20 20 VAL HG1  H   1.027 0.020 2
193 20 20 VAL HG2  H   1.027 0.020 2
194 20 20 VAL CA   C  64.950 0.400 1
195 20 20 VAL CB   C  32.620 0.400 1
196 20 20 VAL CG1  C  22.180 0.400 2
197 20 20 VAL CG2  C  22.180 0.400 2
198 20 20 VAL N    N 118.988 0.400 1
199 21 21 LYS H    H   8.117 0.020 1
200 21 21 LYS HA   H   4.197 0.020 1
201 21 21 LYS HB2  H   1.895 0.020 2
202 21 21 LYS HB3  H   1.895 0.020 2
203 21 21 LYS HD2  H   1.711 0.020 2
204 21 21 LYS HD3  H   1.711 0.020 2
205 21 21 LYS HE2  H   2.997 0.020 2
206 21 21 LYS HE3  H   2.997 0.020 2
207 21 21 LYS HG2  H   1.476 0.020 2
208 21 21 LYS HG3  H   1.476 0.020 2
209 21 21 LYS CA   C  58.390 0.400 1
210 21 21 LYS CB   C  32.910 0.400 1
211 21 21 LYS CD   C  29.530 0.400 1
212 21 21 LYS CE   C  42.420 0.400 1
213 21 21 LYS CG   C  25.070 0.400 1
214 21 21 LYS N    N 123.022 0.400 1
215 22 22 GLU H    H   8.308 0.020 1
216 22 22 GLU HA   H   4.173 0.020 1
389 38 38 ILE CA   C  63.450 0.400 1
390 38 38 ILE CB   C  38.540 0.400 1
391 38 38 ILE CD1  C  13.370 0.400 1
392 38 38 ILE CG1  C  28.030 0.400 1
393 38 38 ILE CG2  C  17.780 0.400 1
394 38 38 ILE N    N 120.850 0.400 1
395 39 39 GLU H    H   8.100 0.020 1
396 39 39 GLU HA   H   4.156 0.020 1
397 39 39 GLU HB2  H   2.071 0.020 2
398 39 39 GLU HB3  H   2.071 0.020 2
399 39 39 GLU HG2  H   2.317 0.020 2
400 39 39 GLU HG3  H   2.317 0.020 2
401 39 39 GLU CA   C  58.230 0.400 1
402 39 39 GLU CB   C  29.990 0.400 1
403 39 39 GLU CG   C  36.270 0.400 1
404 39 39 GLU N    N 122.251 0.400 1
405 40 40 GLU H    H   8.435 0.020 1
406 40 40 GLU HA   H   4.108 0.020 1
407 40 40 GLU HB2  H   2.065 0.020 2
408 40 40 GLU HB3  H   2.065 0.020 2
409 40 40 GLU HG2  H   2.329 0.020 2
410 40 40 GLU HG3  H   2.329 0.020 2
411 40 40 GLU CA   C  58.360 0.400 1
412 40 40 GLU CB   C  30.090 0.400 1
413 40 40 GLU CG   C  36.330 0.400 1
414 40 40 GLU N    N 121.271 0.400 1
415 41 41 ALA H    H   8.054 0.020 1
416 41 41 ALA HA   H   4.221 0.020 1
417 41 41 ALA HB   H   1.483 0.020 1
418 41 41 ALA CA   C  54.180 0.400 1
419 41 41 ALA CB   C  18.680 0.400 1
420 41 41 ALA N    N 122.548 0.400 1
421 42 42 LYS H    H   7.924 0.020 1
422 42 42 LYS HA   H   4.168 0.020 1
423 42 42 LYS HB2  H   1.878 0.020 2
424 42 42 LYS HB3  H   1.878 0.020 2
425 42 42 LYS HD2  H   1.615 0.020 2
426 42 42 LYS HD3  H   1.615 0.020 2
427 42 42 LYS HE2  H   2.950 0.020 2
428 42 42 LYS HE3  H   2.950 0.020 2
429 42 42 LYS HG2  H   1.467 0.020 2
430 42 42 LYS HG3  H   1.467 0.020 2
431 42 42 LYS CA   C  57.910 0.400 1
432 42 42 LYS CB   C  33.100 0.400 1
433 42 42 LYS CD   C  29.950 0.400 1
434 42 42 LYS CE   C  42.370 0.400 1
435 42 42 LYS CG   C  25.710 0.400 1
436 42 42 LYS N    N 118.809 0.400 1
437 43 43 LYS H    H   7.860 0.020 1
438 43 43 LYS HA   H   4.203 0.020 1
439 43 43 LYS HB2  H   1.877 0.020 2
440 43 43 LYS HB3  H   1.877 0.020 2
441 43 43 LYS HD2  H   1.731 0.020 2
442 43 43 LYS HD3  H   1.731 0.020 2
443 43 43 LYS HE2  H   2.982 0.020 2
444 43 43 LYS HE3  H   2.982 0.020 2
445 43 43 LYS HG2  H   1.446 0.020 2
446 43 43 LYS HG3  H   1.574 0.020 2
447 43 43 LYS CA   C  57.380 0.400 1
448 43 43 LYS CB   C  33.140 0.400 1
449 43 43 LYS CD   C  29.740 0.400 1
450 43 43 LYS CE   C  42.350 0.400 1
451 43 43 LYS CG   C  25.410 0.400 1
452 43 43 LYS N    N 120.047 0.400 1
453 44 44 LEU H    H   7.919 0.020 1
454 44 44 LEU HA   H   4.336 0.020 1
455 44 44 LEU HB2  H   1.685 0.020 2
456 44 44 LEU HB3  H   1.685 0.020 2
457 44 44 LEU HD1  H   0.915 0.020 2
458 44 44 LEU HD2  H   0.915 0.020 2
459 44 44 LEU CA   C  55.680 0.400 1
460 44 44 LEU CB   C  42.510 0.400 1
461 44 44 LEU CD1  C  25.200 0.400 2
462 44 44 LEU CD2  C  23.560 0.400 2
463 44 44 LEU CG   C  27.340 0.400 1
```

```
217 22 22 GLU HB2  H   2.088 0.020 2
218 22 22 GLU HB3  H   2.088 0.020 2
219 22 22 GLU HG2  H   2.290 0.020 2
220 22 22 GLU HG3  H   2.417 0.020 2
221 22 22 GLU CA   C  58.720 0.400 1
222 22 22 GLU CB   C  29.900 0.400 1
223 22 22 GLU CG   C  36.730 0.400 1
224 22 22 GLU N    N 120.162 0.400 1
225 23 23 ILE H    H   7.994 0.020 1
226 23 23 ILE HA   H   3.931 0.020 1
227 23 23 ILE HB   H   2.001 0.020 1
228 23 23 ILE HD1  H   0.930 0.020 1
229 23 23 ILE HG12 H   1.678 0.020 2
230 23 23 ILE HG13 H   1.678 0.020 2
231 23 23 ILE HG2  H   1.209 0.020 1
232 23 23 ILE CA   C  63.670 0.400 1
233 23 23 ILE CB   C  38.420 0.400 1
234 23 23 ILE CD1  C  13.132 0.400 1
235 23 23 ILE CG1  C  28.590 0.400 1
236 23 23 ILE CG2  C  17.660 0.400 1
237 23 23 ILE N    N 121.937 0.400 1
238 24 24 GLU H    H   8.261 0.020 1
239 24 24 GLU HA   H   4.131 0.020 1
240 24 24 GLU HB2  H   2.094 0.020 2
241 24 24 GLU HB3  H   2.094 0.020 2
242 24 24 GLU HG2  H   2.320 0.020 2
243 24 24 GLU HG3  H   2.320 0.020 2
244 24 24 GLU CA   C  58.680 0.400 1
245 24 24 GLU CB   C  29.970 0.400 1
246 24 24 GLU CG   C  36.360 0.400 1
247 24 24 GLU N    N 122.682 0.400 1
```

```
464 44 44 LEU N    N 121.328 0.400 1
465 45 45 ILE H    H   7.880 0.020 1
466 45 45 ILE HA   H   4.171 0.020 1
467 45 45 ILE HB   H   1.910 0.020 1
468 45 45 ILE HD1  H   0.914 0.020 1
469 45 45 ILE HG12 H   1.514 0.020 2
470 45 45 ILE HG13 H   1.514 0.020 2
471 45 45 ILE HG2  H   1.205 0.020 1
472 45 45 ILE CA   C  61.230 0.400 1
473 45 45 ILE CB   C  38.880 0.400 1
474 45 45 ILE CD1  C  13.114 0.400 1
475 45 45 ILE CG1  C  28.100 0.400 1
476 45 45 ILE CG2  C  17.850 0.400 1
477 45 45 ILE N    N 121.011 0.400 1
478 46 46 ALA H    H   8.349 0.020 1
479 46 46 ALA HA   H   4.392 0.020 1
480 46 46 ALA HB   H   1.423 0.020 1
481 46 46 ALA CA   C  52.380 0.400 1
482 46 46 ALA CB   C  19.550 0.400 1
483 46 46 ALA N    N 129.057 0.400 1
484 47 47 CYS H    H   8.043 0.020 1
485 47 47 CYS CA   C  59.035 0.400 1
486 47 47 CYS CB   C  28.638 0.400 1
487 47 47 CYS N  N 123.478 0.400 1
```

Appendix C

Chemical shift assignments of E_{1-52} of A_1A_O ATP synthase from *M. jannaschii*

The chemical shift assignment are listed in the NMR-STAR format (Columns are as follows: Atom shift assign ID, Residue author sequence code, Residue sequence code, Residue label, Atom name, Atom type, Chemical shift value, Chemical shift value error, Chemical shift ambiguity code

1	1	1	MET	H	H	9.714	0.020	1
2	1	1	MET	HA	H	4.455	0.020	1
3	1	1	MET	HB2	H	2.135	0.020	2
4	1	1	MET	HB3	H	2.135	0.020	2
5	1	1	MET	HG2	H	2.609	0.020	2
6	1	1	MET	HG3	H	2.609	0.020	2
7	1	1	MET	C	C	177.900	0.400	1
8	1	1	MET	CA	C	57.020	0.400	1
9	1	1	MET	CB	C	31.410	0.400	1
10	1	1	MET	N	N	121.448	0.400	1
11	2	2	ALA	H	H	7.905	0.020	1
12	2	2	ALA	HA	H	4.072	0.020	1
13	2	2	ALA	HB	H	1.506	0.020	1
14	2	2	ALA	C	C	179.200	0.400	1
15	2	2	ALA	CA	C	55.420	0.400	1
16	2	2	ALA	CB	C	17.460	0.400	1
17	2	2	ALA	N	N	121.462	0.400	1
18	3	3	VAL	H	H	7.612	0.020	1
19	3	3	VAL	HA	H	3.722	0.020	1
20	3	3	VAL	HB	H	2.139	0.020	1
21	3	3	VAL	HG1	H	0.962	0.020	2
22	3	3	VAL	HG2	H	0.962	0.020	2
23	3	3	VAL	C	C	179.200	0.400	1
24	3	3	VAL	CA	C	65.420	0.400	1
25	3	3	VAL	CB	C	31.260	0.400	1
26	3	3	VAL	N	N	115.271	0.400	1
27	4	4	LYS	H	H	7.656	0.020	1
28	4	4	LYS	HA	H	4.135	0.020	1
29	4	4	LYS	HB2	H	1.972	0.020	2
30	4	4	LYS	HB3	H	1.972	0.020	2
31	4	4	LYS	HD2	H	1.751	0.020	2
32	4	4	LYS	HD3	H	1.751	0.020	2
33	4	4	LYS	HE2	H	2.978	0.020	2
34	4	4	LYS	HE3	H	2.978	0.020	2
35	4	4	LYS	HG2	H	1.539	0.020	2
36	4	4	LYS	HG3	H	1.539	0.020	2
37	4	4	LYS	C	C	178.600	0.400	1
38	4	4	LYS	CA	C	58.660	0.400	1
39	4	4	LYS	CB	C	31.630	0.400	1
40	4	4	LYS	N	N	119.781	0.400	1
41	5	5	LEU	H	H	8.051	0.020	1
42	5	5	LEU	HA	H	4.150	0.020	1
43	5	5	LEU	HB2	H	1.805	0.020	2
44	5	5	LEU	HB3	H	1.805	0.020	2
45	5	5	LEU	HD1	H	1.008	0.020	2
46	5	5	LEU	HD2	H	0.843	0.020	2
47	5	5	LEU	C	C	179.200	0.400	1
48	5	5	LEU	CA	C	57.160	0.400	1
49	5	5	LEU	CB	C	41.450	0.400	1
50	6	6	MET	H	H	8.081	0.020	1
51	6	6	MET	HA	H	4.315	0.020	1
52	6	6	MET	HB2	H	2.174	0.020	2
53	6	6	MET	HB3	H	2.174	0.020	2
54	6	6	MET	HG2	H	2.722	0.020	2
55	6	6	MET	HG3	H	2.572	0.020	2
56	6	6	MET	C	C	177.900	0.400	1
57	6	6	MET	CA	C	57.400	0.400	1
58	6	6	MET	CB	C	32.380	0.400	1
59	6	6	MET	N	N	117.170	0.400	1
60	7	7	GLY	H	H	7.987	0.020	1
61	7	7	GLY	HA2	H	3.995	0.020	2
62	7	7	GLY	HA3	H	3.995	0.020	2
260	27	27	ILE	H	H	8.357	0.020	1
261	27	27	ILE	HA	H	3.827	0.020	1
262	27	27	ILE	HB	H	1.948	0.020	1
263	27	27	ILE	HD1	H	0.937	0.020	1
264	27	27	ILE	HG12	H	1.564	0.020	2
265	27	27	ILE	HG13	H	1.779	0.020	2
266	27	27	ILE	HG2	H	1.208	0.020	1
267	27	27	ILE	C	C	179.000	0.400	1
268	27	27	ILE	CA	C	64.620	0.400	1
269	27	27	ILE	CB	C	37.770	0.400	1
270	27	27	ILE	N	N	119.945	0.400	1
271	28	28	SER	H	H	8.155	0.020	1
272	28	28	SER	HA	H	4.258	0.020	1
273	28	28	SER	HB2	H	4.080	0.020	2
274	28	28	SER	HB3	H	4.080	0.020	2
275	28	28	SER	C	C	175.800	0.400	1
276	28	28	SER	CA	C	61.530	0.400	1
277	28	28	SER	CB	C	62.570	0.400	1
278	28	28	SER	N	N	115.157	0.400	1
279	29	29	GLU	H	H	8.272	0.020	1
280	29	29	GLU	HA	H	4.164	0.020	1
281	29	29	GLU	HB2	H	2.326	0.020	2
282	29	29	GLU	HB3	H	2.168	0.020	2
283	29	29	GLU	HG2	H	2.518	0.020	2
284	29	29	GLU	HG3	H	2.518	0.020	2
285	29	29	GLU	C	C	178.200	0.400	1
286	29	29	GLU	CA	C	59.120	0.400	1
287	29	29	GLU	CB	C	28.650	0.400	1
288	29	29	GLU	N	N	122.866	0.400	1
289	30	30	ALA	H	H	8.303	0.020	1
290	30	30	ALA	HA	H	4.206	0.020	1
291	30	30	ALA	HB	H	1.578	0.020	1
292	30	30	ALA	C	C	179.200	0.400	1
293	30	30	ALA	CA	C	55.120	0.400	1
294	30	30	ALA	CB	C	17.180	0.400	1
295	30	30	ALA	N	N	123.088	0.400	1
296	31	31	GLU	H	H	8.439	0.020	1
297	31	31	GLU	HA	H	4.108	0.020	1
298	31	31	GLU	HB2	H	2.176	0.020	2
299	31	31	GLU	HB3	H	2.176	0.020	2
300	31	31	GLU	HG2	H	2.401	0.020	2
301	31	31	GLU	HG3	H	2.401	0.020	2
302	31	31	GLU	C	C	178.200	0.400	1
303	31	31	GLU	CA	C	58.950	0.400	1
304	31	31	GLU	CB	C	28.890	0.400	1
305	31	31	GLU	N	N	117.999	0.400	1
306	32	32	ALA	H	H	7.988	0.020	1
307	32	32	ALA	HA	H	4.191	0.020	1
308	32	32	ALA	HB	H	1.175	0.020	1
309	32	32	ALA	C	C	179.200	0.400	1
310	32	32	ALA	CA	C	54.970	0.400	1
311	32	32	ALA	CB	C	17.260	0.400	1
312	32	32	ALA	N	N	122.398	0.400	1
313	33	33	GLU	H	H	8.060	0.020	1
314	33	33	GLU	HA	H	4.144	0.020	1
315	33	33	GLU	HB2	H	2.188	0.020	2
316	33	33	GLU	HB3	H	2.188	0.020	2
317	33	33	GLU	HG2	H	2.381	0.020	2
318	33	33	GLU	HG3	H	2.381	0.020	2
319	33	33	GLU	C	C	178.200	0.400	1
320	33	33	GLU	CA	C	58.950	0.400	1
321	33	33	GLU	CB	C	28.890	0.400	1

```
 63   7   7 GLY C    C 175.700 0.400 1
 64   7   7 GLY CA   C  46.050 0.400 1
 65   7   7 GLY N    N 107.082 0.400 1
 66   8   8 VAL H    H   7.928 0.020 1
 67   8   8 VAL HA   H   3.807 0.020 1
 68   8   8 VAL HB   H   2.204 0.020 1
 69   8   8 VAL HG1  H   1.029 0.020 2
 70   8   8 VAL HG2  H   1.029 0.020 2
 71   8   8 VAL C    C 179.200 0.400 1
 72   8   8 VAL CA   C  65.270 0.400 1
 73   8   8 VAL CB   C  31.190 0.400 1
 74   8   8 VAL N    N 120.229 0.400 1
 75   9   9 ASP H    H   8.260 0.020 1
 76   9   9 ASP HA   H   4.415 0.020 1
 77   9   9 ASP HB2  H   2.726 0.020 2
 78   9   9 ASP HB3  H   2.726 0.020 2
 79   9   9 ASP C    C 179.300 0.400 1
 80   9   9 ASP CA   C  56.950 0.400 1
 81   9   9 ASP CB   C  40.240 0.400 1
 82   9   9 ASP N    N 119.778 0.400 1
 83  10  10 LYS H    H   7.901 0.020 1
 84  10  10 LYS HA   H   4.133 0.020 1
 85  10  10 LYS HB2  H   1.983 0.020 2
 86  10  10 LYS HB3  H   1.983 0.020 2
 87  10  10 LYS HD2  H   1.692 0.020 2
 88  10  10 LYS HD3  H   1.692 0.020 2
 89  10  10 LYS HE2  H   2.977 0.020 2
 90  10  10 LYS HE3  H   2.977 0.020 2
 91  10  10 LYS HG2  H   1.542 0.020 2
 92  10  10 LYS HG3  H   1.542 0.020 2
 93  10  10 LYS C    C 178.600 0.400 1
 94  10  10 LYS CA   C  59.870 0.400 1
 95  10  10 LYS CB   C  31.780 0.400 1
 96  10  10 LYS N    N 119.866 0.400 1
 97  11  11 ILE H    H   7.778 0.020 1
 98  11  11 ILE HA   H   3.837 0.020 1
 99  11  11 ILE HB   H   2.005 0.020 1
100  11  11 ILE HD1  H   0.917 0.020 1
101  11  11 ILE HG12 H   1.660 0.020 2
102  11  11 ILE HG13 H   1.660 0.020 2
103  11  11 ILE HG2  H   1.188 0.020 1
104  11  11 ILE C    C 179.000 0.400 1
105  11  11 ILE CA   C  64.300 0.400 1
106  11  11 ILE CB   C  37.140 0.400 1
107  11  11 ILE N    N 120.208 0.400 1
108  12  12 LYS H    H   8.311 0.020 1
109  12  12 LYS HA   H   3.964 0.020 1
110  12  12 LYS HB2  H   1.934 0.020 2
111  12  12 LYS HB3  H   1.934 0.020 2
112  12  12 LYS HD2  H   1.642 0.020 2
113  12  12 LYS HD3  H   1.642 0.020 2
114  12  12 LYS HE2  H   2.933 0.020 2
115  12  12 LYS HE3  H   2.933 0.020 2
116  12  12 LYS HG2  H   1.433 0.020 2
117  12  12 LYS HG3  H   1.433 0.020 2
118  12  12 LYS C    C 178.600 0.400 1
119  12  12 LYS CA   C  59.870 0.400 1
120  12  12 LYS CB   C  31.780 0.400 1
121  12  12 LYS N    N 119.425 0.400 1
122  13  13 SER H    H   7.964 0.020 1
123  13  13 SER HA   H   4.154 0.020 1
124  13  13 SER HB2  H   4.006 0.020 2
125  13  13 SER HB3  H   4.006 0.020 2
126  13  13 SER C    C 175.800 0.400 1
127  13  13 SER CA   C  61.320 0.400 1
128  13  13 SER CB   C  62.540 0.400 1
129  13  13 SER N    N 112.107 0.400 1
130  14  14 LYS H    H   7.681 0.020 1
131  14  14 LYS HA   H   4.205 0.020 1
132  14  14 LYS HB2  H   1.973 0.020 2
133  14  14 LYS HB3  H   1.973 0.020 2
134  14  14 LYS HD2  H   1.755 0.020 2
322  33  33 GLU N    N 118.582 0.400 1
323  34  34 LYS H    H   7.986 0.020 1
324  34  34 LYS HA   H   3.972 0.020 1
325  34  34 LYS HB2  H   2.169 0.020 2
326  34  34 LYS HB3  H   1.992 0.020 2
327  34  34 LYS HD2  H   1.625 0.020 2
328  34  34 LYS HD3  H   1.625 0.020 2
329  34  34 LYS HE2  H   2.932 0.020 2
330  34  34 LYS HE3  H   2.932 0.020 2
331  34  34 LYS HG2  H   1.593 0.020 2
332  34  34 LYS HG3  H   1.593 0.020 2
333  34  34 LYS C    C 178.600 0.400 1
334  34  34 LYS CA   C  58.110 0.400 1
335  34  34 LYS CB   C  30.980 0.400 1
336  34  34 LYS N    N 119.109 0.400 1
337  35  35 ALA H    H   7.977 0.020 1
338  35  35 ALA HA   H   4.067 0.020 1
339  35  35 ALA HB   H   1.552 0.020 1
340  35  35 ALA C    C 179.200 0.400 1
341  35  35 ALA CA   C  55.180 0.400 1
342  35  35 ALA CB   C  16.830 0.400 1
343  35  35 ALA N    N 119.701 0.400 1
344  36  36 LYS H    H   7.752 0.020 1
345  36  36 LYS HA   H   4.098 0.020 1
346  36  36 LYS HB2  H   1.996 0.020 2
347  36  36 LYS HB3  H   2.093 0.020 2
348  36  36 LYS HD2  H   1.733 0.020 2
349  36  36 LYS HD3  H   1.733 0.020 2
350  36  36 LYS HE2  H   2.949 0.020 2
351  36  36 LYS HE3  H   2.949 0.020 2
352  36  36 LYS HG2  H   1.592 0.020 2
353  36  36 LYS HG3  H   1.592 0.020 2
354  36  36 LYS C    C 178.600 0.400 1
355  36  36 LYS CA   C  59.110 0.400 1
356  36  36 LYS CB   C  31.780 0.400 1
357  36  36 LYS N    N 117.583 0.400 1
358  37  37 ILE H    H   7.960 0.020 1
359  37  37 ILE HA   H   3.719 0.020 1
360  37  37 ILE HB   H   2.044 0.020 1
361  37  37 ILE HD1  H   0.911 0.020 1
362  37  37 ILE HG12 H   1.576 0.020 2
363  37  37 ILE HG13 H   1.798 0.020 2
364  37  37 ILE HG2  H   1.164 0.020 1
365  37  37 ILE C    C 179.000 0.400 1
366  37  37 ILE CA   C  64.980 0.400 1
367  37  37 ILE CB   C  37.460 0.400 1
368  37  37 ILE N    N 120.897 0.400 1
369  38  38 LEU H    H   8.153 0.020 1
370  38  38 LEU HA   H   4.106 0.020 1
371  38  38 LEU HB2  H   1.846 0.020 2
372  38  38 LEU HB3  H   1.587 0.020 2
373  38  38 LEU HD1  H   0.906 0.020 2
374  38  38 LEU HD2  H   0.906 0.020 2
375  38  38 LEU C    C 179.200 0.400 1
376  38  38 LEU CA   C  57.730 0.400 1
377  38  38 LEU CB   C  40.980 0.400 1
378  38  38 LEU N    N 120.156 0.400 1
379  39  39 GLU H    H   8.115 0.020 1
380  39  39 GLU HA   H   4.071 0.020 1
381  39  39 GLU HB2  H   2.206 0.020 2
382  39  39 GLU HB3  H   2.206 0.020 2
383  39  39 GLU HG2  H   2.390 0.020 2
384  39  39 GLU HG3  H   2.390 0.020 2
385  39  39 GLU C    C 178.200 0.400 1
386  39  39 GLU CA   C  59.260 0.400 1
387  39  39 GLU CB   C  29.000 0.400 1
388  39  39 GLU N    N 119.628 0.400 1
389  40  40 LYS H    H   7.951 0.020 1
390  40  40 LYS HA   H   4.095 0.020 1
391  40  40 LYS HB2  H   1.989 0.020 2
392  40  40 LYS HB3  H   1.989 0.020 2
393  40  40 LYS HD2  H   1.734 0.020 2
```

```
135 14 14 LYS HD3  H   1.755 0.020 2
136 14 14 LYS HE2  H   2.956 0.020 2
137 14 14 LYS HE3  H   2.738 0.020 2
138 14 14 LYS HG2  H   1.547 0.020 2
139 14 14 LYS HG3  H   1.547 0.020 2
140 14 14 LYS C    C 178.600 0.400 1
141 14 14 LYS CA   C  58.750 0.400 1
142 14 14 LYS CB   C  31.590 0.400 1
143 14 14 LYS N    N 123.289 0.400 1
144 15 15 ILE H    H   8.163 0.020 1
145 15 15 ILE HA   H   3.813 0.020 1
146 15 15 ILE HB   H   1.976 0.020 1
147 15 15 ILE HD1  H   0.907 0.020 1
148 15 15 ILE HG2  H   1.174 0.020 1
149 15 15 ILE C    C 179.000 0.400 1
150 15 15 ILE CA   C  64.780 0.400 1
151 15 15 ILE CB   C  37.300 0.400 1
152 15 15 ILE N    N 119.789 0.400 1
153 16 16 LEU H    H   8.033 0.020 1
154 16 16 LEU HA   H   4.183 0.020 1
155 16 16 LEU HB2  H   1.838 0.020 2
156 16 16 LEU HB3  H   1.589 0.020 2
157 16 16 LEU HD1  H   0.904 0.020 2
158 16 16 LEU HD2  H   0.904 0.020 2
159 16 16 LEU C    C 179.200 0.400 1
160 16 16 LEU CA   C  57.410 0.400 1
161 16 16 LEU CB   C  41.260 0.400 1
162 16 16 LEU N    N 119.388 0.400 1
163 17 17 ASP H    H   8.210 0.020 1
164 17 17 ASP HA   H   4.495 0.020 1
165 17 17 ASP HB2  H   2.939 0.020 2
166 17 17 ASP HB3  H   2.760 0.020 2
167 17 17 ASP C    C 179.300 0.400 1
168 17 17 ASP CA   C  57.440 0.400 1
169 17 17 ASP CB   C  40.470 0.400 1
170 17 17 ASP N    N 120.521 0.400 1
171 18 18 ASP H    H   8.589 0.020 1
172 18 18 ASP HA   H   4.501 0.020 1
173 18 18 ASP HB2  H   2.940 0.020 2
174 18 18 ASP HB3  H   2.700 0.020 2
175 18 18 ASP C    C 179.300 0.400 1
176 18 18 ASP CA   C  57.150 0.400 1
177 18 18 ASP CB   C  39.550 0.400 1
178 18 18 ASP N    N 121.775 0.400 1
179 19 19 ALA H    H   8.393 0.020 1
180 19 19 ALA HA   H   4.196 0.020 1
181 19 19 ALA HB   H   1.560 0.020 1
182 19 19 ALA C    C 179.200 0.400 1
183 19 19 ALA CA   C  54.970 0.400 1
184 19 19 ALA CB   C  17.040 0.400 1
185 19 19 ALA N    N 123.962 0.400 1
186 20 20 LYS H    H   8.277 0.020 1
187 20 20 LYS HA   H   4.041 0.020 1
188 20 20 LYS HB2  H   1.983 0.020 2
189 20 20 LYS HB3  H   1.983 0.020 2
190 20 20 LYS HD2  H   1.709 0.020 2
191 20 20 LYS HD3  H   1.709 0.020 2
192 20 20 LYS HE2  H   2.940 0.020 2
193 20 20 LYS HE3  H   2.940 0.020 2
194 20 20 LYS HG2  H   1.543 0.020 2
195 20 20 LYS HG3  H   1.543 0.020 2
196 20 20 LYS C    C 178.600 0.400 1
197 20 20 LYS CA   C  59.300 0.400 1
198 20 20 LYS CB   C  31.820 0.400 1
199 20 20 LYS N    N 118.639 0.400 1
200 21 21 ALA H    H   7.989 0.020 1
201 21 21 ALA HA   H   4.193 0.020 1
202 21 21 ALA HB   H   1.593 0.020 1
203 21 21 ALA C    C 179.200 0.400 1
204 21 21 ALA CA   C  54.910 0.400 1
205 21 21 ALA CB   C  17.410 0.400 1
206 21 21 ALA N    N 121.825 0.400 1
394 40 40 LYS HD3  H   1.734 0.020 2
395 40 40 LYS HE2  H   2.993 0.020 2
396 40 40 LYS HE3  H   2.993 0.020 2
397 40 40 LYS HG2  H   1.595 0.020 2
398 40 40 LYS HG3  H   1.595 0.020 2
399 40 40 LYS C    C 178.600 0.400 1
400 40 40 LYS CA   C  59.380 0.400 1
401 40 40 LYS CB   C  31.820 0.400 1
402 40 40 LYS N    N 120.185 0.400 1
403 41 41 ALA H    H   8.550 0.020 1
404 41 41 ALA HA   H   4.120 0.020 1
405 41 41 ALA HB   H   1.540 0.020 1
406 41 41 ALA C    C 179.200 0.400 1
407 41 41 ALA CA   C  55.060 0.400 1
408 41 41 ALA CB   C  17.490 0.400 1
409 41 41 ALA N    N 122.931 0.400 1
410 42 42 LYS H    H   8.257 0.020 1
411 42 42 LYS HA   H   4.036 0.020 1
412 42 42 LYS HB2  H   2.006 0.020 2
413 42 42 LYS HB3  H   2.006 0.020 2
414 42 42 LYS HE2  H   2.948 0.020 2
415 42 42 LYS HE3  H   2.948 0.020 2
416 42 42 LYS HG2  H   1.542 0.020 2
417 42 42 LYS HG3  H   1.542 0.020 2
418 42 42 LYS C    C 178.600 0.400 1
419 42 42 LYS CA   C  57.040 0.400 1
420 42 42 LYS CB   C  31.700 0.400 1
421 42 42 LYS N    N 119.296 0.400 1
422 43 43 GLU H    H   7.979 0.020 1
423 43 43 GLU HA   H   4.070 0.020 1
424 43 43 GLU HB2  H   2.010 0.020 2
425 43 43 GLU HB3  H   2.010 0.020 2
426 43 43 GLU HG2  H   2.370 0.020 2
427 43 43 GLU HG3  H   2.370 0.020 2
428 43 43 GLU C    C 178.200 0.400 1
429 43 43 GLU CA   C  58.960 0.400 1
430 43 43 GLU CB   C  28.940 0.400 1
431 43 43 GLU N    N 119.676 0.400 1
432 44 44 GLU H    H   8.278 0.020 1
433 44 44 GLU HA   H   4.069 0.020 1
434 44 44 GLU HB2  H   2.159 0.020 2
435 44 44 GLU HB3  H   2.159 0.020 2
436 44 44 GLU HG2  H   2.366 0.020 2
437 44 44 GLU HG3  H   2.366 0.020 2
438 44 44 GLU C    C 178.200 0.400 1
439 44 44 GLU CA   C  58.960 0.400 1
440 44 44 GLU CB   C  28.940 0.400 1
441 44 44 GLU N    N 118.681 0.400 1
442 45 45 ALA H    H   8.160 0.020 1
443 45 45 ALA HA   H   4.079 0.020 1
444 45 45 ALA HB   H   1.547 0.020 1
445 45 45 ALA C    C 179.200 0.400 1
446 45 45 ALA CA   C  54.910 0.400 1
447 45 45 ALA CB   C  17.360 0.400 1
448 45 45 ALA N    N 121.693 0.400 1
449 46 46 GLU H    H   8.106 0.020 1
450 46 46 GLU HA   H   4.083 0.020 1
451 46 46 GLU HB2  H   2.164 0.020 2
452 46 46 GLU HB3  H   2.164 0.020 2
453 46 46 GLU HG2  H   2.415 0.020 2
454 46 46 GLU HG3  H   2.415 0.020 2
455 46 46 GLU C    C 178.200 0.400 1
456 46 46 GLU CA   C  58.170 0.400 1
457 46 46 GLU CB   C  28.920 0.400 1
458 46 46 GLU N    N 117.210 0.400 1
459 47 47 LYS H    H   7.905 0.020 1
460 47 47 LYS HA   H   4.106 0.020 1
461 47 47 LYS HB2  H   2.165 0.020 2
462 47 47 LYS HB3  H   2.165 0.020 2
463 47 47 LYS HD2  H   1.710 0.020 2
464 47 47 LYS HD3  H   1.710 0.020 2
465 47 47 LYS HE2  H   3.005 0.020 2
```

```
207 22 22 GLU H    H   8.123 0.020 1
208 22 22 GLU HA   H   4.166 0.020 1
209 22 22 GLU HB2  H   2.170 0.020 2
210 22 22 GLU HB3  H   2.170 0.020 2
211 22 22 GLU HG2  H   2.363 0.020 2
212 22 22 GLU HG3  H   2.363 0.020 2
213 22 22 GLU C    C 178.200 0.400 1
214 22 22 GLU CA   C  58.840 0.400 1
215 22 22 GLU CB   C  28.900 0.400 1
216 22 22 GLU N    N 119.363 0.400 1
217 23 23 ALA H    H   8.251 0.020 1
218 23 23 ALA HA   H   4.109 0.020 1
219 23 23 ALA HB   H   1.511 0.020 1
220 23 23 ALA C    C 179.200 0.400 1
221 23 23 ALA CA   C  55.100 0.400 1
222 23 23 ALA CB   C  16.900 0.400 1
223 23 23 ALA N    N 121.717 0.400 1
224 24 24 ASN H    H   8.228 0.020 1
225 24 24 ASN HA   H   4.460 0.020 1
226 24 24 ASN HB2  H   2.873 0.020 2
227 24 24 ASN HB3  H   2.873 0.020 2
228 24 24 ASN HD21 H   7.477 0.020 2
229 24 24 ASN HD22 H   6.671 0.020 2
230 24 24 ASN C    C 177.200 0.400 1
231 24 24 ASN CA   C  55.760 0.400 1
232 24 24 ASN CB   C  38.330 0.400 1
233 24 24 ASN N    N 114.245 0.400 1
234 24 24 ASN ND2  N 110.540 0.400 1
235 25 25 LYS H    H   7.888 0.020 1
236 25 25 LYS HA   H   4.102 0.020 1
237 25 25 LYS HB2  H   2.111 0.020 2
238 25 25 LYS HB3  H   2.009 0.020 2
239 25 25 LYS HD2  H   1.764 0.020 2
240 25 25 LYS HD3  H   1.764 0.020 2
241 25 25 LYS HE2  H   2.935 0.020 2
242 25 25 LYS HE3  H   2.935 0.020 2
243 25 25 LYS HG2  H   1.595 0.020 2
244 25 25 LYS HG3  H   1.595 0.020 2
245 25 25 LYS C    C 178.600 0.400 1
246 25 25 LYS CA   C  59.390 0.400 1
247 25 25 LYS CB   C  31.430 0.400 1
248 25 25 LYS N    N 121.836 0.400 1
249 26 26 ILE H    H   7.949 0.020 1
250 26 26 ILE HA   H   3.828 0.020 1
251 26 26 ILE HB   H   2.015 0.020 1
252 26 26 ILE HD1  H   0.938 0.020 1
253 26 26 ILE HG12 H   1.572 0.020 2
254 26 26 ILE HG13 H   1.774 0.020 2
255 26 26 ILE HG2  H   1.212 0.020 1
256 26 26 ILE C    C 179.000 0.400 1
257 26 26 ILE CA   C  64.310 0.400 1
258 26 26 ILE CB   C  37.420 0.400 1
259 26 26 ILE N    N 119.989 0.400 1
```

```
466 47 47 LYS HE3  H   3.005 0.020 2
467 47 47 LYS C    C 178.600 0.400 1
468 47 47 LYS CA   C  57.350 0.400 1
469 47 47 LYS CB   C  32.200 0.400 1
470 47 47 LYS N    N 119.153 0.400 1
471 48 48 ARG H    H   7.846 0.020 1
472 48 48 ARG HA   H   4.234 0.020 1
473 48 48 ARG HB2  H   1.950 0.020 2
474 48 48 ARG HB3  H   1.950 0.020 2
475 48 48 ARG HD2  H   3.207 0.020 2
476 48 48 ARG HD3  H   3.207 0.020 2
477 48 48 ARG HG2  H   1.728 0.020 2
478 48 48 ARG HG3  H   1.728 0.020 2
479 48 48 ARG C    C 178.600 0.400 1
480 48 48 ARG CA   C  57.350 0.400 1
481 48 48 ARG CB   C  29.640 0.400 1
482 48 48 ARG N    N 117.908 0.400 1
483 49 49 LYS H    H   7.772 0.020 1
484 49 49 LYS HA   H   4.177 0.020 1
485 49 49 LYS HB2  H   1.922 0.020 2
486 49 49 LYS HB3  H   1.922 0.020 2
487 49 49 LYS HG2  H   1.520 0.020 2
488 49 49 LYS HG3  H   1.520 0.020 2
489 49 49 LYS C    C 178.600 0.400 1
490 49 49 LYS CA   C  57.350 0.400 1
491 49 49 LYS CB   C  32.120 0.400 1
492 49 49 LYS N    N 119.320 0.400 1
493 50 50 ALA H    H   7.785 0.020 1
494 50 50 ALA HA   H   4.365 0.020 1
495 50 50 ALA HB   H   1.480 0.020 1
496 50 50 ALA C    C 179.200 0.400 1
497 50 50 ALA CA   C  52.270 0.400 1
498 50 50 ALA CB   C  18.650 0.400 1
499 50 50 ALA N    N 121.668 0.400 1
500 51 51 GLU H    H   7.924 0.020 1
501 51 51 GLU HA   H   4.343 0.020 1
502 51 51 GLU HB2  H   2.160 0.020 1
503 51 51 GLU HB3  H   1.978 0.020 1
504 51 51 GLU HG2  H   2.354 0.020 1
505 51 51 GLU HG3  H   2.354 0.020 1
506 51 51 GLU C    C 178.200 0.400 1
507 51 51 GLU CA   C  56.370 0.400 1
508 51 51 GLU CB   C  29.790 0.400 1
509 51 51 GLU N    N 118.979 0.400 1
510 52 52 ILE H    H   7.431 0.020 1
511 52 52 ILE HA   H   4.108 0.020 1
512 52 52 ILE HB   H   1.875 0.020 1
513 52 52 ILE HD1  H   0.923 0.020 1
514 52 52 ILE HG12 H   1.473 0.020 2
515 52 52 ILE HG13 H   1.473 0.020 2
516 52 52 ILE HG2  H   1.168 0.020 1
517 52 52 ILE C    C 179.000 0.400 1
518 52 52 ILE CA   C  62.830 0.400 1
519 52 52 ILE CB   C  39.260 0.400 1
520 52 52 ILE N    N 124.557 0.400 1
```

Author's publications during Ph.D period January 2006 to June 2010

1. Biuković, G., M. Rössle, **S. Gayen**, Y. Mu and G. Grüber (2007). *Small-Angle X-ray Scattering Reveals the Solution Structure of the Peripheral Stalk Subunit H of the A_1A_O ATP Synthase from Methanocaldococcus jannaschii and Its Binding to the Catalytic A Subunit*. Biochemistry **46**: 2070-2078.
2. **Gayen, S***., S. Vivekanandan*****, G. Biuković, G. Grüber, and H. S. Yoon (2007). *Backbone ^{1}H, ^{13}C, and ^{15}N resonance assignments of subunit F of the A_1A_O ATP synthase from Methanosarcina mazei Gö1*. Biomol. NMR Assign. **1**: 23-25. (*****Authors have equal contribution)
3. **Gayen, S***., S. Vivekanandan*****, G. Biuković, G. Gruber and H. S. Yoon (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. (*****Authors have equal contribution)
4. **Gayen, S**., A. M. Balakrishna, G. Biuković, W. Yulei, C. Hunke and G. Grüber (2008). *Identification of critical residues of subunit H in its interaction with subunit E of the A-ATP synthase from Methanocaldococcus jannaschii*. FEBS J. **275**: 1803-1812.
5. Rishikesan, S., R. Y. Thaker, R. Priya, **S. Gayen,** M. S. S. Manimekalai, C. Hunke and G. Grüber (2008). *Spectroscopical identification of residues of subunit G of the yeast V-ATPase in its connection with subunit E*. Mol. Mem. Biol. **25**: 400-410.
6. Priya, R., V. Tadwal, M. Rössle, **S. Gayen**, C. Hunke, W. Peng, J. Torres and G. Grüber (2008). *Low resolution structure of subunit b ($b_{22–156}$) of Escherichia coli F_1F_O ATP synthase in solution and the b –δ assembly*. J. Bioenerg. Biomembr. **40**: 245-255.
7. Rishikesan, S., **S. Gayen,** R. Y. Thaker, S. Vivekanandan, M. S. S. Manimekalai, Y. H. Yau, S. Greifman Shochat and G. Grüber (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.
8. Biuković, G*****., **S. Gayen***, K. Pervushin and G. Grüber (2009). *The domain features of the peripheral stalk subunit H of the methanogenic A_1A_O ATP synthase and the NMR solution structure of H_{1-47}*. Biophys. J. **97**: 286-294. (*****Authors have equal contribution)

9. **Gayen, S**., A. M. Balakrishna and G. Grüber (2009). *NMR solution structure of the N–terminal domain of subunit E (E_{1-52}) of A_1A_O ATP synthase from Methanocaldococcus jannaschii*. J. Bioenerg. Biomembr. **41**: 342-348.
10. Priya, R., G. Biuković, **S. Gayen**, S. Vivekanandan and G. Grüber (2009). *NMR solution structure of the b30-82 domain of subunit b of Escherichia coli F_1F_O ATP synthase*. J. Bacteriol. **191**: 7538-7544.
11. **Gayen, S** and Grüber G (2009). *Disulfide linkage in the coiled coil domain of subunit H of A_1A_O ATP synthase from Methanocaldococcus jannaschii and the NMR structure of C-terminal segment H_{85-104}*. FEBS Lett. **584**: 713-718.
12. Devanathan, R, **S. Gayen**, G. Grüber and C. S. Verma (2009). *Cross-talk along the stalk: Dynamics of the interaction of Subunit B and Subunit F in A_1A_O ATP Synthase of Methanosarcina mazei* Gö1. Biochemistry. **49**: 4181-4190.
13. Balakrishna, A. M., M.S.S. Manimekalai, C. Hunke, **S. Gayen,** M. Rössle, J. Jeyakanthan, and G. Grüber (2010). *Crystal and solution structure of the C-terminal part of the Methanocaldococcus jannaschii A_1A_O ATP synthase subunit E revealed by X-ray diffraction and small-angle X-ray scattering.* J. Bioenerg. Biomembr., in press.

Conference presentation

1. Joint Third AOHUPO and Fourth Structural Biology and Functional Genomics Conference NUS, Singapore from 4-7th Dec 2006.

   Title: Solution structure of subunit F of A_1A_O ATP synthase from *Methanosarcina mazei* Gö1.

2. 16th triennial conference for the international society of magnetic resonance (ISMAR 2007) Kenting, Taiwan from 14th – 19th October 2007, pp218.

   Title: Structural features of subunit H of A_1A_O ATP synthase in solution.