STRUCTURAL STUDIES OF SELECTED MITOCHONDRIAL COMPLEX-I SUBUNITS

THESIS SUBMITTED TO
UNIVERSITY OF PUNE

FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN BIOTECHNOLOGY

BY

TULIKA JAOKAR

RESEARCH GUIDE
Dr. C.G. SURESH

RESEARCH CO-GUIDE
Dr Y.S. SHOUCHE

DIVISION OF BIOCHEMICAL SCIENCES
CSIR-NATIONAL CHEMICAL LABORATORY
DR HOMI BHABHA ROAD, PUNE 411008
MAHARASHTRA, INDIA

MAY 2014

Dedicated to
Papa and Mummy

CONTENTS

LIST OF TABLES

LIST OF FIGURES

सीएसआयआर–राष्ट्रीय रासायनिक प्रयोगशाला
(वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद)
डॉ. होमी भाभा रोड, पुणे – 411 008. भारत
CSIR-NATIONAL CHEMICAL LABORATORY
(Council of Scientific & Industrial Research)
Dr. Homi Bhabha Road, Pune - 411008. India

CERTIFICATE

This is to certify that the work incorporated in the thesis "**Structural studies of selected mitochondrial Complex-I subunits**" submitted **by Ms. Tulika Milind Jaokar** was carried out by the candidate under our supervision/guidance. Such material obtained from other sources has been duly acknowledged in the thesis.

Dr. C.G. Suresh
(Research Supervisor)
Chief Scientist

Dr Y.S. Shouche
(Research Co-Guide)
Scientist G

Date:

x

Communications Channels
☎ Fax
NCL Level DID : 2590
NCL Board No. : +91 20 25902000
EPABX : +91 20 25893300
: +91 20 25893400
FAX
Director's Office : +91 20 25902601
COA's Office : +91 20 25902660
COS&P's Office : +91 20 25902664
WEBSITE
www.ncl-india.org

DECLARATION BY THE CANDIDATE

I hereby declare that the thesis entitled "**Structural studies of selected mitochondrial Complex-I subunits**" submitted by me for the degree of Doctor of Philosophy is the record of work carried out by me during the period from November 2010 to May 2014 under the guidance of **Dr. C.G. Suresh, Chief Scientist** and has not formed the basis for the award of any degree, diploma, associateship, fellowship, titles in this or any other University or other institution of higher learning.

I further declare that the material obtained from other sources has been duly acknowledged in the thesis.

Tulika Milind Jaokar

Division of Biochemical Sciences
National Chemical Laboratory
Pune - 411 008
May 2014

"All our dreams can come true –if we have the courage to pursue them" - Walt Disney

ACKNOWLEDGEMENT

This piece of work is a combined effort of many people. I was extremely fortunate to work with some of the greatest minds in CSIR-National Chemical Laboratory which not only helped me advance my knowledge but also helped me develop as a person. I would like to thank my research guide Dr C.G. Suresh for his constant guidance and support throughout my 5 years as a PhD student in NCL, Pune. I'm really grateful to him for giving me freedom to think, plan and execute my work. He has not only taught me good work ethics but also has been a source of inspiration to do work in a disciplined and timely manner. His helpful nature and keen interest in my work helped me try new things without the fear of failure.

I would also like to thank my Co-Guide Dr Yogesh Shouche. I started the initial part of my work in his lab, Lab#3, National Centre for Cell Science, Pune. He gave me a free hand at utilizing the resources in his laboratory including the sequencing facility which gave an impetus to my work. His inputs during my upgradation have surely helped me improvise on my work. His student, Dr Deepak Patil, helped in the very initial stages of the cloning and site-directed mutagenesis. His meticulous work practices, enthusiasm and speed of work gave a momentum to my work. He taught me many new techniques and protocols and his inquisitive nature helped me explore many new aspects giving new dimensions to the work.

I will always be grateful to Dr Sushama M Gaikwad. She has not only introduced me to new spectroscopic techniques but also taught me to plan and execute my experiments. Madam always had time to discuss experiments and gave many new insights while analyzing data. Her

enthusiasm is infectious. She along with her students; Madhurima, Sonali, Sayli, Priya and Ekta have been very helpful. I'm grateful to Dr Siddharth Bhosale for his help with the time-resolved fluorescence instrument.

Special thanks are due for Dr Arvind Sahu and Dr Moneesha Fernandes for allowing the use of the Circular dichroism facility in NCCS and NCL, Pune respectively.

I really appreciate and am grateful for the timely and invaluable advice given by Dr Sureshkumar Ramasamy. I would always remember the company I had of the fellow research scholars in NCL. I'm grateful to all my seniors and other lab members Poorva, Urvashi, Nishant, Manas, Payal, Priyabrata, Ranu, Ruby, Prachi, Deepak, Manu, Ameya, Shridhar, Aditi, Tejashree, Deepanjan, Yashpal and Vijay for providing a great working environment in the laboratory. Special thanks are due for Ranu who helped me with the bioinformatics work and gave a lot of moral support during tough times. I will always be grateful to Manas for his constant support and advices from time to time. I would always remember the project trainees who were a part of our group for a short period of time; Mithila, Shaheen, Sharad, Mahesh, Milan. Their constant enthusiasm and zeal to learn new things was a source of inspiration.

I would like to thank Director, National Chemical Laboratory (NCL, Pune) for giving me the opportunity to work in this great institution. I'm grateful to Dr Vidya Gupta, Head, Division of Biochemical Sciences, NCL, Pune and Dr J.K. Pal, Head, Department of Biotechnology, University of Pune, for being members of my evaluation committee. Their inputs have been invaluable. I would also like to thank Indira Mam, Sheetal ,Yashashree and the SAC office members Puranik mam and Kolhe mam for their help in administrative work.

My friends Preeti, Mugdha, Ketaki, Agraja, Sneha and Vibhusha have always been there for me. Their support and friendship is invaluable. Finally I would like to thank my family Aaji, Papa, Mummy, Parth, Meena Atu and Deepak Kaka for their constant support. They have been extremely patient through all my highs and lows and provided me great encouragement. They' always encouraged my dreams and will always be my strength. I would also like to thank my fiancé Rohit and his parents for their love and understanding. Their encouraging words and belief in me have helped me reach my goals.

Finally I would like to thank all those who have directly or indirectly supported me.

Tulika Jaokar

ABSTRACT

Mitochondrial Complex-I (NADH: ubiquinone oxidoreductase) is the protein complex that carries out the first step in the respiratory chain process. During cellular respiration Complex-I transfers two electrons from NADH to ubiquinone with simultaneous pumping of four protons across the inner mitochondrial membrane, generating a proton motive force driving ATP synthase to synthesize ATP. Being a large complex structure it has many constituent subunits, 44 in eukaryotes and 14 in prokaryotes. This multi subunit complex is divided into 3 domains: dehydrogenase domain (N module), hydrogenase/ ubiquinone reduction domain (Q module) and proton pumping domain (P module). Determining the atomic resolution structure of the entire Complex-I is a daunting task owing to its large size. Due to the absence of information on subunit organization; Complex-I was considered an L-shaped black box. Only recently the three-dimensional structure of a prokaryotic Complex-I from *Thermus thermophilus* has become available (PDB ID: 2FUG, 4HE8, 4HEA). This structure is a 550 kDa complex consisting of 14 subunits conserved across organisms and some additional accessory subunits. Owing to poor resolution the individual subunits are not distinguishable in the available eukaryotic Complex-I structure from yeast. Thus, further exploration is required to understand eukaryotic complex, in particular the human mitochondrial Complex-I.

The subunits of human Complex-I are known mutational hot spots for several disorders. The mutations may be mitochondrial or nuclear in origin resulting in genetic diseases which may be inherited maternally or in an autosomal recessive fashion. The mutations result in a dysfunctional Complex-I causing Complex-I deficit disorders. The symptoms observed range from encephalopathy, cardiomyopathy, neurological disorders or a combination of two or more of these symptoms. Leigh syndrome is one such genetic neurological disorder caused by mutations of Complex-I. Two among every 1,00,000 newborns is affected by this disorder. There is also speculation about the damage that can be caused by reactive oxygen species (ROS) produced when Complex-I is not properly assembled. Only correctly assembled hydrophilic domain can prevent ROS production. Thus, these disorders are thought to be due to disruption of Complex-I assembly by mutations in subunits affecting function and causing damage due to ROS generation. Analyzing the differences

between wild-type (w-t) and mutant subunits would be a step towards understanding at molecular level the potential defects in complex assembly caused by pathogenic mutations and its resultant dysfunction. It would also help associating the broad nature of symptoms specifically to functional deficit of Complex-I in the diseases like Leigh syndrome.

In the present thesis, *in silico* and biophysical approaches have been adopted to study some of the core subunits of the ubiquinone reduction module (Q module) of the human mitochondrial Complex-I. The Q module possesses the iron sulphur cluster N2 vital for its activity. The module acts as a connecting link between the N and P module and assembles during the early stages of Complex-I assembly to form the peripheral arm. The core subunits of Q module are *NDUFS2, 3, 7 & 8*. These subunits are highly conserved in bacteria to mammals; especially their C-terminal part shows an identity of 42% in certain cases. Utilizing the similarity between the human and *Thermus thermophilus* Complex-I subunits, the four core human subunits in the Q module were modelled using molecular modelling based on the *T thermophilus* structure (3I9V, 3IAM, 3FUG) in MODELLER9.10 and PRIME v3.1 of the Schrodinger suite. The mutants causing Leigh syndrome were prepared *in silico* using the FoldX program. The mutations were then classified into those in the interior, those at the interface and those near the binding site depending on their positions in the Q module.

Molecular dynamic simulations were performed in GROMACS v4.5 to understand the effects of these double/point mutations on the structure. The w-t and mutant structures were compared in terms of root mean square fluctuation (RMSF) of residues, radius of gyration (Rg), hydrogen bonding, solvent accessibility etc along a 15 ns trajectory. The substrate n-decyl ubiquionone (n-DBQ) was docked at the interface of *NDUFS2 & 7* using Glide v5.8. Energy calculations were performed in Bioluminate 1.0 to study the effects of subunit mutations on the stability of the Q module, iron sulphur cluster, mutual affinities of subunits and affinity to n-DBQ.

In the dynamic simulation studies *NDUFS2* showed a high RMSF mainly at the location of mutation, whereas *NDUFS3, 7 & 8* showed fluctuations spread over the whole protein. In the case of *NDUFS3 & 8* the w-t proteins were more compact (measured by Rg) compared to mutant proteins. All the mutant subunits were found energetically unstable compared to their w-t counterparts. Some of them showed

reduced affinity to n-DBQ or iron sulphur cluster N2. Although the mutations caused minimum perturbation to the overall secondary structure, the root mean square fluctuations of certain segments were substantial, especially those in the loop regions. The mutations affected the hydrogen bonded interactions, solvent accessible area and the observed radius of gyration to various extents. Molecular dynamic simulations of the *NDUFS2* and *7* complex along with n-DBQ revealed the reduced affinity of the mutants towards n-DBQ. Also, high fluctuations of residues near the n-DBQ binding site was observed in the mutants during MD simulations; indicating that even when the mutation positions are not in the vicinity of binding sites, they certainly influence the binding.

The cumulative effect of all the changes due to mutations on the formation of complex assembly is reflected in the observed unfavourable energy variations, instability of subunit association and a reduced affinity for substrates such as ubiquinone. Thus the analysis indicates that Leigh syndrome mutations lead to formation of structurally and functionally defective complex, which in turn results in disease phenotype. Further attempt was made to observe the nature of the subunits *in vitro* to extrapolate the *in silico* observations in solution.

Although we tried to clone and express all the subunits of Q module, finally we were successful in completing studies only on two subunits, the structural subunit *NDUFS3* and the functional subunit *NDUFS7*, which were cloned and expressed via autoinduction in *E coli*. The full length genes of both the subunits were amplified from the cDNA obtained from RNA of the *HT 29* cell line. The genes were further cloned into pGEM-T vector followed by subcloning in pET28b(+) expression vector between Nde I and Xho I sites and transformed in *BL21DE3* expression cells . The subunits engineered by incorporating mutations identified with Leigh syndrome were prepared by site-directed mutagenesis. A double mutant of *NDUFS3* (T145I+R199W) was prepared by a two step PCR protocol. Two point mutants of *NDUFS7*: V122M and R145H were prepared in a single PCR step. The products of site-directed mutagenesis were ligated and further subcloned in expression vector pET28(+) between Nde I and Xho I sites. The w-t proteins and corresponding mutants were further expressed by autoinduction and purified. *NDUFS3* and its mutant were purified by Q-Sepharose ion-exchange chromatography followed by Sephadex G-200 size exclusion chromatography. *NDUFS7* and its mutants were purified by Ni-NTA

affinity chromatography followed by Superose 12 size exclusion chromatography. The expression and purity of proteins were confirmed by 12% SDS-PAGE and Western blot; molecular weights were determined using MALDI TOF/TOF.

The w-t and mutant proteins have been compared for their stability, aggregation propensities, binding with n-decylubiquinone and other biophysical properties. The mutant proteins of both the subunits showed differences with respect to the corresponding w-t counterparts in secondary and tertiary structure and higher aggregation propensity. Apart from these, mutant *NDUFS3* showed a different native structure as compared to the w-t although the conformational transitions at different pH, temperature and Gdn-HCl concentrations were similar. The tryptophan microenvironment and secondary and tertiary structure of w-t *NDUFS3* were substantially different from that of the double mutant. Despite the point mutants of *NDUFS7* showing a structure similar to the native w-t *NDUFS7* the peak at 325 nm characterising cluster N2 was absent in them. Both the mutants of *NDUFS7* showed reduced binding affinity towards n-DBQ. In the w-t, an increase in the binding affinity to n-DBQ was observed with an increase in temperature, however, the mutants showed an opposite trend of decrease in binding affinity with temperature. Rayleigh scattering, Thioflavin-T and Congo red dye binding studies helped to compare and quantify the differences in aggregation tendency of the proteins. The mutants showed a high aggregation propensity in both *NDUFS3* and *7* as compared to the w-t. At physiological temperature (37°C), an increase in aggregation was observed only in the mutant proteins. A higher aggregation propensity of the mutants would lead to lesser number of molecules available for assembly of Complex-I.

These different properties highlighted the change in the functional nature of the mutant proteins *in vitro*. These changes have been correlated with complex functional defects or complex disassembly observed in Leigh syndrome. Using computational and biophysical approaches, the structural and functional defects in the mutant subunits have been demonstrated. This has contributed to understanding molecular level effects of pathological mutations causing Leigh syndrome.

ABBREVIATIONS

Abbreviation	Long form
ANS	8-anilino-1-naphthalene sulfonic acid
ATP	Adenosine 5'-triphosphate
BN-PAGE	Blue Native Polyacrylamide Gel Electrophoresis
bp	Base pair
CD	Circular dichroism
cDNA	complementary DNA
CR	Congo red
CsCl	Caesium chloride
CSF	Cerebrospinal fluid
DTT	Dithiothreitol
EDTA	Ethylene-diamine-tetra-acetate
Fe-S	Iron-sulphur cluster
FMN	Flavin mononucleotide
FP	Flavoprotein
Gdn-HCl	Guanidium- Hydrochloride
IP	Iron-sulphur protein
KI	Potassium iodide
LHON	Leber's Heriditary optic neuropathy
MALDI-TOF	Matrix-assisted laser-desorption ionization-time of flight
MD	Molecular dynamics
MELAS	Mitochondrial encephalopathy, lactic acidosis and stroke like episodes
mtDNA	Mitochondrial DNA
NAD	Nicotinamide adenine dinucleotide
n-DBQ	n-decyl-ubiquinone
nDNA	Nuclear DNA
NCBI	National Centre for Biotechnology Information
NIDDM	Non insulin dependent diabetes mellitus
NDUFS	NADH dehydrogenase (ubiquinone) iron sulphur (Fe-S)
NMR	Nuclear magnetic resonance
ns	Nanosecond
ORF	Open reading frame
OXPHOS	Oxidative phosphorylation
PBS	Phosphate buffer saline

PCR	Polymerase Chain reaction
ps	Picoseconds
PVDF	Polyvinylidene fluoride
Rg	Radius of gyration
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
ROS	Reactive oxygen species
SDM	Site-directed mutagenesis
SDS	Sodium dodecyl sulphate
Th-T	Thioflavin-T
Unord	Unordered region
woFe-S	Without iron-sulphur cluster
w-t	Wild type

Chapter 1

Introduction to mitochondrial function and genetic disease

Mitochondria are eukaryotic cellular organelles bound by a double membrane and possessing their own circular genome (mt-DNA). They are sometimes referred to as "cellular power plants" as they generate most of the cells' supply of adenosine 5' triphosphate (ATP) used as a source of chemical energy by the cells. The word mitochondrion was coined by Carl Benda in 1898 and comes from the Greek word *mitos*, i.e. thread and *chondrion*, i.e. granule. They are present in all eukaryotes, however, their number varies in different organisms and tissues. Their size ranges from 0.5 to 1.0 μm in diameter. Every mitochondrion is composed of 4 compartments that carry out specialized functions; outer membrane, inner membrane, cristae and matrix (Figure 1.1).

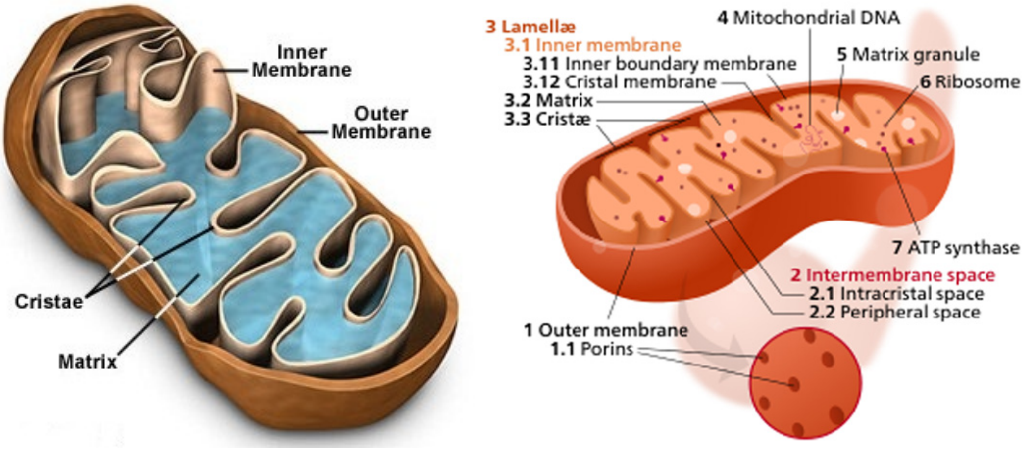


Figure 1.1: Structural features of mitochondria. Reproduced from Molecular Expressions™ Cell Biology: mitochondria and Wikipedia, the free encyclopedia.

Mitochondria play a pivotal role in generating 90% of the cellular energy in the form of ATP by oxidative phosphorylation (OXPHOS). Along with energy generation, they are also involved in apoptosis, free radical production, calcium signalling, intermediary metabolism and cellular homeostasis [Rossignaol *et al*, 2003; Wallace, 2013].

The mitochondria utilize five multi-domain enzyme complexes present on the inner mitochondrial membrane for OXPHOS to synthesize energy in the form of ATP (Figure 1.2). The complexes are:

1. Complex-I or NADH dehydrogenase; EC 1.6.5.3
2. Complex-II or succinate dehydrogenase; EC 1.3.5.1
3. Complex-III or cytochrome bc_1 complex; EC 1.10.2.2
4. Complex IV or cytochrome c oxidase; EC 1.9.3.1

5. Complex V or ATP synthase; EC 3.6.3.14

The first four complexes are involved in the step wise transfer of electrons to the ultimate electron acceptor oxygen, simultaneously pumping protons across the membrane generating a proton motive force which drives ATP synthase to synthesize ATP.

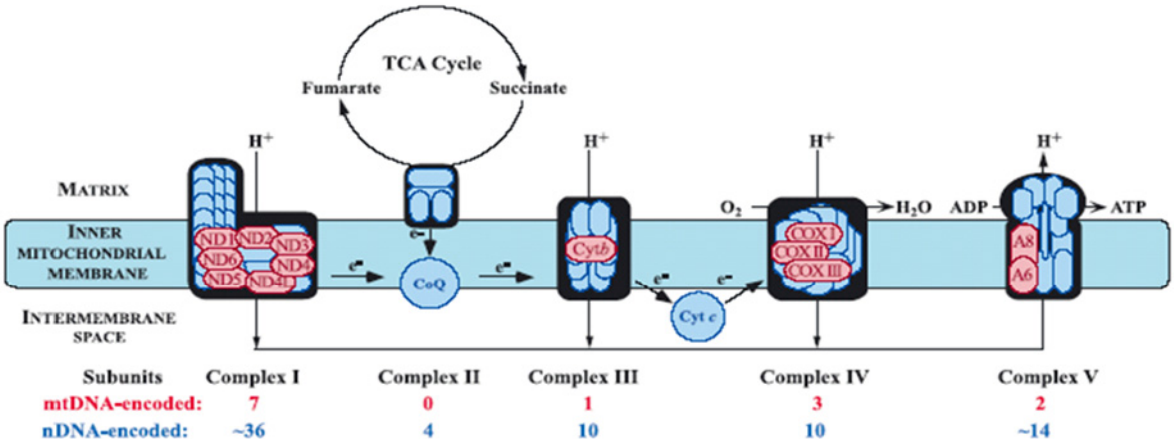


Figure 1.2: Schematic of the mitochondrial respiratory electron transport chain showing the 5 complexes. Reproduced from Dimauro *et al*, 2003.

1.1 Mitochondrial genome

Mitochondria possess their own circular DNA. It is widely accepted that mitochondria evolved from bacteria that were engulfed by nucleated ancestral cells. Thus, they possess their own genome (mtDNA) as well as their own biosynthetic machinery for making RNA and organelle proteins.

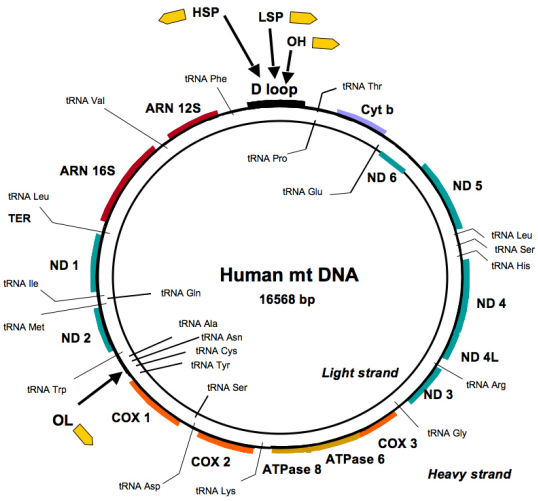


Figure 1.3: Human mtDNA map. ND1-6, NADH dehydrogenase subunits; COX 1-3, Cytochrome c oxidase subunits, ATP6 and ATP8, subunits of ATP synthase, Cyt b, Cytochrome b. Reproduced from Bellance *et al*, 2009.

The human mtDNA is 16,568 bp in size (Figure 1.3). It is a double stranded circular molecule coding for 13 polypeptides of the mitochondrial respiratory chain, 22 tRNA's and 2 rRNA. The non coding regulatory region also known as the D loop is of ~600 bp. HSP and LSP are the heavy and light strand promoters, respectively and OH is a purine rich region which is the origin of replication. Transcription is polycistronic and translation has a different codon usage than the universal codons. For amino acids Met, Trp and for termination the codons are different than the universal ones (Table 1.1). The remaining amino acids have the standard codons.

Codon	Translation	
	Universal	Mitochondria
AGA	Arg (R)	Ter (*)
AGG	Arg (R)	Ter (*)
AUA	Ile (I)	Met (M)
UGA	Ter (*)	Trp (W)

Table 1.1: Universal codon usage versus codon usage in mitochondria.

The nuclear genome codes for the remaining proteins of the mitochondrial respiratory chain, metabolic enzymes, DNA/RNA polymerase, ribosomal proteins, mtDNA regulatory factors and mitochondrial transcription factor A (Tfam). Thus, the mitochondrial and nuclear genomes co-operate and co-ordinate for the smooth functioning of the OXPHOS. Since the mitochondrial DNA was discovered in 1963 and its entire genome sequenced by 1981, a large number of studies on migrations of populations were reported, and several genetic disorders caused by mutations in mitochondrial DNA were also defined.

1.1.1 Population studies

Population studies involved tracing back the origin, pre-historic migrations, haplotypes, genetic relationships and genetic variations of various population groups. Since the mtDNA is maternally inherited with no recombination events, migration and genetic relationships of various populations could be mapped by restriction site mapping and D loop sequencing followed by analysis of the phylogentic networks. Such studies were carried out in various ethnic groups such as Africans, Caucasians, Chinese, Indians, South-East Asians, Australian tribes etc. These studies lead to the identification of several population-specific mtDNA polymorphisms which were further linked to the predisposition of a broad range of metabolic and degenerative

diseases. These variants were ancient and have accumulated along with radiating maternal lineages during the human expansion out of Africa [Wallace, 2013].

1.1.2 Genetic disorder studies

Mitochondrial genetic diseases are a group of, often untreatable, disorders affecting the eyes, nervous system, heart, muscles etc occurring due to mutations in the mtDNA or nDNA. Thornburn [Thornburn, 2004] defined the minimum prevalence of such mitochondrial disorders as 13.1 in 1,00,000 individuals. Of which the prevalence of primary OXPHOS related disorders is 1 in 7,634 individuals. Thus, the misnomer that mitochondrial diseases are rare is slowly being wiped out. Disease based epidemiological studies estimate the prevalence of mitochondrial disease to be ~1 in 5000. This number however doesn't take into account asymptomatic carriers or new mutations that may cause disease [Elliott *et al*, 2008; Thorburn *et al*, 2004]. The clinical presentation of these diseases is diverse. They may affect children as well as adults and are progressive and systemic in nature.

Since the first report about pathogenic mutation in mitochondrial DNA in 1988, a large number of primary mutations causing diseases have been described. Since, both mitochondrial as well as nuclear genome code for the proteins in OXPHOS; disorders related to OXPHOS proteins may be due to either mutations in mtDNA or in nDNA or in both. The mtDNA is known to have a mutation rate 10-20 times more than the nDNA. This is due to the absence of protective histones, the poor repair mechanisms and lack of proof reading activity of mtDNA polymerase [Schmiedel *et al*, 2003]. Thus, a large number of mutations accumulate in the mitochondrial genome. These may vary from neutral polymorphisms to pathogenic mutations. A mutation is considered pathogenic when it satisfies certain criteria:

1. It should not be seen in the control population.
2. It should be observed in unrelated pedigrees with similar disease presentation.
3. The nature and location of the disease suggests a logical mechanism of the disease.
4. It should be heteroplasmic.

Neutral mutations go on accumulating in the mtDNA, which is maternally inherited. Since the mtDNA codes for the most important mitochondrial energy genes the effects of mutations in these are lethal. Clinical diagnosis, pedigree analysis,

biochemical and histological analysis, restriction enzyme site polymorphism (RSP) and sequencing of mtDNA may be done to diagnose and establish that the disease has a mitochondrial origin [Taylor *et al*, 2004; Munnich *et al*, 2001]. The mutations hamper the smooth functioning of the complexes or their assembly thereby affecting OXPHOS and cause an energy deficit. Muscles and nerves are the worst hit in the case of such mutations due to their large energy dependency. Hence most of the diseases show up as encephalopathy and cardiomyopathy along with several other associated symptoms.

1.1.3 Human mitochondrial genome database

MITOMAP: **A human mitochondrial genome database** is a compilation of all the genetic variations reported in mtDNA [Kogelnik *et al*, 1995]. It brings together information on mitochondrial genome structure, function, pathogenic mutations, their clinical characteristics, population associated variations and gene-gene interactions. Till date, ~277 disease causing mutations have been reported in the coding and control regions of the mtDNA and 300 mutations in rRNA and tRNA. Approximately 264 DNA rearrangements have been enlisted which include deletions, inversions, insertions and complex rearrangements.

1.2 Mitochondrial proteome

The mitochondrial proteome consists of approximately 1500 gene products. These proteins have a dual genetic origin, either from nDNA or mtDNA. Amongst them 13 polypeptides are encoded by the mtDNA and the rest of the subunits and assembly factors are encoded by the nDNA. Mutations have been reported in the different subunits of the 5 complexes involved in OXPHOS as well as the assembly factors leading to energy deficit and mitochondrial diseases.

1.2.1 Mitochondrial proteome database

MitoP2: Mitochondrial proteome database gives information on mitochondrial proteins, their molecular functions and associated diseases. This database has several useful features like the identification of putative orthologous proteins to study evolutionary conserved functions and pathways, integration of data from systematic genome-wide studies such as proteomics and deletion phenotype screening, prediction of novel mitochondrial proteins and systematic searches to find genes underlying mitochondrial diseases [Prokisch *et al*, 2005].

This thesis focuses on the subunits of the Complex-I and its mutants causing mitochondrial disorders in humans.

1.3 Complex-I

Complex-I or NADH: ubiquinone oxidoreductase or NADH dehydrogenase is the first multi-domain enzyme complex in the electron respiratory chain pathway. Its function is to transfer two electrons from NADH to ubiquinone (Q) through various protein bound prosthetic groups with simultaneous pumping of four protons across the bacterial or inner mitochondrial membrane to generate a proton motive force which activates ATP synthase to synthesize ATP [Ripple *et al*, 2013]. It is also involved in minimizing reactive oxygen species (ROS) production in the cell.

$$NADH + H^+ + Q + 4H^+_{\ in} \longrightarrow NAD^+ + QH_2 + 4H^+_{out}$$

1.3.1 Subunit Composition and Nomenclature

Complex-I is one of the largest multi subunit enzyme complexes comprising of 14 subunits (~550 kDa) in prokaryotes and 45 subunits (~950 kDa) in eukaryotes. Of the 45 subunits constituting this giant complex in eukaryotes, 7 are encoded by the mitochondrial genome and the rest 38 by the nuclear genome. This gives the Complex-I a dual genetic origin, where majority of the subunits are encoded by the nDNA, synthesized in the cytoplasm and transferred to the mitochondria where they get assembled with the mtDNA encoded subunits.

The 45 subunits in eukaryotes are divided into 3 domains which carry out specialized functions; the N module having dehydrogenase activity, the Q module involved in ubiquinone reduction and the P module involved in proton pumping across the inner mitochondrial membrane. Out of these 45 subunits, 14 are core subunits, highly conserved across all species and play a vital role in the functioning of the complex from bacteria to humans. These subunits are considered to be 'minimal' subunits required to build the complex with electron transfer and proton translocation activities. The nomenclature of the subunits across species is enlisted in Table 1.2.

Table 1.2: Nomenclature of the 14 conserved subunits of Complex-I across the N, Q and P module. Adopted from Baradaran *et al*, 2013. Nqo stands for NADH:quinone oxidoreductase. NDU in eukaryotes stands for NADH dehydrogenase ubiquinone and MT-ND mitochondrial NADH dehydrogenase.

Domain	*Escherichia coli*	*Thermus thermophilus*	*Bos taurus*	*Homo sapiens*
		Hydrophilic domain		
N module	Nqo1	NqoF	51 kDa	NDUFV1
	Nqo2	NqoE	24 kDa	NDUFV2
	Nqo3	NqoG	75 kDa	NDUFS1
Q module	Nqo4	NdoD fused with NqoC	49 kDa	NDUFS2
	Nqo5	NqoC fused with NqoD	30 kDa	NDUFS3
	Nqo6	NqoB	PSST	NDUFS7
	Nqo9	NqoI	TKYK	NDUFS8
		Membrane domain		
P module	Nqo7	NqoA	ND3	MT-ND3
	Nqo8	NqoH	ND1	MT-ND1
	Nqo10	NqoJ	ND6	MT-ND6
	Nqo11	NqoK	ND4L	MT-ND4L
	Nqo12	NqoL	ND5	MT-ND5
	Nqo13	NqoM	ND4	MT-ND4
	Nqo14	NqoN	ND2	MT-ND2

Apart from these 14 conserved subunits, ~30 additional accessory subunits present in eukaryotes form a protective shell around the core. They help in the biogenesis of the eukaryotic Complex-I and give structural stability. These are also referred to as 'supernumerary subunits' as they have no homologues in prokaryotes [Baradaran *et al*, 2013].

The arrangement of the 14 conserved subunits in *E coli* and bovine is shown in Figure 1.4.

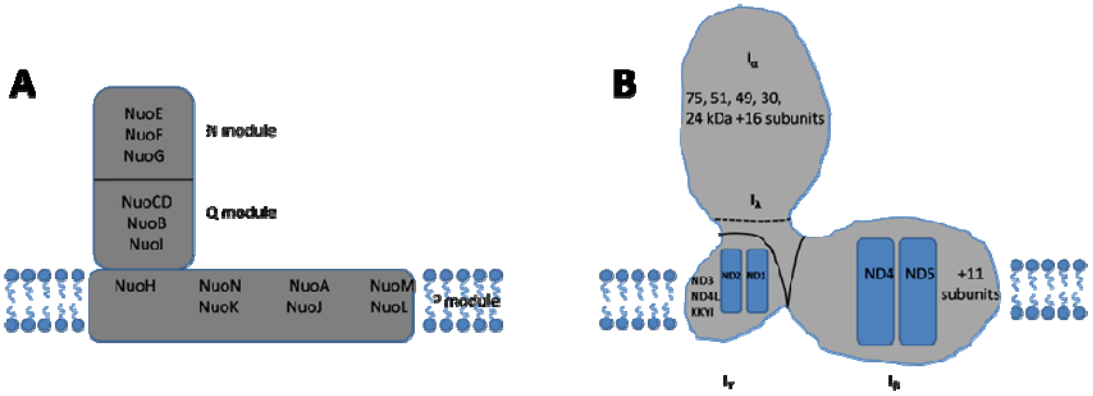


Figure 1.4: Arrangement of the conserved subunits of Complex-I in A: *E. coli* and B: Bovine. Reproduced from Leif *et al*, 1995; Saznov *et al*, 2006; Carroll *et al*, 2006.

1.3.2 Structural studies of Complex-I

Complex-I is one of the largest multi subunit enzyme complexes with a molecular weight of ~550 kDa in prokaryotes and 1 mDa in eukaryotes. The initial studies on the structure of Complex-I involved cryo-electron microscopy to understand the shape of this large enzyme. Reverse phase HPLC, two-dimensional gel analysis with isoelectric focusing and mass spectroscopy were used to determine the composition of the Complex-I especially in eukaryotes. In low resolution cryo-electron microscopic images from *E.coli*, *Neurospora crassa*, *Yarrowia lipolytica* and bovine, Complex-I appears to be an L-shaped structure divided into 3 domains [Grigorieff, 1998; Zickermann *et al*, 2009; Radermacher *et al*, 2006]

1. Dehydrogenase domain (N module)
2. Hydrogenase domain (Q module)
3. Membrane domain (P module)

These three domains co-ordinate and function together maintaining the activity of Complex-I.

In the L-shaped structure of Complex-I, the two arms are arranged almost perpendicular to each other. The dehydrogenase (N) and hydrogenase (Q) domain form the hydrophilic portion protruding out of the membrane into the cytosol. The membrane fraction (P) is embedded in the bacterial or in the inner mitochondrial membrane (Figure 1.5).

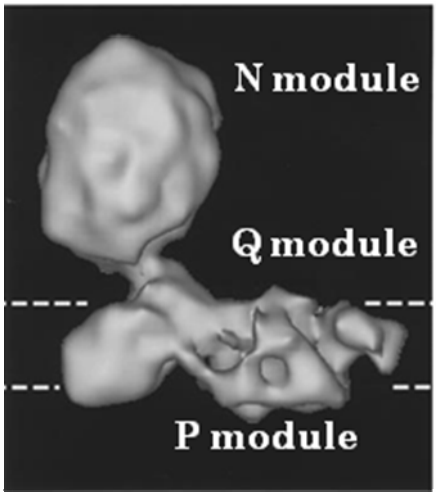


Figure 1.5: Cryo-electron microscopy of the bovine respiratory Complex-I (22Å); L-shaped structure showing the 3 domains. The dotted lines represent the membrane [Grigorieff, 1998]

Mass spectroscopic studies have been done, especially in eukaryotic systems like bovine, to get insights into the nature of large number of subunits in Complex-I.

1.3.2.1 Structure of prokaryotic Complex-I

Owing to the presence of a large number of subunits and the resultant large size of Complex-I, its atomic resolution structure was not successfully determined until recently. The structure of *Thermus thermophilus* Complex-I determined at high resolution for the first time provided insights into both the structure and the functioning of Complex-I in prokaryotes. The 3.3 Å resolution structure (2FMN, 3I9V, 3IAM) of the hydrophilic domain of Complex-I from *T. thermophilus* was published by Saznov *et al* [Saznov *et al*, 2006]. The architecture of the entire complex was determined subsequently at 4.5 Å resolution. However, in this structure only the arrangement of subunits and helices in the membrane domain could be resolved [Efremov *et al*, 2011]. This was followed by a 3.0 Å resolution structure of the membrane domain from *E. coli* Complex-I (3RKO). Three large domains NuoM, L and N were identified, which could function as putative proton translocation channels. The pumping of the 4th proton still remained undetermined as only three channels were defined [Efremov *et al*, 2011]. In 2013, Baradaran *et al* [Baradaran *et al*, 2013] determined the entire structure of respiratory Complex-I at resolution 3.3 Å (4HE8, 4HEA).

The 536 kDa structure of Complex-I comprises of 9 hydrophilic subunits with 9 iron sulphur clusters and FMN along with 7 membrane subunits possessing 64

transmembrane helices. An additional assembly-like factor Nqo16 was also present in the structure (Figure 1.6).

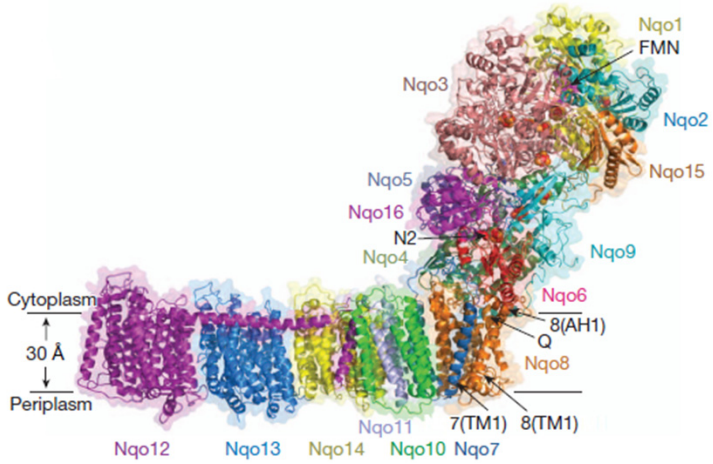


Figure 1.6: Crystal structure of the entire respiratory Complex-I from *Thermus thermophilus*. Reproduced from Baradaran *et al*, 2013.

A. Hydrophilic domain

The structure of the hydrophilic domain was initially described in structure file 2FMN. However, the missing features like loops of Nqo4 present at the interface of hydrophilic and membrane domain were resolved only by the structure of the entire complex. This domain contained eight subunits (Nqo1, 2, 3, 4, 5, 6, 9 and 15), a non-covalently bound FMN, nine iron-sulphur clusters and the NADH binding site. The peripheral arm was a Y shaped assembly made up of Nqo1, 2 and the C-terminal of Nqo3. The N-terminal of Nqo3 acts as a connecting link between the subunits. The base of the domain is formed by Nqo4 and 6 (Figure 1.7A). The FMN is coordinated by Nqo1 at the deep end of a solvent cavity containing the NADH binding site. The pathway defined by NADH-FMN-N3-N1b-N4-N5-N6a-N6b-N2-quinone is the main route for electron transfer within the enzyme (Figure 1.7B). Binuclear cluster N1a acts as an antioxidant. The large number of iron sulphur clusters reflects the modular assembly of Complex-I built from many small building blocks.

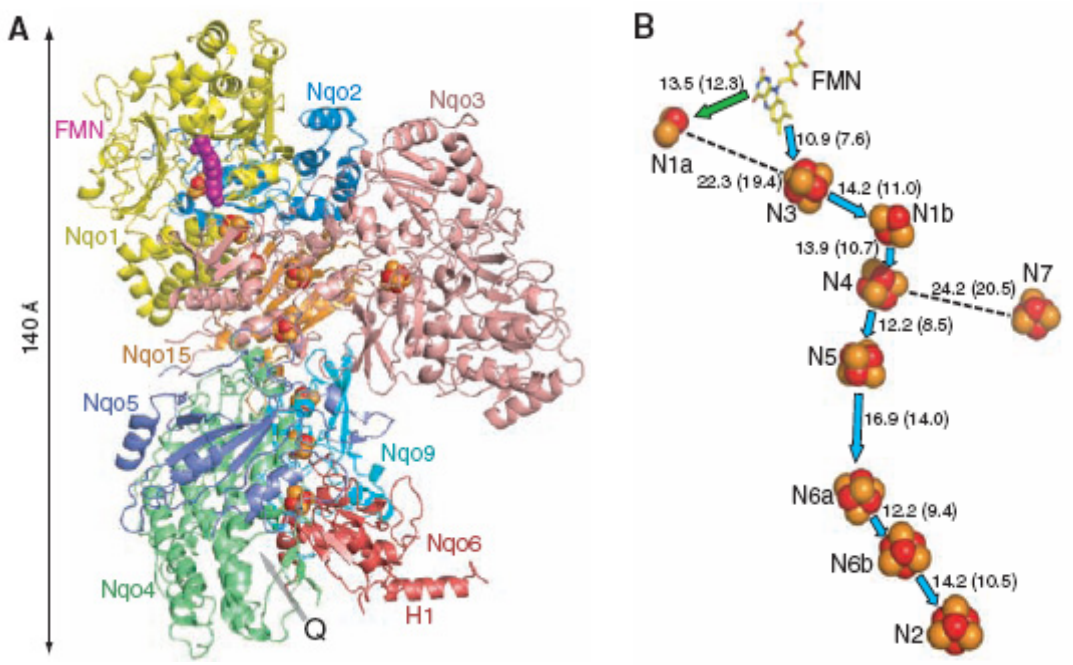


Figure 1.7: A: Structure of the hydrophilic domain of Complex-I from *Thermus thermophilus.* Eight subunits are shown. The grey arrow shows the entry point for ubiquinone (Q). B: The arrangement of the iron sulphur clusters in Complex-I is shown. The arrows show the path of electrons that travel to reach the ultimate electron acceptor N2. Reproduced from Saznov *et al*, 2006.

B. Membrane domain

The seven membrane subunits (Nqo7, 8, 10, 11, 12, 13 and 14) form the membrane domain of the complex in *Thermus thermophilus.* On one side of the domain a long connecting helix HL is present; which is an extension of the carboxyl terminal of Nqo12. The helix is involved in the conformational coupling of electron transfer and proton pumping. On the opposite side of the domain a β-hairpin-helix (β-H) and connecting elements (CH) are present which link these subunits together (Figure 1.8B).

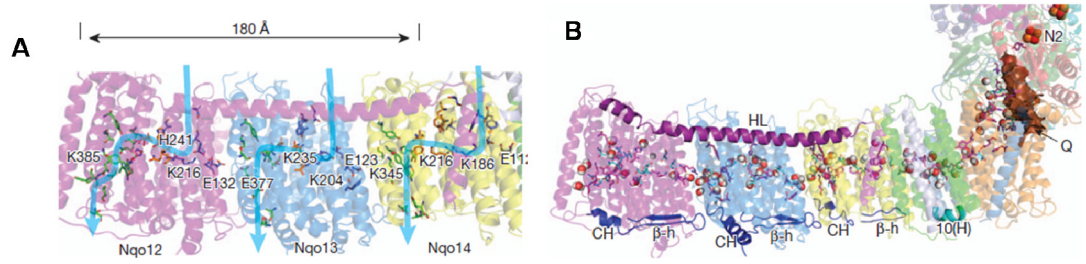


Figure 1.8: Membrane domain of *Thermus thermophilus* Complex-I. A: Major subunits Nqo12, 13 and 14 containing two half channels each related by inverse symmetry. B: The domains are connected by βH elements on one side of the domain and the C terminal extension of Nqo12; HL connects the membrane subunits together. Reproduced from Baradaran *et al*, 2013.

Nqo12, 13 and 14 are three largest of the subunits containing two half channels each related by inverted symmetry (Figure 1.8A). The fourth channel is formed by a half channel of Nqo8 connected to the second half channel between Nqo10 and Nqo11. Thus, the membrane possesses four channels, the fourth being at the interface of Nqo11 and Nqo14. Upon electron transfer in the hydrophilic domain, these half channels undergo conformational changes to form single proton translocation channels which are lined by charged residues. In Nqo8, the two half channels are linked by a Glu/Asp quadret and is referred to as the 'E channel'. Residues in the E channel are conserved and are essential for the activity. Thus, the membrane domain comprises of 64 transmembrane helices, which couple with the hydrophilic domain for activity.

C. Quinone reaction centre

The quinone reaction chamber is an enclosed 30 Å long chamber with a narrow entry point between subunits Nqo7 and Nqo8. The residues facing the lipid bilayer are hydrophobic, whereas those lining the inside of the chamber are hydrophilic. The front side of the cavity is negatively charged, top neutral and area near cluster N2 is positively charged (Figure 1.9). A small hydrophobic patch is present within the chamber to accommodate the quinone tail which remains in an extended conformation. During the binding of native ubiquinone, the last 1-3 isoprenoid units protrude out of the cavity into the lipid. Structural rearrangements are essential for the movement of the quinone head group in and out of the chamber. A probable second quinone binding site is unlikely to be present.

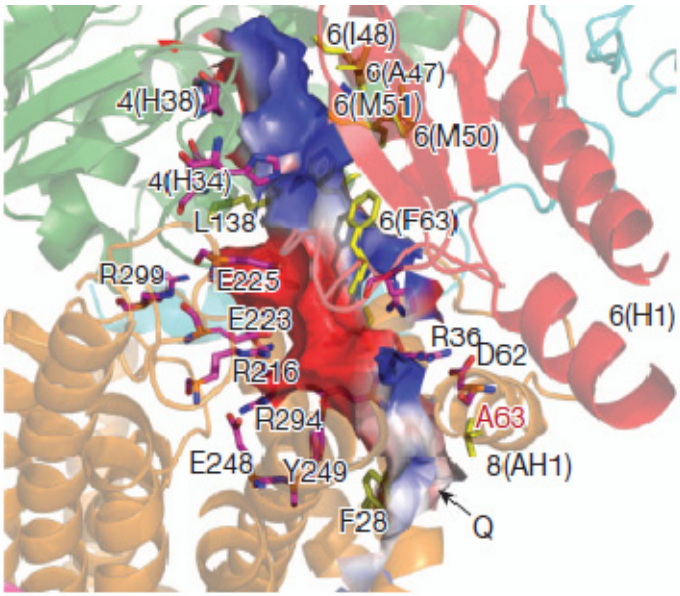


Figure 1.9: The structure of Q site/ quinone reduction centre. Red coloured part represents the negatively charged front region of the cavity, white region is the neutral top part of the cavity and blue colour represents the positively charged region near iron-sulphur cluster N2. The arrow indicates the entry point of quinone. Reproduced from Baradaran *et al*, 2013.

D. Coupling mechanism in prokaryotes

Tight coupling is observed between quinone reduction and proton pumping in Complex-I. Conformational changes involve the central hydrophilic axis of the hydrophilic domain and the helix HL and βH elements of the membrane domain. The electrons are transferred via the iron sulphur centres to cluster N2 in the hydrophilic domain. The cluster N2 on reduction, drives the movement of helices of Nqo4 and 6, which interact with Nqo8 providing a driving force for conformational changes. As the electrons are transferred from cluster N2 to quinone for reduction, the head group of the quinone is protonated by either Tyr87 or His38 further resulting in conformational changes. The Q site is linked to the Glu/Asp quadret in the centre of the E channel of Nqo8 by a hydrophilic funnel. This hydrophilic tunnel runs through the entire membrane domain till the tip of Nqo12 giving the membrane domain overall flexibility along the axis (Figure 1.10).

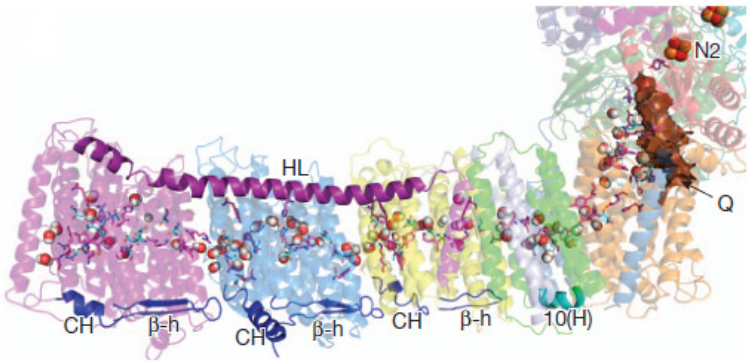


Figure 1.10: Hydrophilic and membrane domains of Complex-I highlighting Helix HL, βH structure of the membrane domain involved in conformational coupling. The Q site is shown in brown. The hydrophilic channel from the Q site links with the Nqo8 site and spans the entire membrane domain. Reproduced from Baradaran *et al*, 2013.

The N2 and electrostatically driven E channel cause conformational changes in Nqo14, 13 and 12 one after the other through a central axis. The movement of the helix HL results in proton pumping. In the oxidised state, the half channels of the membrane on the periplasmic side are open which take up the protons (Figure 1.11A). On reduction of quinone, there is movement of helix HL, which closes the half channels on the periplasmic side and the protons are pumped out (Figure 1.11B).

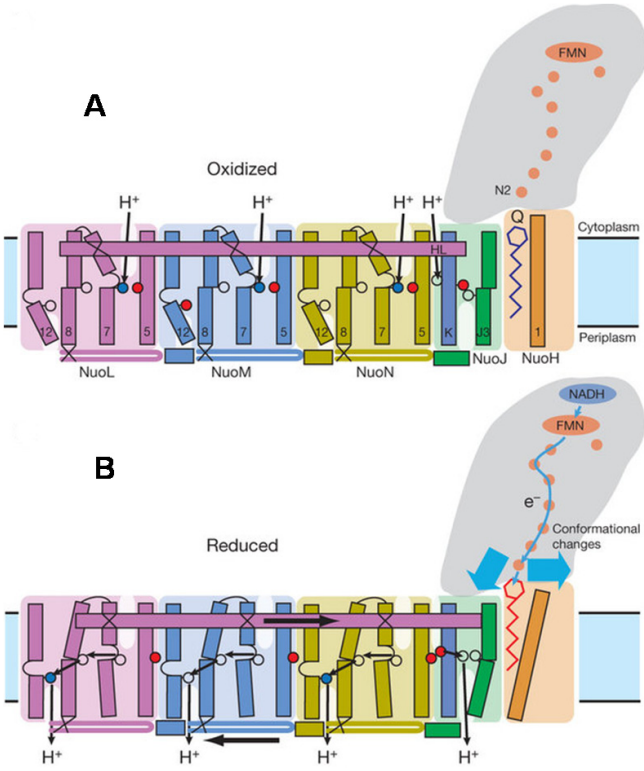


Figure 1.11: Conformational coupling of hydrophilic domain with membrane arm for quinone reduction and proton pumping. A: Oxidised form. B: Reduced form. Reproduced from Saznov *et al*, 2006; Efremov *et al*, 2011]

1.3.2.2 Structure of eukaryotic Complex-I

I. Complex-I of *Neurospora crassa*

The Complex-I from *N. crassa* is made up of 35 subunits, 7 of which are encoded by mtDNA [Weiss *et al*, 1991; Ise *et al*, 1985]. Under the electron microscope, the *N. crassa* Complex-I structure is L-shaped with two arms perpendicular to each other. The hydrophilic arm protrudes out into the cytosol whereas the membrane arm is embedded in the mitochondrial membrane. The two arms are equivalent to subcomplexes 1λ and 1γ of the bovine complex with some minor differences in the prosthetic groups present [Videira, 1998].

N. crassa Complex-I possesses the 14 'minimal' subunits required for oxidative phosphorylation, in addition it has a large number of accessory subunits. FMN and the four iron-sulphur clusters (binuclear N1 and tetranuclear N2, N3 and N4) are present in the hydrophilic domain [Wang *et al*, 1991]. Arrangement of prosthetic groups in different domains in the structure is enlisted in the Table 1.3.

***N. crassa* subunit (kDa)**	**Bovine homologue**	**Prosthetic group**
78	75 IP	N4, N1
51	51 FP	FMN, N3
24	24 FP	N1
21.3c	TKYK	N2
19.3	PSST	N2
9.3	B9	Q

Table 1.3: List of subunits of Complex-I from *N. crassa* possessing different prosthetic groups; FMN, iron sulphur clusters or ubiquinone (Q). Reproduced from Videira, 1998.

The cluster N2 is the direct reducer of ubiquinone and contributor to the proton translocation activity of Complex-I. It is most likely bound to subunits TYKY and PSST. A small 9.3 kDa subunit is identified to possess the ubiquinone binding pocket [Heinrich *et al*, 1992]. However, this subunit doesn't have a homologue in prokaryotes. Around 11 subunits of this Complex-I show homologues in bacteria. The

nuclear encoded subunits of Complex-I in *N. crassa* are related to either NAD^+-reducing hydrogenase of *Alcaligenes eutrophus* or to formate hydrogenlyase of *E. coli* [Schulte *et al*, 1994](Figure 1.12).

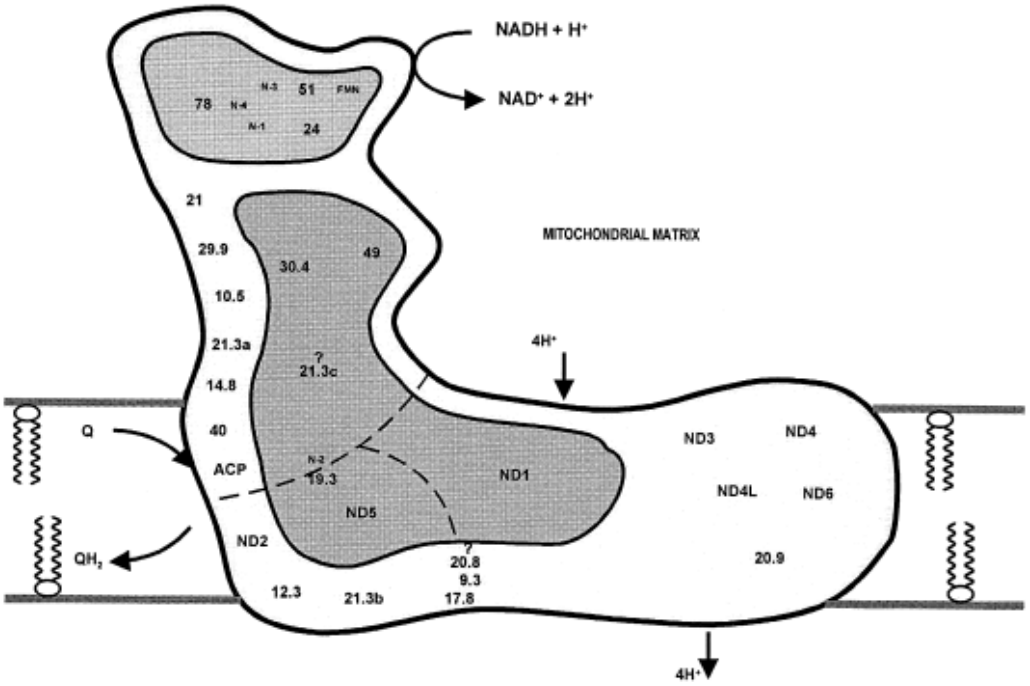


Figure 1.12: Structural model of *N. crassa* Complex-I. The peripheral matrix arm and membrane arms are separated by broken lines. The small and large shaded regions represent subunits related to NAD^+ reducing hydrogenase and formate hydrogenlyase respectively. Reproduced from Videira, 1998.

The 75 kDa, 51 kDa and 24 kDa subunits of Complex-I are homologues of the NADH oxidoreductase (αϒ dimer) of the NAD^+ -reducing hydrogenase [Friedrich *et al*, 1995]. The 49 kDa, 30 kDa, TKYK, PSST, ND1 and ND5 are homologous to bacterial formate hydrogenlyase [Weidner *et al*, 1993]. Apart from these subunits, the Complex-I also possesses an acyl-carrier protein. This reflects the additional roles played by the Complex-I in the fungus, apart from OXPHOS function. Thus, apart from OXPHOS, Complex-I is also involved in lipid biosynthesis [Brody *et al*, 1988].

Some portions of the *N. crassa* mitochondrial Complex-I are related to enzyme complexes from chloroplast and bacteria indicating that they are functional modules with a common origin. The structure is thus, speculated to be a result of association of pre-existing enzymes during evolution [Friedrich *et al*, 1997].

II Complex-I structure of *Yarrowia lipolytica*

The Complex-I of *Yarrowia lipolytica* has a molecular mass of 946.5 kDa constituting 40 subunits. Amongst these 40 subunits, 14 were the conserved core subunits essential for the activity of the assembly. The remaining 26 accessory subunits are arranged around these 14 subunits and function in assembly, stabilization, regulation and are involved in additional metabolic pathways.

The X-ray crystallographic structure at 6.3 Å revealed an L-shaped molecule, with the two arms oriented at 100° to each other. The membrane arm is 180 Å in length [Hunte *et al*, 2010].

A. Hydrophilic domain

Superimposition of the yeast structure with the bacterial structure revealed that the two parts of the peripheral arm form rigid bodies corresponding to the N and Q modules. As compared to the bacterial structure, these modules are opened 3° wider and twisted slightly relative to each other. One FMN and eight iron sulphur clusters are present in the peripheral hydrophilic arm. The distance between clusters were similar as seen in the complex of *Thermus thermophilus*. Seven of the clusters form a continuous electron chain between the catalytic sites (Figure 1.13).

B. Membrane domain

The membrane domain is divided into 2 portions of equal size: proximal and distal domains (Figure 1.13). These are connected by a narrow, centrally located interface in the membrane arm. The membrane domain comprises of 17 membrane integral subunits with more than 70 transmembrane helical segments. Some of the helices are interrupted in the middle of the membrane core indicating discontinuous helices similar to bacterial complex. The helices appear curved from the side view with the concave side facing the mitochondrial matrix.

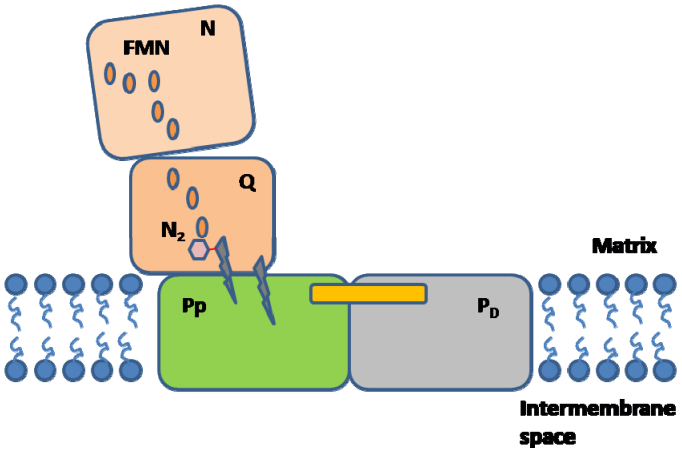


Figure 1.13: Schematic representation of the 4 functional modules of Complex-I from *Y. lipolytica.* A chain of 7 iron sulphur clusters (orange circles) lead form the N module to the cluster N2 in the Q module with FMN as a primary electron acceptor. Ubiquinone reduction site involves cluster N2 and Y144 (pink hexagon) in the Q module. The ubiquinone head group interacts with the Y144. The membrane domain is further divided into 2 sub-domains proximal Pp (green) and distal P_D (grey). A transmission element (yellow) forms a bridge across the two subdomains in the Q module. Reproduced from Hunte *et al*, 2010.

C. Quinone binding site

The ubiquinone binding site is at the interface of the highly conserved 49 kDa subunit and PSST subunit of the Q module. Cluster N2 present in the PSST subunit is the immediate electron donor to ubiquinone. A funnel like cavity leading from the N-terminal of the 49 kDa subunit towards the Tyr144 (Tyr87 in *T. thermophilus*) is the entry point of ubiquinone. The cavity is long enough to accommodate nine isoprenoid units of ubiquinone. The long ubiquinone tail acts as a tether that slides along the channel whereas the head group diffuses through the water phase. The binding cavity can be divided into a hydrophobic region around Tyr144 and a hydrophilic region formed by the C-terminal of the 49 kDa subunit [Hunte *et al*, 2010].

D. Coupling mechanism in eukaryotes

Energy released during redox reactions in the N and Q modules is transmitted to the proton-pumping machinery via long-range conformational changes. A long helix present, similar to helix HL of *T. thermophilus,* which is lateral to the membrane and embedded in the transmembrane segments near the matrix facing surface, is the key player in the conformational coupling.

III Bovine Complex-I of *Bos Taurus*

The bovine Complex-I is made up of ~45 subunits of which 7 are encoded by the mtDNA [Carroll *et al*, 2003]. It also comprises of a non-covalently bound FMN and eight iron sulphur clusters. Of the 45 subunits, 14 form the catalytic core while the remaining are supernumerary subunits.

Bovine Complex-I is also an L-shaped structure with one arm embedded in the mitochondrial membrane and the other arm protruding out into the mitochondrial matrix [Grigorieff, 1998]. The complex can be dissociated using mild chaotropic agents into four subcomplexes [Carroll *et al*, 2006].

i. Subcomplex 1λ: corresponding to the peripheral arm contains 15 subunits and has all the redox centres and the NADH binding site.

ii. Subcomplex 1α: A part of the subcomplex 1λ has 9 additional subunits.

iii. Subcomplex 1β has 13 subunits.

iv. Subcomplex 1γ: Consists of 6 subunits, which constitute the membrane arm.

The bovine Complex-I structure is closely related to the human Complex-I and gives insights into the structural complexity and arrangement of subunits in the complex. However, an atomic resolution structure is still not available.

Thus, the eukaryotic Complex-I requires further exploration.

1.3.3 Assembly and formation of Complex-I in eukaryotes

The assembly of Complex-I is a complicated and dynamic process. The large number of subunits and their dual genetic origin further complicates the assembly process. The hydrophilic subunits of nuclear origin together with the hydrophobic subunits of mitochondrial origin form a fully functional and mature Complex-I. The assembly of this complex has been extensively studied in *Neurospora crassa* [Videira, 1998], *Caenorhabditis elegans* [Van den Ecker *et al*, 2012], *Chlamydomonas reinharditii* [Cardol *et al*, 2008; Remacle *et al*, 2008] and in cultures of mammalian cell lines [Mimaki *et al*, 2012].

I. Complex-I assembly in *Neurospora crassa*

The Complex-I of *N. crassa* contains 35 subunits, of which 7 are encoded by the mitochondrial genome. *N. crassa* Complex-I assembles via the evolutionarily conserved core subunits. The hydrophilic matrix arm and the membrane arm are assembled independently. The membrane arm is formed from two assembly intermediates. One of them is a ~200 kDa complex consisting of ND2 and ND5 and the other a ~350 kDa complex containing ND1, ND4, ND4L and ND6. The 350 kDa complex utilizes CIA chaperones to assemble. Although the study of *N. crassa* Complex-I gave the first insight into the assembly process of Complex-I it couldn't be extended to the mammalian system because of the presence of extra subunits in mammalian system, which are absent in fungus [Videira, 1998].

II. Complex-I assembly in *Caenorhabditis elegans*

C. elegans is an established model to study mitochondrial function in a range of OXPHOS diseases [Van den Ecker *et al*, 2012]. There is a large conservation of mitochondrial composition between mammalian and *C. elegans*. Almost 84% of human Complex-I subunits are found in *C. elegans.* Thus, they probably follow the same assembly pattern as in human Complex-I.

III. Complex-I assembly in *Chlamydomonas reinhardtii*

The assembly of the mitochondrially encoded subunits is well characterized in green algae; *C. reinhardtii.* A 200 kDa nuclear encoded subcomplex containing *NDUFS1* and *NDUFS2* subunits binds to the membrane by combining with ND1, ND3, ND4L and ND6. ND3 and *NDUFA9* are located at the junction between the matrix and membrane. A 700 kDa mega-complex is formed when the subunits *NDUFV1, NDUFV2, NDUFS3, NDUFS4, NDUFS7, NDUFS8* and *NDUFA12* get associated with the *NDUFS1* and *NDUFS2*. This 700 kDa mega-complex further expands by the addition of ND4 and ND5 to form the holo Complex-I enzyme. Thus, the 700 kDa complex is firmly anchored to the membrane by the ND4 and ND5 subunits [Cardol *et al*, 2008; Remacle *et al*, 2008].

IV. Complex-I assembly in humans

Several assembly models have been proposed in humans by studying conditional assembly systems in cell lines where the assembly process is disturbed [Ugalde *et al*, 2004a]. Several experiments have been done by tracing assembly intermediates in wild type cell lines by pulse chase experiments, by tracking the assembly of the individual nuclear encoded subunits by using an in vitro mitochondrial import and assembly assay [Lazarou *et al*, 2007], by monitoring the assembly pattern of the green fluorescent protein tagged NDUFS3 subunit and its progression into a holoenzyme [Vogel *et al*, 2007] for understanding the assembly process.

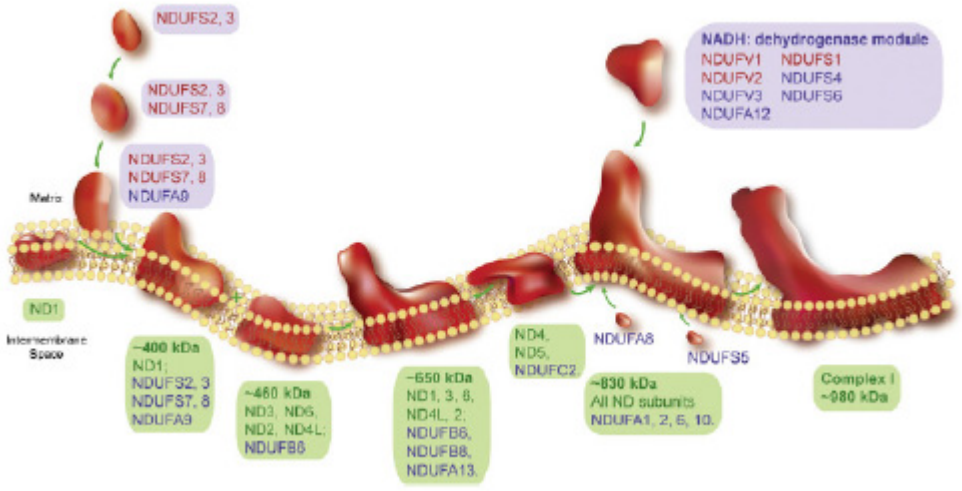


Figure 1.14: Diagrammatic representation of the assembly process of Complex-I in humans. The core subunits are shown in red, nDNA encoded subunits in blue and mtDNA encoded subunits in green. Reproduced from Mimaki *et al*, 2012.

A general consensus model from all these studies revealed that like in *N. Crassa* the hydrophilic matrix arm and membrane arm assemble independently in humans as well (Figure 1.14). Evidence shows that it occurs in a stepwise fashion, where the intermediates N, Q and P modules assemble individually and then come together to form a viable complex [Vogel *et al*, 2007]. During the early stages of assembly, *NDUFS2* and *NDUFS3* form a small hydrophilic sub-complex which further expands on the addition of *NDUFS7, NDUFS8* and *NDUFA9*. This peripheral sub-complex binds to mitochondrially encoded ND1 subunit to form a 400 kDa assembly intermediate. On the other hand, subunits ND3, 6, 2, 4L and *NDUFB6* assemble to form a 450 kDa complex which combines with the peripheral assembly intermediate

to form a 650 kDa sub-complex. Subunits ND4, 5 and *NDUFC2* assemble independently and then join the 650 kDa complex to form a 830 kDa assembly intermediate. In the meanwhile, *NDUFV1, 2, 3, NDUFS1, 4, 6* and *NDUAF12* combine independently to form the N module which joins the 830 kDa complex along with NDUFA8 and NDUFS5 to form a fully functional mature Complex-I.

1.3.4 Complex-I related diseases

Complex-I is the largest multi-subunit enzyme complex in the mitochondria. Due to large number of subunits, their dual genetic origin and also due to the dynamic assembly process Complex-I is prone to malfunction resulting in OXPHOS disorders. Mutations may affect either the subunits or assembly factors. It has been pointed out that in 40% of the mitochondrial diseases it is the Complex-I malfunctioning that happens to be the reason for the disease. Isolated Complex-I deficiency is the most commonly identified biochemical defect in childhood onset mitochondrial disease. It is clinically heterogeneous. Majority of the affected individuals, develop symptoms during the first one year of their life and have a rapidly progressive disease course resulting in a fatal outcome. Certain selected mutations in the 14 conserved subunits and assembly factors along with their clinical phenotype are enlisted in Table 1.3.

Table 1.4: Mutations in the subunits of Complex-I.

Subunit	Mutation	Clinical Phenotype	References
N module			
NDUFV1	Y204C-C206G	Leigh like syndrome	Benit *et al*, 2001
	A341	Leukoencephalopathy with macrocephaly	Schuelke *et al*, 1999
	R386C/H	Leukoencephalopathy	Vilain *et al*, 2012; Marin *et al*, 2013
NDUFV2	IVS2+5_8delGTAA (skipping of exon 2)	Cardiomyopathy and encephalopathy	Benit *et al*, 2003; Pagniez-Mammeri *et al*, 2009
NDUFS1	R241W-R557X .	Unspecified encephalomyopathy	Benit *et al*, 2001
	M707V-LS deletion	Leigh like syndrome	Benit *et al*, 2001

	Q522K	Leukoencephalopathy	Bugiani *et al*, 2004
	L231V	Leigh syndrome	Martín *et al*, 2005
	D619N-R557X	Leukoencephalopathy	Hoefs *et al*, 2010
	R408C	Leigh syndrome and leukoencephalopathy	Hoefs *et al*, 2010
Q module			
	R228Q-S413P	Leigh syndrome	Loeffen *et al*, 2001
	R138Q-R333Q	Leigh like syndrome	Tuppen *et al*, 2010
NDUFS2	M292T-RR118Q	Leigh syndrome	
	M292T-R443K	Leigh syndrome	
	M292T-E148K	Leigh syndrome	
NSUFS3	T145I-R199W	Leigh syndrome	Benit *et al*, 2004
NDUFS7	V122M	Leigh syndrome	Triepels *et al*, 1999
	R145H	Leigh like syndrome	Lebon *et al*, 2007b
NDUFS8	P79L-R102H	Leigh syndrome	Loeffen *et al*, 1998
	P85L-R138H	Leigh syndrome	Procaccio *et al*, 2004
P module			
MT-ND1	A52T	LHON	Howell *et al*, 1991
	E143K	LHON	Achilli *et al*, 2012
MT-ND2	L71P	Leigh syndrome	Ugalde *et al*, 2007
	L128Q	LHON	Kumar *et al*, 2010
MT-ND3	A47T	LHON and dystonia	Wang *et al*, 2009
	S45P	Leigh syndrome	Nesbitt *et al*, 2012
MT-ND4	R340H	LHON	Kumar *et al*, 2010
MT-ND4L	V65A	LHON	Achilli *et al*, 2012
MT-ND5	V253A	Leigh syndrome	Ching *et al*, 2013
	I596V	OXPHOS diseases	Hinttala *et al*, 2006
	D393N	OXPHOS diseases	Brautbar *et al*, 2008

MT-ND6	M64V/I	LHON	Kumar *et al*, 2012
	L60S	LHON	Achilli *et al*, 2012
	A72V	LHON and dystonia	Achilli *et al*, 2012
Supernummery subunits			
NDUFS4	W97S	Leigh like syndrome	Budde *et al*, 2000
	R106X	Leigh like syndrome	
	W14X	Leigh like syndrome	Petruzzella *et al*, 2001
NDUFS6	C115Y	Severe neonatal lactic acidosis	Spiegel *et al*, 2009
NDUFA12	R60X	Leigh syndrome	Ostergaard *et al*, 2001
NDUFA11	IVS1+5G>A (splicing abnormality)	Lethal infantile mitochondrial disease	Berger *et al*, 2008
NDUFA2	C208+5G>A	Leigh like syndrome	Hoefs *et al*, 2008
NDUFA10	C1A>G and C425A>G	Leigh syndrome	Hoefs *et al*, 2011
NDUFA1	G8R	Leigh like syndrome	Fernandez-Moreira *et al*, 2007
	R37S	Encephalomyopathy	
	G32R	Progressive neurodegeneration	Potluri *et al*, 2009

Mutations may lead to diseases with the involvement of single organ or tissue such as hypertrophic cardiomyopathy (HCM) or Leber's hereditary optic neuropathy (LHON). Diverse clinical presentations are also observed in cases of fatal neonatal lactic acidosis, infantile onset Leigh syndrome, MELAS syndrome (mitochondrial encephalomyopathy, lactic acidosis and stroke like episodes) and in certain cases of adult onset of encepahlomyopathy syndrome.

I. Mutations in the nDNA encoded subunits

Of the 38 nuclear encoded Complex-I subunits, 17 have mutations reported in them responsible for Complex-I deficiency. Till date more than 100 patients with mitochondrial diseases have been identified of which ~60% have mutations in the

core subunits and the remaining 40% have mutations in the accessory subunits. Most mutations have been reported in the *NDUFS1* and *4* subunits (24 and 21 cases, respectively). Most frequently observed phenotypes are Leigh syndrome, leukoencephalopathy and hypertrophic cardiomyopathy. These mutations follow an autosomal recessive inheritance pattern.

II. Mutations in the mtDNA encoded subunits

Mutations have been reported in all the seven genes of Complex-I encoded by the mtDNA. These mutations were linked to LHON and were homoplasmic in nature. Heteroplasmic mutations were reported to cause dystonia, MELAS and Leigh syndrome. Many patients may have overlapping features of two syndromes like LHON plus dystonia. These mutations are maternally inherited.

III. Mutations in tRNA and rRNA

22 tRNA's and 2 rRNA's encoded in mtDNA are involved in translation. Approximately 300 mutations have been reported in these RNA's [MITOMAP 2013]. The tRNA genes are one of the hot spots of mutations resulting in mitochondrial disorders [Ukei *et al*, 2006]. There is a large clinical spectrum of the diseases due to mutations in tRNA genes, ranging from mitochondrial myopathy, myglobinuria, MELAS, tubulointerstitial nephtritis, deafness, movement disorder (dystonia), Leigh's syndrome, Diabetes mellitus and deafness (DMDF), Chronic Progressive External Opthalmoplegia (CPEO), Ocular myopathy, encephalomyopathy, LHON, gastrointestinal syndrome to multiple sclerosis. Of all the mutations in tRNA, 39 are located in tRNA Leu gene (UUR and CUN). Mutations in tRNA and rRNA genes lead to dominantly acting, qualitative translation defects like amino acid misincorporation or premature translation termination products. Each type of tRNA mutation may result in the synthesis of a different set of aberrant translation products, thereby resulting in diverse clinical effects. A point mutation in a tRNA gene may thus, influence protein synthesis either quantitatively or qualitatively.

IV. Mutations in assembly factors

A large number of assembly factors are involved in the assembly of the holo-Complex-I enzyme. They are not a part of the final Complex-I structure but are involved in Complex-I biogenesis and stability of intermediates.

Mutations in the assembly factors prevent the assembly of Complex-I at different stages (Table 1.5) resulting in varied clinical phenotypes.

Table 1.5: A list of selected mutations reported in the assembly factors of Complex-I and their clinical manifestations.

Assembly factor	Mutation	Clinical Phenotype	Assembly defect	References
C20orf7	L229P	Lethal infantile mitochondrial disease	Translation or stabilization of ND1 subunit is impaired.	Sugiania *et al*, 2008; Gerards *et al*, 2011; Saada *et al*, 2012
	L159F	Leigh syndrome		
	G250V	Leigh syndrome		
NDUFAF3	G77R	Lethal infantile mitochondrial disease	Peripheral subcomplex containing *NDUFS2, 3, 7 & 8* fails to insert in the membrane.	Saada *et al*, 2008; 2009
	R122P			
	M1T-R122P			
NDUFAF4	L65P	Lethal infantile mitochondrial disease, antenatal cardiomyopathy	Peripheral subcomplex containing *NDUFS2, 3, 7 & 8* fails to insert in the membrane.	Saada *et al*, 2008; 2009
NDUFAF1	K253R-T207P	Cardioencephalo myopathy	Assembly intermediate containing ND2, ND3, ND4L is not formed.	Vogel *et al*, 2005; Dunning *et al*, 2007
ACAD9	R518H	Cardiomyopathy, hearing loss and exercise intolerance	Decreased levels of NDUFAF1, Ecsit and mature Complex-I.	Nouws *et al*, 2010; Haack *et al*, 2010; Gerards *et al*, 2011
		Encephalopathy,		

	E63X-E413H	cardiomyopathy		
	R532W	Excercise intolerance		
	R127Q-R469W	Excercise intolerance		
NDUFAF2	R45X	Encephalopathy	Late stages of Complex-I assembly are impaired (N module cannot assemble).	Ogilvie *et al* 2005; Hoefs *et al* ,2009; Calvo *et al,* 2010; Barghuti *et al*, 2008.
	M1L	Encephalopathy		
	Y38X	Leigh syndrome		
	W74X	Leigh syndrome		
NUBPL	G56R	Encephalopathy	Peripheral subcomplex cannot assemble or insert into the membrane and the N module cannot assemble.	Calvo *et al*, 2010, Bych *et al*, 2008

Structural and functional understanding of the subunits of Complex-I and their assembly factors would help to understand at molecular level the effects of pathogenic mutations leading to disease condition.

1.3.5 Leigh syndrome

Leigh syndrome is a fatal subacute necrotising encephalopathy. It is a rare genetic disorder resulting in progressive neurodegeneration. It is caused by the mutations in the subunits or assembly factors of the complexes of the respiratory chain (Table 1.6). It is inherited in an autosomal recessive fashion or maternally inherited. This disorder was first described by Denis Leigh in 1951 [Leigh, 1951]. The prevalence of Leigh syndrome is 2.05 cases per 1,00,000 [Castro-Gago *et al*, 2006] with a pre-school incidence of 1 in 32,000 [Darin *et al*, 2001].

Table 1.6: List of mutations in the subunits of the respiratory chain causing Leigh syndrome.

Respiratory chain Complex deficiency	Disorder	Genes	References
I	Complex-I deficient Leigh syndrome	NDUFV1, NDUFS1, 2, 3, 4, 7, 8, NDUFA1, NDUFA2, NDUFA10, NDUFAF2, C8orf38, C20orf7, FOXRED1	Loeffen *et al*, 2000; Benit *et al*, 2001; 2004; Fernandez-Moreira *et al*, 2007; Hoefs *et al*, 2008; 2009; 2011; Pagliarini *et al*, 2008; Calvo *et al*, 2010; Gerards *et al*, 2010; Tuppen *et al*, 2010
II	Complex-II deficient Leigh syndrome	SDHA	Bourgeron *et al*, 1995; Pagnamenta *et al*, 2006
IV	Cytochrome c oxidase-deficient Leigh syndrome	SURF1, COX10, COX15	Pequignot *et al*, 2001; Antonicka *et al*, 2003; Oquendo *et al*, 2004
	French-Canadian or Saguenay-Lac Saint Jean type	LRPPRC	Mootha *et al*, 2003
II + III	Coenzyme Q_{10} deficiency	PDSS2	Van Maldergem *et al*, 2002; Lopez *et al*, 2006
I, III+IV	Mitochondrial translation defect	C12orf65	Antonicka *et al*, 2010

I. Diagnosis of Leigh syndrome

The typical symptoms include psychomotor regression, abnormal muscle tone, weakness, dystonia, brainstem and cerebral dysfunction (ataxia), visual loss, missed milestones, tachypnea, seizures, lactic academia and acute deterioration following common infections. Symptoms appear within the first year of life with difficulties in feeding, vomiting and failure to thrive. The progression of the disease is fast and typically death occurs within a few years due to progressive respiratory failure [Shrikhande *et al*, 2010; Fassone *et al*, 2012].

However, certain individuals with clinical features of Leigh syndrome do not fulfil stringent diagnostic criteria because of atypical neuropathology or normal

neuroimaging, blood and CSF lactate levels are said to have symptoms of Leigh like syndrome [Thorburn *et al*, 2011].

Rahman *et al* 1996, defined stringent diagnostic criteria for Leigh syndrome. Diagnosis includes identifying initial symptoms of motor and intellectual developmental delay, signs and symptoms of brain-stem and/or basal ganglia disease, characteristic symmetric nectrotic lesions in the basal ganglia and/or brain stem. Laboratory analysis helps in diagnosis by indicators like metabolic acidosis with elevated lactate and pyruvate concentration in blood or CSF.

A. Biochemical diagnosis

Diagnosis is generally done by estimating the lactate levels in blood and CSF samples. The lactate levels are usually elevated. Plasma amino acids may show increased alanine concentration. Urine organic acid analysis may also be done to detect lactic aciduria. Activities of the respiratory chain enzyme complexes may also be checked. Generally, skeletal muscles or skin fibroblasts are the tissue of choice for activity assays.

B. Muscle biopsy

Histological examination shows accumulation of intra-cytoplasmic neutral liquid droplets. Ragged red fibres which are a hallmark in adult-onset of mitochondrial diseases are not observed in the case of Leigh syndrome.

C. Molecular Genetic testing

Targeted mutational analysis is done on DNA extracted from leukocytes. Sequencing of mitochondrial genes MT-ND1, 2, 3, 4, 5, 6 or nuclear genes like NDUFS2, 3, 7, 8 etc is done to detect mutations.

D. Brain imaging

Magnetic resonance imaging (MRI) shows characteristic bilateral symmetric hypodensities in the basal ganglia or bilateral symmetric hyperintense signal abnormality in the brainstem or basal ganglia [Arii *et al*, 2000; Rossi *et al*, 2003]. Brain lesions may also be observed in certain cases.

E. Management of disease

Post diagnosis, the progression of the disorder is monitored by developmental assessment, routine neurological evaluation (MRI, MRS, EEG etc), metabolic

evaluation (plasma and CSF lactate and pyruvate concentrations and urine organic acids), ophthalmologic and cardiac activity monitoring.

II. Therapies

Leigh syndrome, unfortunately, is a fatal disorder with no cure. However, several therapies have been put forth for the management of the disorder. Management for specific symptoms include treatment of acidosis, seizures, dystonia and cardiomyopathy with specific drugs. Some alternative therapies under consideration are:

A. Antioxidants

Mitochondrial-targeted antioxidants such as coenzyme Q, mitoQ or analogues like idebenone show protection against oxidative stress in cultured cells and animal models. A range of vitamins and other antioxidants like riboflavin, thiamine and coenzyme Q10 are often used to improve mitochondrial function. A high fat diet along with biotin, creatine and succinate is also recommended. Role of α-ketoglutarate and aspartate in increasing the flux of citric acid cycle, thereby increasing ATP production independent of OXPHOS has also been explored [Sgarbi *et al*, 2009].

B. Gene therapy

Gene therapy provides a potential approach to decreasing the proportion of mutant mtDNA in the cells of an individual. Studies have shown the utility of mitochondrially targeted restriction endonuclease delivered by an adenoviral vector in recognizing and degrading mutant mtDNA while leaving the wild type DNA intact [Tanaka *et al*, 2002; Alexeyev *et al*, 2008]. However, the clinical applicability of such approaches is still questionable.

C. Genetic counselling

Genetic counselling is the process of providing individuals and families with a history of the disorder; information on the nature, inheritance and implications of genetic disorders to help them make informed medical and personal decisions. Pre-natal diagnosis plays an important role in cases where an affected family member has been

identified. Analysis is done on the DNA extracted from non-cultured foetal cells obtained by amniocentesis or chronic villus sampling.

1.3.6 Mouse models of Complex-I deficiency

Recently a number of mouse models for Complex-I deficiency have been developed which help in understanding disease mechanisms in Complex-I deficiency [Fassone *et al*, 2012]. One of the first models was that of the Harlequin mouse, possessing the hypomorphic mutation in the *Aif* gene encoding the apoptosis inducing factor [Benit *et al*, 2008]. However, the human homologue of the gene AIFM1 resulted in progressive encephalopathy with multiple respiratory chain defects rather than isolated Complex-I deficiency [Ghezzi *et al*, 2010]. Hence, it is not the best model for studying human complex-I deficiency. Since then many knockout mouse models of nuclear encoded mitochondrial genes were developed, however, they were lethal at the embryo stage itself.

A mouse with a conditional knockout of the *NDUFS4* gene was born healthy, however; it developed ataxia at 5 weeks and progressive encephalopathy leading to its death by 7 weeks. Its neuropathological features resembled that of the Leigh syndrome [Kruse *et al*, 2008; Quintana *et al*, 2010]. Another model, *NDUFS6* gene trap mouse model showed isolated Complex-I deficiency; cardiomyopathy starting from day 30 with heart failure, weight loss or sudden death at 4-8 months. Complex-I activity was 10% of control values due to very low levels of assembled complex and reduced ATP production [Ke *et al*, 2012]. Such mouse models are useful tools for understanding the effects of pathogenic mutations or the prognosis of diseases in cases of isolated Complex-I deficiency.

1.3.7 Structural effects of Complex-I mutations

Apart from the clinical perspective, several researchers have tried to explore the structural consequences of the mutations on Complex-I. However, due to the lack of a high resolution structure of Complex-I for several years, the structural aspects of the pathological mutations could not be explored to a great extent. Owing to the recent developments and using the recently reported structure from *T thermophilus* and *Y. lipolytica* several studies have been possible mapping the mutation sites onto the

structures. This has helped in exploring the structural aspects of mutations and hindrances caused by them in assembly or activity of Complex-I.

Early studies included, noting the changes in the amino acids on mutation, observing the conservation of mutated residue, predicting possible change in secondary structure and several such preliminary observations correlating the dysfunction to possible structural changes. DeHaan *et al* [DeHaan *et al*, 2004] extensively studied the mutations associated with LHON in the subunit ND6. They modelled the subunit ND6 and showed the changes in the secondary structure on mutation. Mutation in ND6 significantly changed the accessible surface area of the protein and orientation of individual amino acids, which resulted in secondary structure changes from helix to coil. These secondary structure changes were thought to alter the proton conductance properties resulting in defective respiration control. In another study by Ngu *et al* [Ngu *et al*, 2012], several mutations in the human *NDUFS2* responsible for Leigh syndrome were mapped onto the modelled structure in order to understand the location of the mutations with respect to the quinone binding site and the other subunits. Such attempts were made to correlate the structural effects of the mutation to dysfunction in Complex-I and disease phenotype.

The crystal structure of Complex-I from *T. thermophilus* at atomic resolution of 3.3 Å aided Braradaran *et al* [Braradaran *et al,* 2013] not only to understand the structure and functioning of Complex-I but has also helped in understanding the location of mutations in the ND1 subunit. Different mutations associated with diseases like LHON, MELAS, NIDDM, encephalopathy, cardiomyopathy etc were studied and correlated with dysfunction of Complex-I.

In the present thesis, *in silico* and biophysical studies have been conducted on the core subunits of the ubiquinone reduction module (Q module) of the human mitochondrial Complex-I. The 4 subunits *NDUFS2, 3, 7 & 8* in humans have been studied by molecular modelling and dynamic simulations to understand the structural consequences of the mutations on the subunit assembly and complex function. Attempts were made to clone and express these proteins for detailed experimental studies. Finally, only two subunits *NDUFS3* (structural) *& NDUFS7* (functional) could be cloned and expressed in *E coli*. The subunits engineered with mutations

associated with Leigh syndrome have been prepared by site-directed mutagenesis. The w-t and mutant proteins have been compared in terms of their stability, aggregation propensities, binding with n-decylubiquinone and other biophysical properties. The differences in their properties highlight the change in nature of the mutant proteins *in vitro* in solution. The changes have been correlated with complex functional defects or complex disassembly.

CHAPTER 2

Materials and Methods

In the research presented here, several approaches have been used to study the structure, stability and function of the ubiquinone reduction (Q) module. The materials and methods utilized in the entire study are enlisted below.

2.1 Phylogenetic analysis

Protein sequences of subunits *NDUFS2*, *3*, *7* and *8* available from bacteria to mammals were downloaded from UniProtKB (http://uniprot.org) protein database. The sequences for each have been aligned using ClustalW, the program that is part of the MEGA: Molecular Evolutionary Genetics Analysis software [Thompson *et al*, 1994; Tamura *et al*, 2011]. Phylogenetic analysis was performed with the MEGA 5.5 software [Tamura *et al*, 2011; Hall BG, 2013], using the neighbor joining statistical method with 1000 bootstrap replications. The phylogenetic trees generated were visualized in FigTree v1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/) [Rambaut, 2009].

2.2 Molecular modelling

2.2.1 Core subunits of the Q module

The Q module is made up of four core subunits *NDUFS2, 3 7 & 8*, the constituents of complex-I common from bacteria to mammals. The protein sequences of these four core subunits of Q module were acquired from the UniProtKB protein databank (Entry number: O75306, O75489, O75251, O00217). A similarity search carried out using the BLAST server [Altschul *et al*, 1990] to choose model templates, identified the bacterial complex (PDB ID: 2FUG, 3I9V, 3IAM) of the organism *Thermus thermophilus* in PROTEIN DATABANK (PDB) (http://www.rcsb.org). These had 40-50% sequence identity with the corresponding four subunits of human Complex-I Q module [Berrisford *et al*, 2009; Saznov *et al*, 2006]. Subunits *NDUFS2* and *NDUFS7* were modelled using D and F chains of 3I9V (resolution: 3.1 A°) whereas subunits *NDUFS3* and *NDUFS8* were modelled using E and G chains of 3IAM (3.1A°), templates were chosen based on higher sequence similarity and better resolution of the crystal structure. The signal sequences for subunits were predicted using MITOPROT: prediction of mitochondrial targeting sequences server [Claros *et al*, 1996]. Signal sequences were not part of modelling (Table 2.1).

Table 2.1: Signal sequences are shown predicted by MITOPROT for the four core subunits. These regions were excluded from modelling.

Subunit	Signal sequence	Cleavage site
NDUFS2	MAALRALCGFRGVAAQVLRP	21
NDUFS3	MAAAAVARLWWRGILGASALTRGTGRPSV LLLPVRRE	38
NDUFS7	MAVLSAPGLRGFRILGLRSSVGPAVQARGV HQSVATDGPSSTQPALPKARA	52
NDUFS8	MRCLTTPMLLRALAQAARAGPPGGRSLHS SAVAA	35

The molecular models were prepared using the software MODELLERv9.10 [Eswar *et al*, 2007; 2008]. 20 models were generated in each case. Energy minimization of both wild type and mutants was carried out by using GROMACS v.4.5 [Hess *et al*, 2008] by the steepest descent minimization for 100ps with maximum force field cut off being 1 KJ/mol.

2.2.2 Modelling Q module assembly

The subunit assembly of Q module was modelled by comparing with the arrangement of chains D, E, F and G (PDB ID: 2I9V) using the alignment program in molecular modelling suite PRIME v3.1 (Schrödinger). Ten initial models of Q module along with three iron sulphur clusters were first generated. The models were then energy minimized and prepared in the "protein preparation wizard" of Maestro 9.3.

2.3 Evaluation of the generated models

The molecular models, subunits as well as the entire Q module, were evaluated using Discrete Optimized Protein Energy (DOPE) score, ERRAT (version 2.0) [Colovos *et al*, 1993], PDBsum [Laskowski, 2009], ProSA-web Protein structure analysis [Wiederstein *et al*, 2007] and RMSD based on Cα overlap between target and template.

The DOPE score is an independent assessment of the accuracy of the output models in the MODELLERv9.10. The program assigns a score for a model by considering the positions of all the non-hydrogen atoms. Lower scores correspond to predicted models that are more accurate. The ERRAT program verifies the quality of the model. This program plots error values as a function of position in the sequence

by sliding a nine residue window along the sequence. The error function is based on the statistics of non-bonded atom-atom interactions in the template structure. PDBSum is a pictorial database providing an overview of the model structure in terms of Ramachandran plot statistics, main chain parameters etc. ProSA is a tool widely used to check errors in 3D protein models. An overall quality score or Z score is estimated and shown in a plot where scores estimated from experimentally determined structures in PDB are plotted. The Z score is an indication of the overall quality of model and measures the deviation of the total energy of the modelled structure from an energy distribution derived from random conformations.

The 20 models generated for each of the 4 subunits of the Q module and the 10 models generated for the entire Q module were evaluated with the above parameters. The ones with the best statistics were utilized for further study.

2.4 *In silico* mutagenesis of the subunits

The *in silico* mutants were generated using the program FoldX (Table 2.2) [Schymkowitz *et al*, 2005].

Table 2.2: List of *in silico* mutations generated using the FoldX program. All these mutations are known to result in Leigh or Leigh like syndrome.

Subunit	Mutation	Reference
NDUFS2	R228Q+S413P	Loeffen J *et al* 2001
	R138Q+R333Q	
	M292T+R118Q	Tuppen HA *et al* 2010
	M292T+M443K	
	M292T+E18K	
NDUFS3	T145I+R199W	Benit P *et al* 2004
NDUFS7	V122M	Triepels RH *et al* 1999
	R145H	Lebon S *et al* 2007
NDUFS8	P85L+R138H	Loeffen J *et al* 2001
		Procaccio V *et al* 2004

2.5 Molecular Dynamics (MD) simulation of individual subunits

The molecular models of wild-type and mutant proteins were used to perform molecular dynamics simulation using GROMACS v.4.5 with the OPLS-AA/L all-atom force field. The subunits were solvated with SPC water model using the genbox program of GROMACS suite. The default cubic boxes of GROMACS with dimensions: 12.03 nm (*NDUFS2*), 10.55 nm (*NDUFS3*), 9.38 nm (*NDUFS7*) and 7.49 nm (*NDUFS8*) were used. Sodium and chloride ions were added to each system depending on the requirement for charge neutralization (Table 2.3).

Table 2.3: Ions introduced to neutralize the system for each subunit and its *in silico* mutant.

Subunit	**Ions added to neutralize the system**	
	Na^+	**Cl^-**
NDUFS2		
w-t	4	-
R228Q+S413P	5	-
R138Q+R333Q	6	-
M292T+R118Q	5	-
M292T+M443K	3	-
M292T+E18K	2	-
NDUFS3		
w-t	5	-
T145I+R199W	6	-
NDUFS7		
w-t	-	1
V122M	-	1
R145H	-	1
NDUFS8		
w-t	6	-
P85L+R138H	7	-

Energy minimization was performed by the steepest descent method for 50000 steps (the minimization tolerance was set to 1000 kJ/mol-nm). Equilibration was carried out in 2 steps using conditions NVT and NPT, respectively. Long range electrostatics was computed using the Particle Mesh Ewald (PME) method [Darden *et al*, 1993] and Lennard-Jones energy cut off was set to 1.0 nm. Bond lengths were constrained with the LINCS algorithm [Hess *et al*, 1997]. Simulations of 15 ns duration were performed on the wild type and mutant structures at constant temperature of 300 K maintained by modified Berendsen thermostat coupling [Berendsen *et al*, 1984] and at a constant pressure of 1 bar by Parrinello-Rahman pressure coupling. The time step employed was 2 fs and coordinates were saved every 2 ps for analysis of MD trajectory. Analyses were performed with the tools available in the GROMACS utilities. RMSD, RMSF, Radius of gyration (Rg), average number of inter and intra-molecular hydrogen bonds formed, solvent accessible area and secondary structure prediction along the trajectory (DSSP) were estimated.

2.6 Docking of n-decyl-ubiquinone (DBQ) in Q module

The molecular model of Q module prepared as already described was used to dock n-decyl-ubiquinone (DBQ) using Glide 5.8 (Schrödinger) [Friesner *et al*, 2004]. The model structure was imported and a centroid receptor grid was generated around the residue Y141 known to interact with DBQ [Angerer *et al*, 2012]. The ligand molecule DBQ (CID 2971) was downloaded from PubChem compound database [Balton *et al*, 2008] and prepared in the LigPrep v2.5 of the Schrödinger suite. This was then docked in the Q module and ten poses were generated. Out of these poses, the most likely one was chosen based on parameters such as glide gscore, glide emodel value and essential interactions were confirmed by experimental mutagenesis data [Angerer *et al*, 2012].

2.7 *In silico* mutation analysis

Residue scanning wizard of BioLuminate 1.0 in Schrödinger suite was used to study the structural effects of mutations on the Q module. For individual mutations the difference in the stability of mutant compared to wild-type protein was estimated. Similarly, the difference in binding affinity of each mutant for the three iron-sulphur clusters, the DBQ and other interacting subunits in comparison with corresponding wild-type proteins was also considered.

2.8 MD simulation of subunit complex with bound nDBQ

In order to determine the effects of mutations on the binding of n-DBQ, MD simulations were performed on the subunit complex of *NDUFS2* and *7*. The mutants of *NDUFS7*, V122M and R145H were the only ones considered in this study as they have been further verified by experimentation.

Desmond Molecular Dynamics systems v3.1 [Guo *et al*, 2010] with Optimized Potentials for Liquid Simulations (OPLS) all atom force field 2005 [Kaminski *et al*, 2001, Jorgensen *et al*, 1996] was used to perform molecular dynamic simulations on w-t complex of *NDUFS2* and *7*, w-t without Fe-S cluster and n-DBQ bound mutant proteins V122M and R145H. A complex of *NDUFS2* and *7* was modelled using Prime v.3.1 [Jacobson *et al*, 2004] n-DBQ was docked at the interface of *NDUFS2* and *7* using Glide v.5.8 [Friesner *et al*, 2004]. Suitable pose was chosen depending on the Glide gscore and the presence of the essential hydrogen bond between Y141 of *NDUFS2* and head carbonyl group of n-DBQ. Mutants were prepared in the mutagenesis wizard of Maestro v9.3. Modelled protein bound with n-DBQ were prepared in the protein preparation wizard of Maestro v 9.3. Preparation of protein structures included addition and optimization of hydrogen atoms, generating metal binding states of the Fe-S cluster and restrained minimization using impref. The prepared structures were then uploaded in Desmond setup wizard and were solvated with SPC water model in an orthorhombic periodic boundary box so as to minimize system volume. Systems were neutralized using appropriate number of counterions. Energy of the prepared system was minimized up to maximum 1000 steps using steepest descent method until a gradient threshold (25 kcal/mol/A°) is reached followed by LBFGS (Low-memory Broyden-Fletcher-Goldfarb-Shanno quasi-Newtonian minimizer) until a convergence threshold of 1 kcal/mol/A° was achieved. The systems were equilibrated with the default parameters in Desmond v3.1 and MD simulations were carried out for 5 ns at a constant temperature of 300 K and a constant pressure of 1 bar with a time step of 2 fs. Long range electrostatic interactions were calculated by the smooth particle mesh Ewald method and a 9 A° cut-off radius was used for Coulombic short range interaction cut-off method.

The quality of simulations in terms of total energy, potential energy, temperature, pressure and volume were analyzed. The root mean square deviation

(RMSD), fluctuation of residues (RMSF), hydrogen bond between ligand and residues and distances were calculated by the Simulation Event Analysis module in Desmond v3.1. All the figures for the molecular modelling, docking and simulation studies were prepared in the PyMOL Molecular Graphics System, Version 1.5.0.4.

Cloning, expression and purification of the four core subunits of the Q module and the ND1 gene was attempted. Due to several problems associated with either cloning or expression of *NDUFS2, NDUFS8* and ND1 (cloning is described in Appendix), further characterization and study of these proteins was not successful. Thus, the other two proteins *NDUFS3* and *7* cloned were utilized for detailed study. The cloning, expression, purification and further biophysical studies performed on the purified proteins are described ahead.

2.9 RNA isolation, cDNA preparation, primer design and PCR amplification

Total RNA was isolated from the human colorectal adenocarcinoma cell line HT29 ($1x10^6$ cells) using Trizol® Reagent (Life Technologies, Cat#10296-010) as per the manufacturer's instructions. Purified RNA samples were analyzed by denaturing agarose gel electrophoresis and concentration was spectrophotometrically determined using Nanodrop (Thermo Scientific, USA). One µg of purified RNA was used for the preparation of cDNA using the SuperScript™ III First Strand Synthesis System (Life Technologies, Cat#18080-051).

2.9.1 Preparation of *NDUFS3* clone

Suitable primers (NDUFS3F: 5' ATC ATA TGG CGG CGG CGG C 3' & NDUFS3R: 5' TGC TCG AGC TAC TTG GCA TCA GGC TTC 3') were designed based on the RNA sequences downloaded from from National centre for biotechnology information NCBI (http://www.ncbi.nlm.nih.gov/) for amplification of the full length *NDUFS3* ORF. The PCR reactions were set in a 50 µl volume containing 1X Pfu buffer (20 mM Tris-HCl pH 8.8 at 25ºC, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100, & 0.1 mg ml^{-1} nuclease free BSA), 1 unit of Pfu polymerase, 200 µM each of dNTP and forward and reverse primers and 100 ng of the amplified cDNA. Cycling conditions were 95°C for 3 min, followed by 30 cycles of 95°C for 10 s, 50°C for 45 s, 72°C for 1 min and final extension of 72ºC for 20 min with hold of 25°C forever. The amplicon of desired size (808 bp) was gel

extracted using the QIAQuick Gel Purification kit (Qiagen, Cat#28704), which was used for TA cloning in pGEM-T vector (Promega).

2.9.2 Preparation of *NDUFS7* clone

Suitable primers (NDUFS7F: 5' GCC ATA TGG CGG TGC TGT CAG CTC 3' & NDUFS7R: 5' TGC TCG AGC TAC CTG CGG TAC CAG ATC 3') were designed based on the RNA sequences downloaded from NCBI (http://www.ncbi.nlm.nih.gov/) for amplification of the full length *NDUFS7* ORF. The PCR reactions were set in a 50 µl volume containing 1X Pfu buffer (20 mM Tris-HCl pH 8.8 at 25ºC, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100, & 0.1 mg ml^{-1} nuclease free BSA), 1 unit of Pfu polymerase, 200 µM each of dNTP and forward and reverse primers and 100 ng of the amplified cDNA. Cycling conditions were 95°C for 3 min, followed by 30 cycles of 95°C for 10 s, 50°C for 45 s, 72°C for 1 min and final extension of 72ºC for 20 min with hold of 25°C forever. The amplicon of desired size (656 bp) was gel extracted using the QIAQuick Gel Purification kit (Qiagen, Cat#28704), which was used for TA cloning in pGEM-T vector (Promega).

2.10 TA cloning and sub-cloning in pET bacterial expression vector

The purified amplicon was A-tailed using Taq polymerase in 1X Thermopol buffer (NEB) containing 200 µM dATP and 5 units of Taq DNA polymerase (NEB) at 72°C for 20 minutes. The A-tailed amplicons were then cloned into pGEM-T vector (Promega), followed by chemical transformation into *E. coli* DH5α (Invitrogen). The plasmids from the positive colonies, screened through colony PCR, were purified by the standard alkaline lysis method. Full length *NDUFS3* cDNA ORF or *NDUFS7* cDNA ORF from pGEM-T vector was further subcloned into pET-28b(+) vector (Novagen) between NdeI and XhoI sites. Sequences were confirmed by DNA sequencing using BigDye™ Terminator Cycle Sequencing Ready Reaction Kit v3.1 (ABI, Cat#4337457) in an automated 3730 DNA analyzer (ABI). Appropriate plasmids were transformed into *E. coli* BL21(DE3*)* for protein expression.

2.11 Site-directed mutagenesis

2.11.1 *NDUFS3*

The double mutant T145I-R199W is known to cause Leigh syndrome. A two step approach was used to prepare the double mutant (T145I-R199W) of *NDUFS3*. Two sets of primers were designed for the same. The SDMF1 5'CGG ATC CGT GTG AAG ATC TAC ACA GAT GAG CTG3', SDMR1 5'CAG CTC ATC TGT GTA GAT CTT CAC ACG GAT CCG3' and SDMF2 5'CTT CGA GGG ACA TCC TTT CTG GAA AGA CTT TCC TCT ATC3', SDMR2 5'GAT AGA GGA AAG TCT TTC CAG AAA GGA TGT CCC TCG AAG3'. Site-directed mutagenesis was carried out with the help of Phusion™ Site-Directed Mutagenesis kit (Finnzymes, Cat#F541). The PCR reactions were set in a 50 µl volume containing the w-t *NDUFS3* gene cloned in pGEM-T vector, 1X HF buffer, 200 µM each of dNTP and forward and reverse primers, 0.02U/µl Phusion Hot Start DNA polymerase. Cycling conditions were 95°C for 3 min, followed by 30 cycles of 95°C for 10 s, 55°C for 5 min, 68°C for 6 min and final extension of 72ºC for 20 min with hold of 25°C forever. In the first PCR cycle the first set of primers (SDMF1 and SDMR1) were used. The purified amplicons were phosphorylated and ligated to circularize the plasmid and transformed into *E. coli* DH5α cells. The plasmids were sequenced and mutant plasmids were used for a second round of PCR with the second set of primers (SDMF2 and SDMR2). The PCR protocol was repeated and the plasmids were sequenced to choose the suitable construct. The double mutant of *NDUFS3* was further sub-cloned in the pET-28b(+) vector (Novagen) between NdeI and XhoI restriction enzyme sites and the plasmids were transformed in *E coli* BL21(DE3) for protein expression. Sequences were confirmed by DNA sequencing using BigDye™ Terminator Cycle Sequencing Ready Reaction Kit v3.1 (ABI, Cat#4337457) in an automated 3730 DNA analyzer (ABI) at each step.

2.11.2 *NDUFS7*

Two mutants containing point mutations V122M and R145H were prepared. Suitable primers were designed for the same. SDMF1 5'GAT CAT GGC CGG CAC ACT CAC CAA CAA GAT GGC 3', SDMR1 5'ATG ACG TCG GAC TGG CGC GGG CTG 3' and SDMF2 5'CCG CAC TAC GTG GTC TCC ATG GGG AG 3', SDMR2 5'CTC CGG CAT CTG GTC GTA GAC CTT GCG AA 3'. SDMF1 and SDMR1

were used for V122M mutation and SDMF2 and SDMR2 were used for R145H mutation. Site-directed mutagenesis was done with the help of Phusion™ Site-Directed Mutagenesis kit (Finnzymes, Cat#F541). The PCR reactions were set in a 50 µl volume containing the w-t *NDUFS7* gene cloned in pGEM-T vector, 1X HF buffer, 200 µM each of dNTP and forward and reverse primers, 0.02U/µl Phusion Hot Start DNA polymerase. Cycling conditions were 95°C for 3 min, followed by 30 cycles of 95°C for 10 s, 55°C for 5 min, 68°C for 6 min and final extension of 68ºC for 20 min with hold of 25°C forever. The purified amplicons were phosphorylated and ligated to circularize the plasmid and transformed into *E. coli* DH5α cells. The plasmids were sequenced to confirm the mutagenesis. Both the mutants of *NDUFS7* were further sub-cloned in the pET-28b(+) vector (Novagen) between NdeI and XhoI restriction enzyme sites and the plasmids were transformed in *E coli* BL21(DE3) for protein expression. Sequences were confirmed by DNA sequencing using BigDye™ Terminator Cycle Sequencing Ready Reaction Kit v3.1 (ABI, Cat#4337457) in an automated 3730 DNA analyzer (ABI) at each step.

2.12 Protein expression and solubility

A fresh sterile plate of LB containing 60 µg/ml Kanamycin was streaked using the glycerol stock of BL21(DE3) *E. coli* cells transformed with suitable gene cloned in pET-28b(+) expression vector. The plate was incubated at 37°C for 14 hours. A single colony from the plate was further inoculated in 5 ml sterile LB containing 60 µg/ml kanamycin. The culture was incubated for 12-14 hours at 37°C with a shaking of 150 rpm.

Autoinduction method was used to obtain soluble expression of the protein. This method as described by Studier [Studier, 2005] utilizes the Rich (ZYM-5052) autoinduction medium for soluble expression of proteins. In short, the media comprises of the following solutions:

Stock solutions

- Sterile 1 M $MgSO_4$ stock solution
- Sterile 50 X 5052 solution

  For 100 ml sequentially dissolve 2.5 g glucose and 10 g α-lactose in 75 ml water and finally add 25 g glycerol.

- Sterile 25 X M solution

  For 200 ml sequentially dissolve in 80 ml water 3.6 g Na_2SO_4 anhydrous, 13.4 g NH_4Cl anhydrous, 17.0 g KH_2PO_4 anhydrous and 17.7 g Na_2HPO_4 anhydrous. Make up the volume to 200 ml.

- ZY medium

  For 1000 ml medium, dissolve 10 g N-Z-amine AS (or tryptone), 5 g Yeast extract in 937 ml water.

- 1000x metal solution

  For 100 ml, autoclaved solutions of 50 ml 0.1 M $FeCl_3$-$6H_2O$ (in 0.1 M HCl), 2 ml 1 M $CaCl_2$, 1 ml 1 M $MnCl_2$-$4H_2O$, 1 ml 1 M $ZnSO_4$-$7H_2O$, 1 ml 0.2 M $CoCl_2$-$6H_2O$, 2 ml 0.1 M $CuCl_2$-$2H_2O$, 1 ml 0.2 M $NiCl_2$-$6H_2O$, 2 ml 0.1 M Na_2MoO_4-$5H_2O$, 2 ml 0.1 M Na_2SeO_3-$5H_2O$ and 2 ml 0.1 M H_3BO_3 were mixed in 36 ml water.

All the above solutions were autoclaved seperately at 120°C for 15 minutes.

The Rich medium for autoinduction was prepared by mixing the above stock solutions aseptically in the following ratios

Table 2.4: Composition of the Rich ZY medium.

Stock solution	For 1000 ml of Rich ZY medium
1M $MgSO_4$	1 ml
1000X metals	1 ml
50 X 5052	20 ml
25 X M	40 ml
Kanamycin (60 mg ml^{-1})	1 ml
ZY medium	Make up the volume to 1000 ml

The expression protocol in short involves, inoculating 1 ml of the primary culture in fresh LB medium with Kanamycin (5 ml) and incubation at 37°C to activate dormant cells. In the meanwhile, autoinduction media is prepared by adding required amount of ZY medium, 50 X 5052 solution, 25 X M solution and Kanamycin (Table 2.4) in a sterile flask under aseptic conditions. 10 ml of the primary culture was added to 1000 ml of the Rich ZYM autoinduction medium. The inoculated medium was kept at 37°C shaker for 2 hours. The culture was then incubated at 16°C for 16 hours.

Following prolonged incubation, the cells were harvested by spinning at 4000 rpm for 15 minutes at 4°C. The cell pellet was re-dissolved in suitable amounts of Lysis Buffer (50 mM Tris-Cl pH 8.5, 100 mM NaCl, 0.01% IGEPAL® CA-630 and 1 mg/ml lysozyme) for *NDUFS3* and (20 mM sodium phosphate buffer pH 7.0, 300 mM NaCl, 0.01% IGEPAL® CA-630 and 1 mg/ml lysozyme) for *NDUFS7* and was kept in ice for 30 minutes with intermittent shaking. Cells were lysed by sonication at 60% power in Esquire Biotech Ultrasonic homogenizer with a pulse of 6 s followed by pause of 8 s for 10 minutes. The sonicated solution was centrifuged at 10,000 rpm for 45 minutes at 4°C. The pellet and supernatant were separated. The expression profile was checked on 12% SDS-PAGE gel followed by Coomassie Brilliant Blue R-250 staining.

2.13 Purification of *NDUFS3* and *NDUFS7*

2.13.1 *NDUFS3*

The protein from the crude cell lysate was purified in two steps of column chromatography, anion exchange followed by size exclusion chromatography. In short, the supernatant obtained from the lysed cells was centrifuged at 10,000 rpm for 15 minutes at 4°C. It was then used for loading directly onto a Q-Sepharose (Sigma-Aldrich, Cat#Q1126) column pre-equilibrated with 50 mM Tris-Cl pH 8.5, 100 mM NaCl and 0.01% IGEPAL® CA-630. The column was washed with the above buffer till the eluate showed no absorbance at 280 nm. The column was subjected to a NaCl gradient from 0.1 M to 0.5 M and final wash was given with the buffer containing 1 M NaCl. Absorbance of the fractions was recorded at 280 nm. The fractions showing OD >0.2 were run on 12% SDS-PAGE gel to detect the presence of the protein. The fractions showing the protein with desired molecular weight (~35 kDa) were pooled and concentrated in a Labcono vacuum concentrator. The OD_{280} was noted and concentration of the protein solution was checked using Bradford assay [Bradford, 1976]. The Sephadex G-200 column (Sigma-Aldrich, Cat#G-200-120) was pre-equilibrated with 50 mM Tris-Cl pH 8.5, 300 mM NaCl and 0.01% IGEPAL® CA-630. The protein was loaded onto the column and washed with three column volumes of equilibration buffer. Absorbance of the fractions was measured at 280 nm. The fractions showing an OD >0.2 were run on 12% SDS-PAGE gel to identify protein

fractions. Both w-t and mutant (T145I-R199W) proteins were purified by this protocol.

2.13.2 *NDUFS7*

The supernatant obtained from the lysed cells was used for loading directly on Ni-NTA agarose (Life technologies, Cat#R109-15)) column pre-equilibrated with 20 mM sodium phosphate buffer pH 7.0, 300 mM NaCl, 20mM imidazole and 0.01% IGEPAL® CA-630. The column was washed with the above buffer till the eluate didn't show any absorbance at 280 nm. The column was subjected to an imidazole gradient from 0.1 M to 0.4 M and final wash was given with buffer containing 0.5 M imidazole. Absorbance of the fractions was taken at 280 nm. The fractions showing an OD >0.2 were run on a 12% SDS-PAGE gel to detect the presence of the protein of interest. The fractions showing the protein with desired molecular weight (~20 kDa) were pooled and dialyzed to remove the imidazole and concentrated in a Labcono vacuum concentrator. The OD_{280} and concentration of the concentrated protein was checked using Bradford's assay [Bradford, 1976]. The Superose 12 column (GE Healthcare, Cat#G-17-5173-01) was pre-equilibrated with 20 mM sodium phosphate buffer pH 7.0, 150 mM NaCl and 0.01% IGEPAL® CA-630. The protein was loaded onto the column and was subjected to 2 column volumes of equilibration buffer washes. Absorbance of the fractions was taken at 280 nm. The fractions showing an OD >0.2 were run on 12% SDS-PAGE gel to detect the presence of the protein. Western Blot with anti-His tag antibody and MALDI TOF/TOF was used to confirm the presence of the protein and its molecular weight.

2.14 Western Blot and MALDI-TOF/TOF™

Western blots were prepared by electroblotting the SDS-PAGE gels onto PVDF membrane [Towbin *et al*, 1979]. The membrane was blocked with 5% fat-free skimmed milk in PBS for 60 minutes followed by 3 washes of 5 minutes each in PBS containing 0.05% TWEEN® 20. The membrane was further incubated with the monoclonal anti-polyhistidine antiboby (Sigma-Aldrich, Cat# H1029) at 1:1000 dilution in PBS containing 1% BSA overnight. The membrane was washed 3 times for 5 minutes each in PBS containing 0.05% Tween®20 and incubated with anti-mouse Ig-G (Fc-specific) Peroxidase conjugate antibody in PBS and Tween®20 at a

concentration of 1:6000 (Sigma-Aldrich, Cat#A0168). The membrane was further treated with Novex HRP Substrate (Invitrogen Cat#WP 20004) to visualize the bands.

For MALDI-TOF/TOF™ the concentration of the purified protein sample was determined by Bradford assay. 10µg ml^{-1} sample was mixed with a saturated solution of α-cyanohydroxy cinapinic acid in a mixture of 0.1% TFA:ACN and spotted onto the MALDI plate. The plate was dried at room temperature and was analyzed with AB SCIEX TOF/TOF™ 5800 System for MALDI mass spectrometry imaging.

2.15 Biophysical characterization

2.15.1 Circular Dichroism (CD) Spectroscopy

CD spectra of the purified proteins were recorded using a Jasco J-815-150S (Jasco, Tokyo, Japan) spectropolarimeter connected to a Peltier Type CD/FL Cell circulating water bath at room temperature. Far UV spectra was recorded in a rectangular quartz cell of 1 mm path length in the range of 200-250 nm at a scan speed of 100 nm min^{-1} with a response time of 1 s and a slit width of 1 nm. Purified w-t *NDUFS3* at a concentration of 0.05 mg ml^{-1} and its mutant (T145I-R199W) at a concentration of 0.1 mg ml^{-1} was used for all the far-UV CD samples. Purified w-t *NDUFS7* at a concentration of 0.08 mg ml^{-1} was used for all the far-UV CD samples. Near UV CD spectrum was recorded in the range of 250-300 nm with a protein concentration of 1 mg ml^{-1} for all the proteins. Each spectrum was recorded as an average of 5 scans.

Conformational transition studies of w-t, mutant *NDUFS3* and w-t, mutant *NDUFS7* at various conditions of pH, chemical denaturant and temperature were performed.

2.15.1.1 pH variation: The purified proteins were incubated against buffers of different pH for 4 hours before recording the spectra. The different buffers used were: 20 mM glycine-HCl (pH 1.0-3.0), 20 mM citrate-phosphate buffer (pH 4.0-6.0), 20 mM tris-HCl (pH 7.0-9.0), 20 mM glycine-NaOH (pH 10.0-12.0). The native tertiary structure of proteins was studied by recording the near UV CD spectrum in the range of 250-300 nm with a working concentration range of 1 mg ml^{-1}.

2.15.1.2. Temperature dependence: The CD spectra of the purified proteins were recorded by increasing the temperature of the samples at the rate of 2°C min^{-1} within

the temperature range of 25-90°C. Ellipticity was recorded at a temperature interval of 5°C and incubation time of 5 minutes was maintained between 200-250 nm.

2.15.1.3. Gdn-HCl concentration: The purified proteins were incubated in Gdn-HCl in the concentration range of 0-4 M at pH 8.5 for 6 hours before recording the spectra. Buffer scans were subtracted from each spectrum for further analysis.

Results were determined in terms of the mean residue ellipticity (MRE). CD in millidegrees was converted to mean residue ellipticity by the formula:

$(\theta)_\lambda = M\theta_\lambda / 10dcr$

where, M is the molecular mass, θ_λ is the ellipticity in millidegrees, d is the cuvette path length in cm and c is the concentration of the protein in mg ml^{-1}, r is the number of residues. Secondary structure content was estimated using the CDPro program (CDSSTR, CONTIN, SELCON3) [Sreerama *et al*, 2004].

2.15.2. Steady State Fluorescence Spectroscopy

The intrinsic fluorescence of the proteins was recorded using a Perkin Elmer LS50 fluorescence spectrophotometer connected to a Julabo F20 water bath. The spectra were recorded using a quartz cuvette at room temperature. The background emission due to the buffer or denaturants was subtracted for further analysis. The protein solutions were excited at 295 nm and the emission spectra were recorded between 300-400 nm setting the slit width 7 nm and speed 100 nm min^{-1}. The effect of pH was recorded by incubating the protein in suitable buffers of different pH (1-12) as described in the section 2.15.1.1. Thermal unfolding was monitored by incubating the protein for 5 min at the desired temperature followed by recording the spectra. Thermal aggregation was studied by the Rayleigh scattering measurement on the same instrument under different conditions of protein concentration (25-200 µg ml^{-1}), in the presence of salt (5-355 mM NaCl) and detergent (0.05-0.16 mM IGEPAL® CA-630). Chemical unfolding was monitored by recording the spectra of the protein incubated in Gdn-HCl (0-6 M) for 6 hours at room temperature. The decomposition analysis of tryptophans was performed by the PFAST: Protein Fluorescence and Structural Toolkit [Shen *et al*, 2008; Hixon *et al*, 2009].

2.15.3 Solute Quenching studies

2.15.3.1. Steady-state Fluorescence Quenching

Fluorescence quenching measurement was done with quenchers: acrylamide (5 M), caesium chloride (5 M) prepared in milliQ water and KI (5 M) prepared in 0.2 M sodium thiosulphate by titrating 80 μg ml^{-1} of the protein solution prepared in 50 mM Tris-Cl pH 8.5, 100 mM NaCl and 0.01% IGEPAL®CA-630. Quencher from the stock solutions was added till the final concentration of quencher in the protein solution reached 0.5 M. The fluorescence intensities were measured at the wavelength corresponding to maximum emission and volume correction was done before analyzing the quenching data [Lakowicz, 1983]. Stern-Volmer analysis of the quenching data was used to estimate K_{sv} and modified Stern-Volmer plots for determining fa (fraction accessibility) [Lehrer *et al*, 1978]. The Stern-Volmer equation given below was used for the same.

$$Fo/Fc=1+K_{sv}[Q] \qquad (1)$$

Where Fo and Fc represent the relative fluorescence intensities corrected for dilution in absence and presence of the quencher [Q]. K_{sv} is the Stern-Volmer constant for the given quencher.

2.15.3.2. Fluorescence lifetime measurement

Lifetime measurements were recorded on Edinburgh Instruments FLS-920 single photon counting spectrofluorimeter. The excitation source was a laser pico second pulsed light emitting diode (model EPLED-295) and fluorescence was detected by a synchronization photomultiplier. Diluted Ludox solution was used for measuring the Instrument response function (IRF). The samples (1 mg ml^{-1}) were excited at 295 nm and emission was recorded at 342 nm, 346 nm and 356 nm for w-t, mutant and denatured protein samples. Slit widths of 15 nm each were used for the excitation and emission monochromators. The resultant decay curves were analyzed by a multiexponential iterative fitting program provided by Edinburgh Instruments. The average lifetimes after each step of quenching was calculated using the equations [Inokuti *et al*, 1965; Grinvald *et al*, 1974]

$$\tau=\sum_i \alpha_i \tau_i/\sum_i \alpha_i \qquad (2)$$

$$<\tau>=\sum_i \alpha_i \tau_i^2/\sum_i \alpha_i\tau_i \quad (3)$$

Where i=1,2,...

The average lifetimes τ and <τ> were obtained by two different approaches and the plot of τo/τ against quencher concentration [Q] gave the K_{sv}. The bimolecular quenching constant, *kq* was calculated using equation given below [Lehrer, 1971].

$$kq=K_{sv}/\tau \quad (4)$$

The dynamic and static components were resolved by using the equation shown below [Lacowicz *et al*, 1973]:

$$Fo/Fc=(1+K_{sv}[Q])(1+K_s[Q]) \quad (5)$$

Where K_{sv} is the Stern-Volmer (dynamic) quenching constant, K_s is the static quenching constant and [Q] is the concentration of quencher. The dynamic quenching constant is the reflection of the degree to which the quencher achieves the encounter distance of the fluorophore and is determined by the fluorescence lifetime measurement fitted to the equation [Lacowicz *et al*, 1973]:

$$\tau o/\tau = (1+Ksv[Q]) \quad (6)$$

where τo is the average lifetime in the absence of quencher and τ is the lifetime in the presence of the quencher at concentration [Q].

2.15.4. Hydrophobic dye binding

Binding of the hydrophobic dye 8-Anilino-1-naphthalene sulfonic acid (ANS) was studied by recording the emission spectra between 400-550 nm post excitation at 375 nm using a steady state spectrofluorimeter. ANS has been shown to bind to hydrophobic regions of partially unfolded proteins that are exposed to the solvent [Semisotnov *et al*, 1991]. The proteins were incubated at various pH (1-12), Gdn-HCl concentrations (0-6 M) and at different temperatures (25-90°C) to study the unfolding of polypeptide chain. 5 µl of 15 mM ANS was mixed with 2 ml of protein solution (0.05 mg ml^{-1}). Buffer spectrum with ANS present was subtracted in each case for further analysis.

2.15.5. Assays for studying protein aggregation

2.15.5.1. Rayleigh scattering

The measurement of scattering of the w-t and mutant proteins were carried out in the time drive module of the Perkin Elmer LS50 fluorescence spectrophotometer connected to a Julabo F20 water bath. Measurements were recorded at 28°C and 37°C. Buffer values were subtracted in each case for further analysis.

2.15.5.2 Thioflavin-T (Th-T) binding assay

The amyloidophilic Thioflavin-T dye shows a characteristic increase in fluorescence upon binding to amyloid fibrils. Binding of Thioflavin-T was studied by recording the fluorescence emission spectra between 450-600 nm with an excitation at 442 nm using a steady state fluorescence spectrofluorimeter [Groenning, 2010; Chang *et al*, 2009; Khurana *et al*, 2005]. In short, a stock solution of 2 mM Th-T was prepared in 150 mM NaCl and 100 mM sodium phosphate buffer (pH 7.0) and filtered through 0.22 μm Millipore filter. A fresh working solution was prepared by adjusting the dye concentration to 200 μM. Protein samples (25-100 μg ml^{-1}) were incubated at 25°C and 37°C at pH 8.5 for 1 hour and further used for the assay. A 250 μl aliquot of sample solution was mixed with a 250 μl Th-T working solution and the spectra were recorded. Buffer spectrum with Th-T was subtracted in each case for further analysis.

2.15.5.3 Congo red (CR) assay

CR is similar to Th-T and binds to β-rich structures, inducing an increase in absorption and a red shit in the CR absorption band from 490 to 540 nm. Binding of Congo red was studied by recording the absorbance spectra on a UV-Vis spectrophotometer between 380-700 nm [Chang *et al*, 2009; Klunk *et al* 1999]. In short a 2 mM CR stock solution prepared in 150 mM NaCl and 5 mM potassium phosphate buffer (pH 7.5) was filtered through 0.22 μM Millipore filter. A fresh working solution was prepared by adjusting dye concentration to 200 μM. Protein samples were incubated at 25°C and 37°C for 1 hour. A 250 μl protein sample (25-100 μg ml^{-1}) was mixed with 50 μl of the working stock solution and 700 μl of the above buffer and the UV spectrum was recorded. Buffer spectrum with CR was subtracted in each case for further analysis.

2.16 Fe-S cluster detection

W-t *NDUFS7* possesses an iron sulphur cluster N2. A spectrum was recorded on a UV-visible spectrophotometer between 300-800 nm for the w-t, V122M, R145H and w-t after removal of the Fe-S cluster for the detection of this cluster. Buffer spectrum was subtracted in each case for further analysis. The Fe-S cluster was removed from the w-t by incubating the protein in 5 mM DTT and 2 mM EDTA for 1 hour. Far and near-UV CD spectra were recorded for the w-t, R145H, V122M with Fe-S cluster and then with w-t without Fe-S cluster to detect structural changes.

2.17 n-DBQ binding assay

NDUFS7 possess a partial binding site for nDBQ. In order to explore the binding affinity of n-DBQ to protein; a fluorescence based assay was designed. The w-t and V122M, R145H mutant proteins in 20 mM sodium phosphate buffer pH 7.4, 100 mM NaCl and 0.01% IGEPAL-CA630 were placed in a quartz cuvette and maintained at a desired temperature (28°, 32°, 37° and 42°C) with a Julabo circulating water bath. A 0.1 mM stock solution of n-DBQ was prepared in dimethylsulfoxide (DMSO). The n-DBQ solution was added in 12-14 aliquots (3-5 µl) each and fluorescence spectra were recorded with each addition. Samples were excited at 295 nm and spectra were recorded between 310- 400nm. The fluorescence intensity at 341 nm (λ_{max} of w-t and mutant proteins) was considered for further analysis. Corrections were made to compensate for the dilution effect upon addition of nDBQ. Each data point was an average of 3 independent sets of experiments with standard deviation less than 5%.

The association constants were calculated by the method of Chipman *et al* (Chipman *et al*, 1967). The abscissa intercept of the plot of log $[C]_f$ against log $\{(\Delta F)/(Fc-F\infty)\}$, where $[C]_f$ is the free concentration of n-DBQ, yielded pKa value for protein-nDBQ interaction according to

$$\log [(Fo-Fc)/(Fc-F\infty)] = \log Ka + \log \{[C]_t-[P]_t (\Delta F/\Delta F\infty)\} \quad (1)$$

where, Fc is the fluorescence intensity of protein at any point during the titration, $[P_t]$ is the total protein concentration, $\Delta F\infty$ is the change in fluorescence intensity at saturation of binding, $[C_t]$ is the total n-DBQ concentration given by,

$$[C]_f = \{[C]_t – [P]_t (\Delta F/ \Delta F\infty)\} \quad (2)$$

Free energy changes of association (ΔG) were determined by:

$$\Delta G = -RT \ln K_a \quad (3)$$

Temperature dependence of the association constants was used to determine thermodynamic parameters. Change in enthalpy (ΔH) was determined from the Van't Hoff plot by equation:

$$\ln K_a = (-\Delta H/RT) + \Delta S/R \quad (4)$$

where, ΔH is enthalpy change, R is gas constant, ΔS is entropy change and T is the absolute temperature. The entropy change was obtained from the equation

$$\Delta G = \Delta H - T\Delta S \quad (5)$$

CHAPTER 3

In silico studies on the effects of Leigh syndrome mutations on the structure and function of human mitochondrial Complex-I

The human mitochondrial Complex-I is a relatively less explored protein complex owing to its large size. For several years it was referred to as an 'L-shaped black box' because of its mutation prone nature and lack of structural information. About 40% of the mitochondrial diseases reported are owing to Complex-I deficiency. Even then the structural studies on this complex are limited owing to its large size and complex nature. The recent structure from the bacterial homologue *Thermus thermophilus* gave several insights in the functioning and structural aspects of the complex. Like the bacterial homologue, the human Complex-I is also divided into 3 domains which co-ordinate for the smooth functioning of the Complex-I. These are,

1. Dehydrogenase domain (N module)
2. Hydrogenase or ubiquinone reduction domain (Q module)
3. Membrane or proton pumping domain (P module)

The 14 core subunits forming the minimal complex in prokaryotes are present in the human Complex-I as well. These core subunits show ~40-45% identity with their prokaryotic counterpart. The eukaryotic proteins, in certain cases, carry an additional sequence located in between the signal sequence and the prokaryotic homologous sequence. This sequence is thought to be essential for binding with the accessory or supernummery subunits. These subunits provide protection to the core region from oxidative injury, are involved in regulation of activity and their role in assembly and stabilization of Complex-I has also been speculated.

3.1 Human Q module

The human Q module or ubiquinone reduction domain also acts as a connecting domain between the N and P modules. It is involved in the transfer of electrons from NADH to the ultimate electron acceptor ubiquinone. It harbors the ubiquinone binding and reduction site. It is made up of 4 core subunits; *NDUFS2, 3, 7* and *8* and possesses three iron sulphur clusters, N6a, N6b and N2. Their function is to transfer electrons for ubiquinone reduction.

The Q module possesses the binding site for ubiquinone as well as several Complex-I inhibitors. The ubiquinone binding site lies at the interface of *NDUFS2* and *7*.

Ubiquinone plays a dual role by acting as a substrate by accepting electrons during NADH oxidation and as a tightly bound co-factor during electron recycling [Prieur *et al*, 2001]. Of the three iron sulphur clusters, cluster N2 acts as an electron donating group to ubiquinone. This cluster, located at the end of the redox centres, transfers electrons from the flavin site to quinine [Hinchliffe *et al*, 2005]. This is also suggested to play a role in coupling the redox reactions that translocates protons across the membrane [Berrisford *et al*, 2009]. Thus, Q module is also the electron donor for the ubiquinone [Tocilescu *et al*, 2010; Angerer *et al*, 2012]. In addition to ubiquinone, *NDUFS2* harbours binding sites for quinone related Complex-I inhibitors such as piericidin and rotenone [Darrouzet *et al*, 1998].

3.1.1 Mutations in the human Q module

A large number of mutations have been reported in the subunits forming the Q module resulting in different neuromuscular disorders. Proper assembly and stabilization of Complex-I is essential for its activity as well as structural stability of the inner mitochondrial membrane. Deleterious mutations perturb the assembly of Complex-I and induce structural changes in the mitochondrial membrane causing proton leakage, accumulation of reactive oxygen species and thereby release of apoptotic factors, which remain the primary causes of these diseases [Ugalde *et al*, 2004a]. Reports of BN-PAGE studies indicate that the mutations affect the assembly process at an early stage resulting in the accumulation of assembly intermediates. The various stages of subunit association are determined such that *NDUFS2* and *NDUFS3* assemble first followed by *NDUFS7* and finally *NDUFS8*, which indicate that a mutation in any one subunit would affect assembling successfully into full functional module [Ugalde *et al*, 2004b]. Lack of a properly assembled Complex-I results in severe energy deficiency affecting muscles and nerves. Also, normally, the iron-sulphur clusters in Complex-I are present close to each other so as to reduce the ROS production, however, lack of proper assembly would lead to the excessive production of ROS causing further damage to the cell. This causes a wide range of symptoms that depict the disorder. The mutations of the Q module and their clinical phenotypes considered in the present study are enlisted in Table 3.1.

Table 3.1: List of mutations and their clinical phenotypes considered in the present study.

Subunit	Mutation	Clinical phenotype	Reference
NDUFS2	R228Q-S413P	Leigh syndrome	Loeffen *et al* 2001
	R138Q-R333Q	Leigh like syndrome	Tuppen *et al* 2010
	M292T-RR118Q	Leigh syndrome	
	M292T-R443K	Leigh syndrome	
	M292T-E148K	Leigh syndrome	
NDUFS3	T145I-R199W	Leigh syndrome	Benit *et al* 2004
NDUFS7	V122M	Leigh syndrome	Triepels *et al* 1999
	R145H	Leigh like syndrome	Lebon *et al* 2007a
NDUFS8	P85L-R138H	Leigh syndrome	Procaccio *et al* 2004

In the present study, the core subunits of the human Q module have been modeled individually. The arrangement of subunits in the human Q module has been studied in order to understand the positions of the mutations with respect to the other subunits and the binding site. Molecular dynamic simulations have been done to study the changes in w-t and mutant structures. *In silico* energy calculations have been useful to observe the destabilizing effect of the mutations on the structure of the complex.

3.1.2 Core subunits of the human Q module

NDUFS2, 3, 7 and *8* are the core subunits of the Q module. *NDUFS2* and *7* are the functional subunits while *NDUFS3* and *8* are the structural subunits. These four subunits are highly conserved across all species. They show a very high identity with their homologues.

Phylogenetic analysis of the four subunits showed that the sequences clustered into primarily 3 groups: Eubacteria, insects and mammals (Figure 3.1). This clustering was mainly due to the extra signal sequence in the eukaryotes. In certain cases like

NDUFS2 and *7*, in eukaryotes, apart from the signal sequence an additional stretch of amino acid sequence is also present, which may probably bind to the accessory subunits present in eukaryotes.

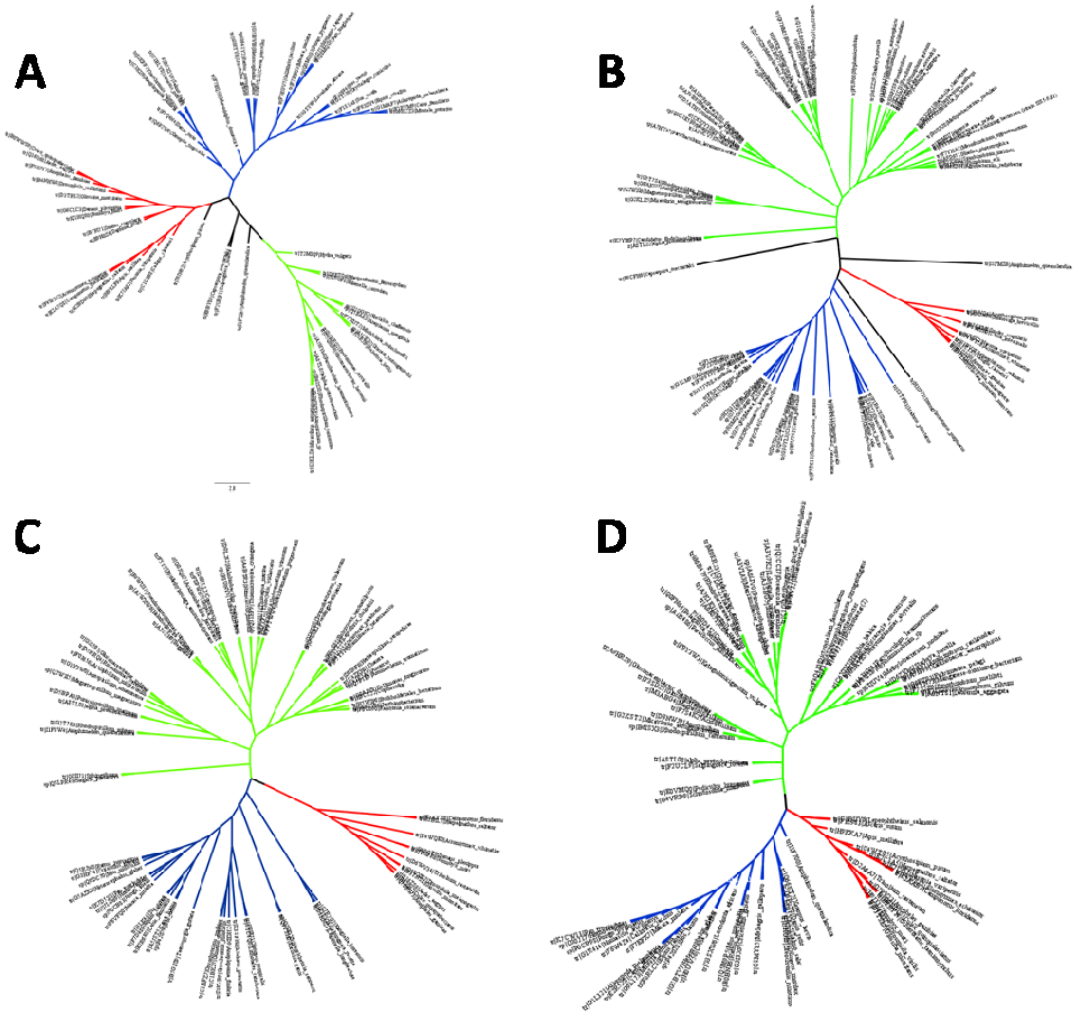


Figure 3.1: Phylogenetic trees for the 4 core subunits of the human Q module; A: *NDUFS2* B: *NDUFS3,* C: *NDUFS7* and D: *NDUFS8*, The blue branch indicates mammals, red branch indicates insects and green branch indicates eubacteria. The trees were prepared in MEGA 5.5 and plotted using Figtree 1.3.1.

Apart from the signal sequence and some extra amino acid sequence, all the 4 subunits show a highly conserved C-terminal part (upto 45% identity) with the homologous subunits of prokaryote *Thermus thermophilus*.

3.1.3 Modeling of the core subunits

The 4 core subunits of human Complex-I were modeled in Modeller v9.10. The subunits showed a high sequence identity with homologous subunits in the *T. thermophilus.* The templates were chosen based on sequence identity, coverage and resolution of template structure (Table 3.2).

Table 3.2: Templates used for modeling the subunits of the human Q module.

Subunit	Templates				
	PDB ID	**Chain**	**Sequence Identity (%)**	**Resolution (A°)**	**Coverage**
NDUFS2	3I9V	D	46	3.1	73-463
NDUFS3	3IAM	E	42	3.1	107-204
NDUFS7	3I9V	F	53	3.1	58-202
NDUFS8	3IAM	G	48	3.1	64-174

Molecular models were generated for each subunit using corresponding templates. The signal sequence was not modeled (Table 2.1). The N-terminal residues, first 49 of *NDUFS2* and 57 of *NDUFS8*, following the signal peptide, could not be modeled due to the absence of suitable template for that part of the sequence (Figure 3.2).

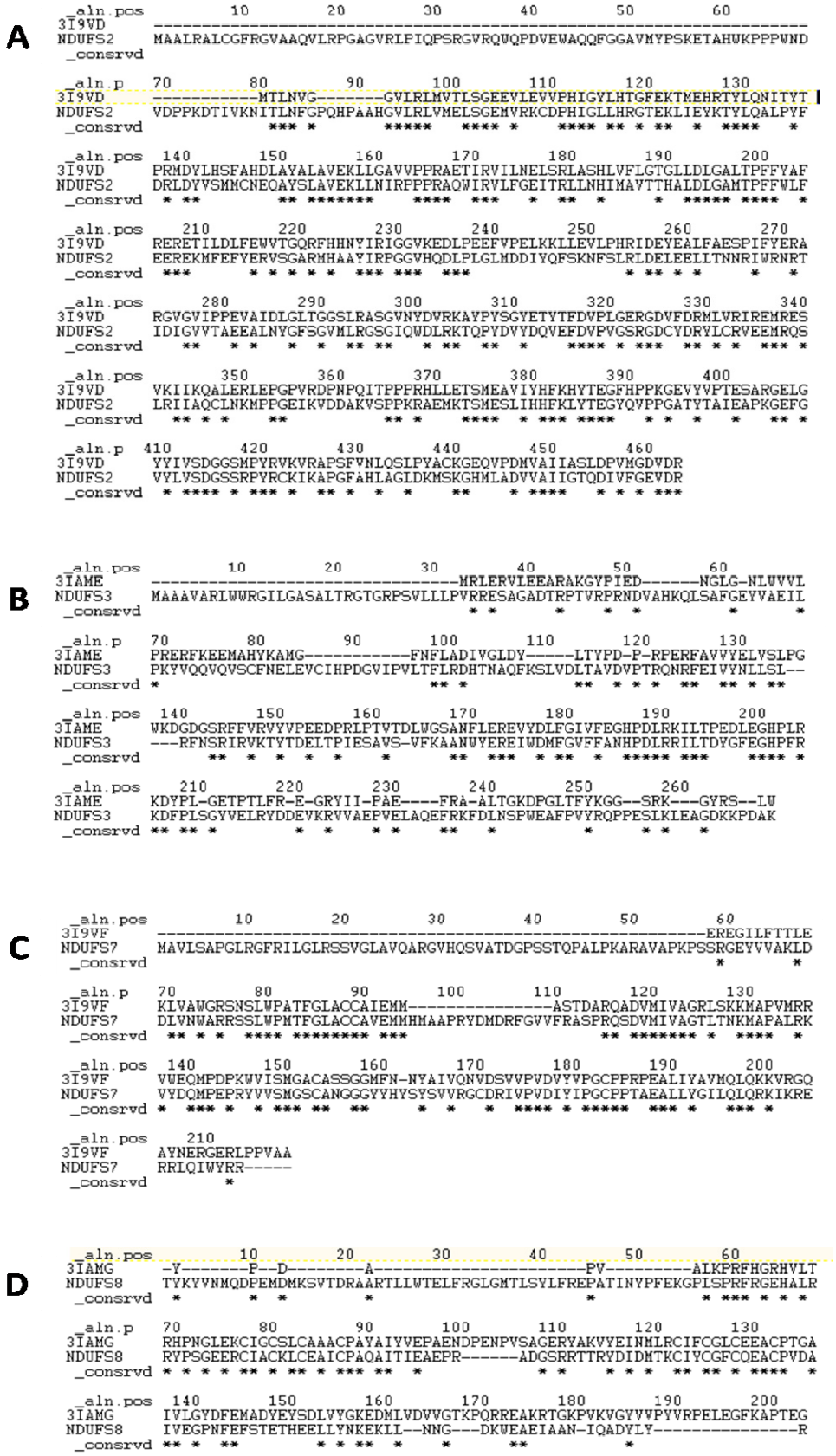


Figure 3.2: Sequence alignment of the human subunits with the respective templates. A: NDUFS2 with 3I9V-D, B: NDUFS3 with 3IAM-E, C: NDUFS7 with 3I9V-F and D: NDUFS8 with 3IAM-G.

The quality of the models was checked and the best model among them were chosen depending on the DOPE score, ERRAT score, PDBsum analysis and ProSA-web protein structure analysis (Table 3.3).

Table 3.3: Estimated values of evaluation parameters to assess the quality of molecular models and validation scores to select the best model.

Subunit	DOPE score	ERRAT Score	Ramachandran Plot number of residues		G score	Z score	RMSD from template (Å)
			Allowed region (%)	Disallowed region (%)			
NDUFS2	-46629	78.41	99.7	0.3	-0.3	-8.53	0.49
NDUFS3	-19345	85.14	99	1	-0.49	-1.68	1.36
NDUFS7	-15989	95.45	100	0	-0.32	-4.58	0.77
NDUFS8	-12353	79.01	97	3	-0.4	-1.85	0.45
Q module	NA	73.086	98.8	1.2	-0.4	-9.29	0.41

The models with best validation parameters are chosen for further study (Figure 3.3). *NDUFS2* and 7 are primarily helical structures with a few β-sheets present while *NDUFS3* and *8* comprise of loops and β-sheets along with few helices.

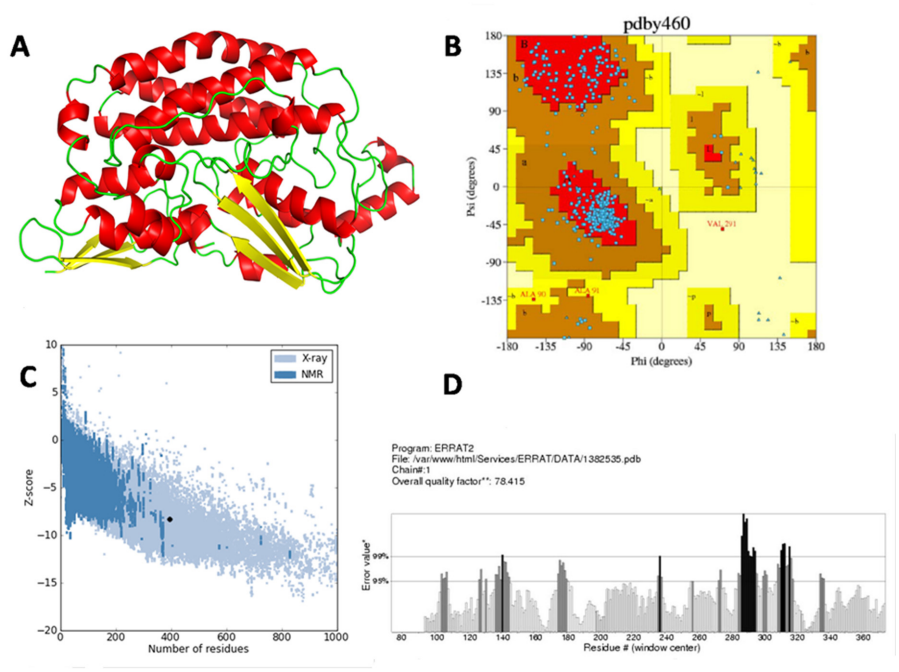


Figure 3.3.1: A: Molecular model of *NDUFS2* prepared using pdb structure 3I9V(D-chain) as template, B: Ramachandran plot, C: ProSa analysis, D: ERRAT score.

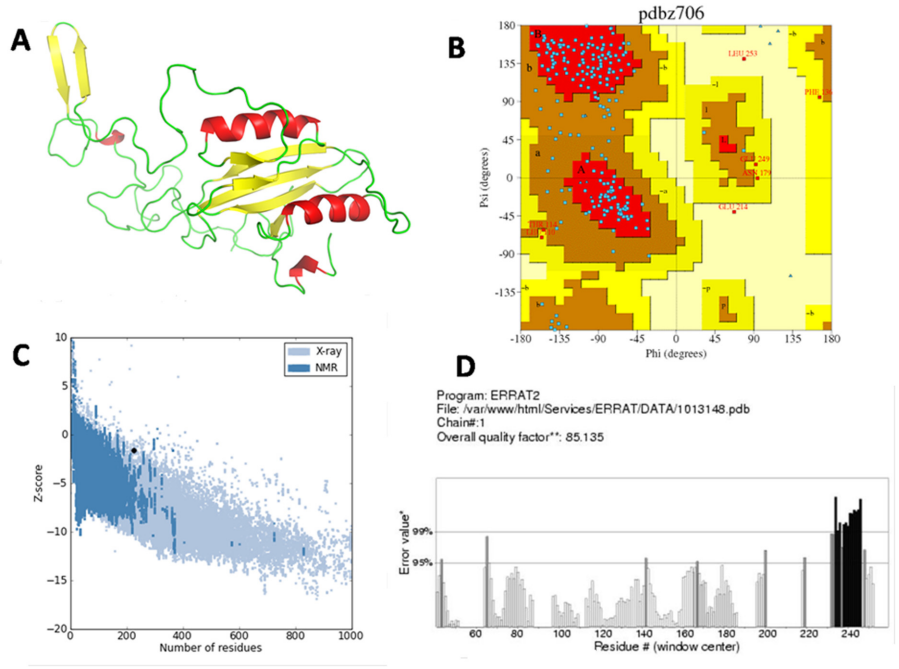


Figure 3.3.2: A: Molecular model of *NDUFS3* prepared using pdb structure 3IAM (E-chain) as template, B: Ramachandran plot, C: ProSa analysis, D: ERRAT score.

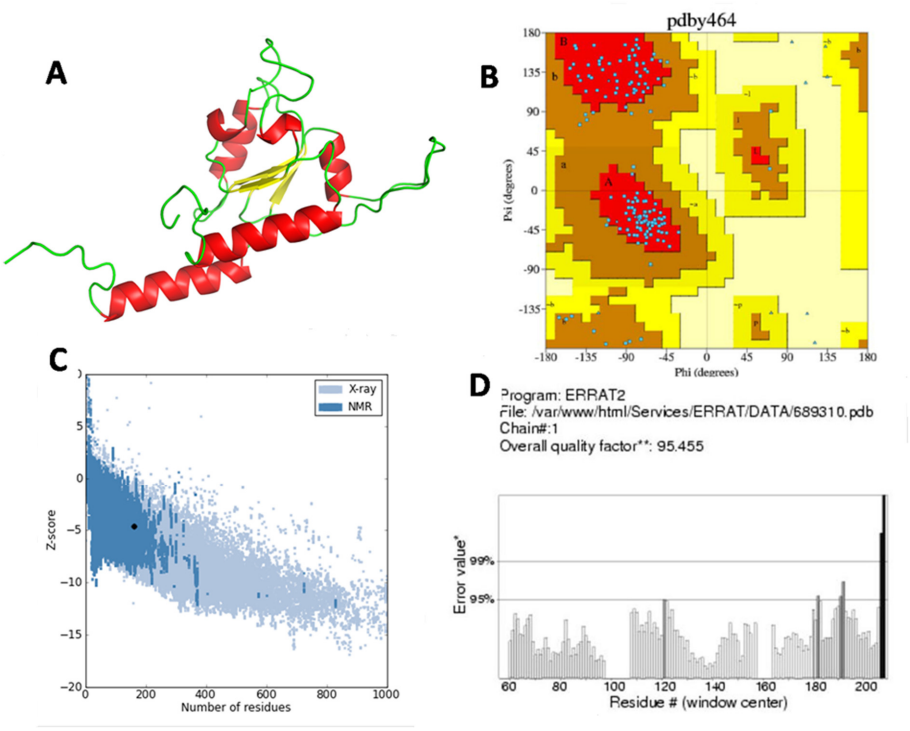


Figure 3.3.3: A: Molecular model of *NDUFS7* prepared using pdb structure 3I9V (F-chain) as template, B: Ramachandran plot, C: ProSa analysis, D: ERRAT score.

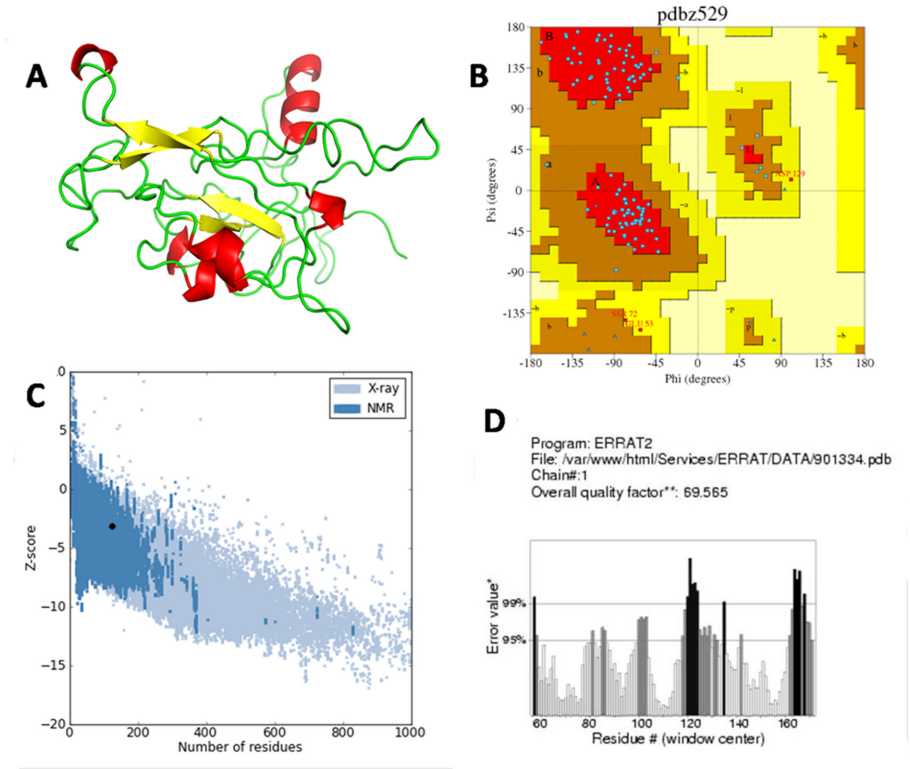


Figure 3.3.4: A: Molecular model of *NDUFS8* prepared using pdb structure 3IAM (F-chain) as template, B: Ramachandran plot, C: ProSa analysis, D: ERRAT score.

The mutant structures were generated using *in silico* mutagenesis for all the four subunits of Q module. The RMSD were insignificant for residues forming secondary structure in the post energy minimized models of these mutants.

3.1.4 Modeling of the human Q module

The subunit arrangement in the human Q module was studied by modeling the Q module using the template from *T. thermophilus* (3I9V). The conserved residues at the interface of four subunits in the Q module assembly modeled (Figure 3.4) were examined. Most interacting interface residues of the *T. thermophilus* were found to be conserved in the human mitochondrial Q module as well. The three iron-sulphur clusters (N6a, N6b and N2) modeled within the Q module remains the path for the entry of DBQ, indicating importance of the correct assembly in the formation of functional complex.

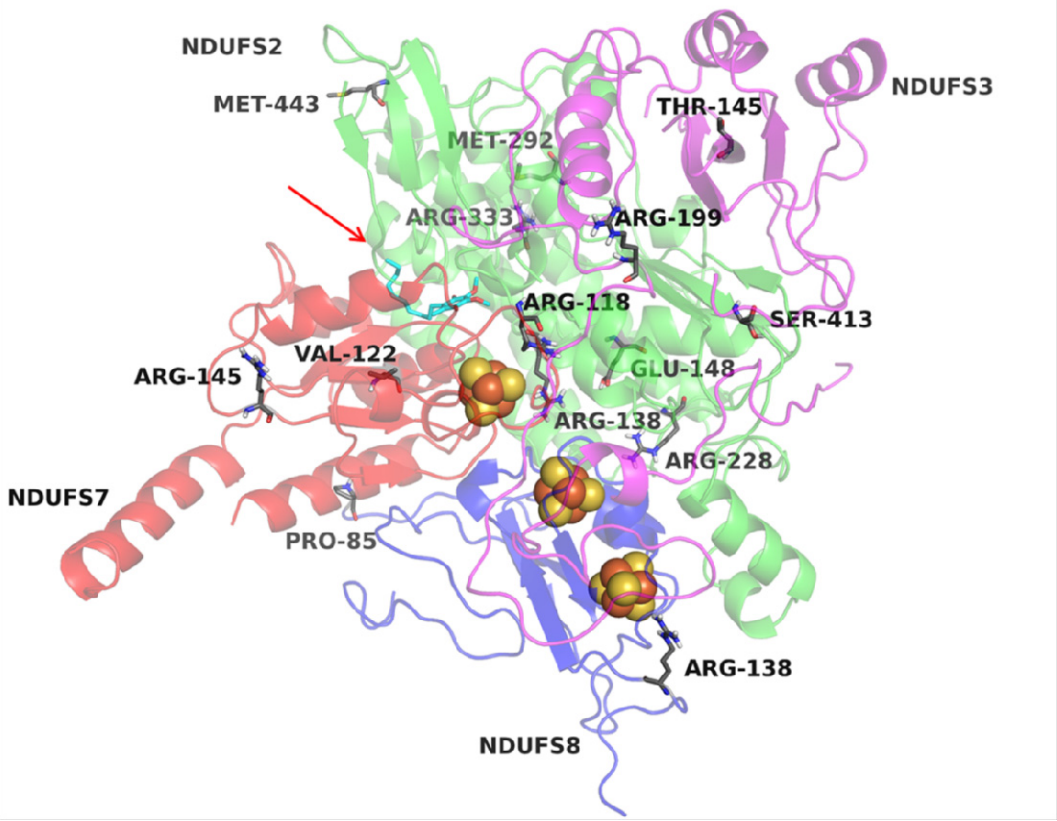


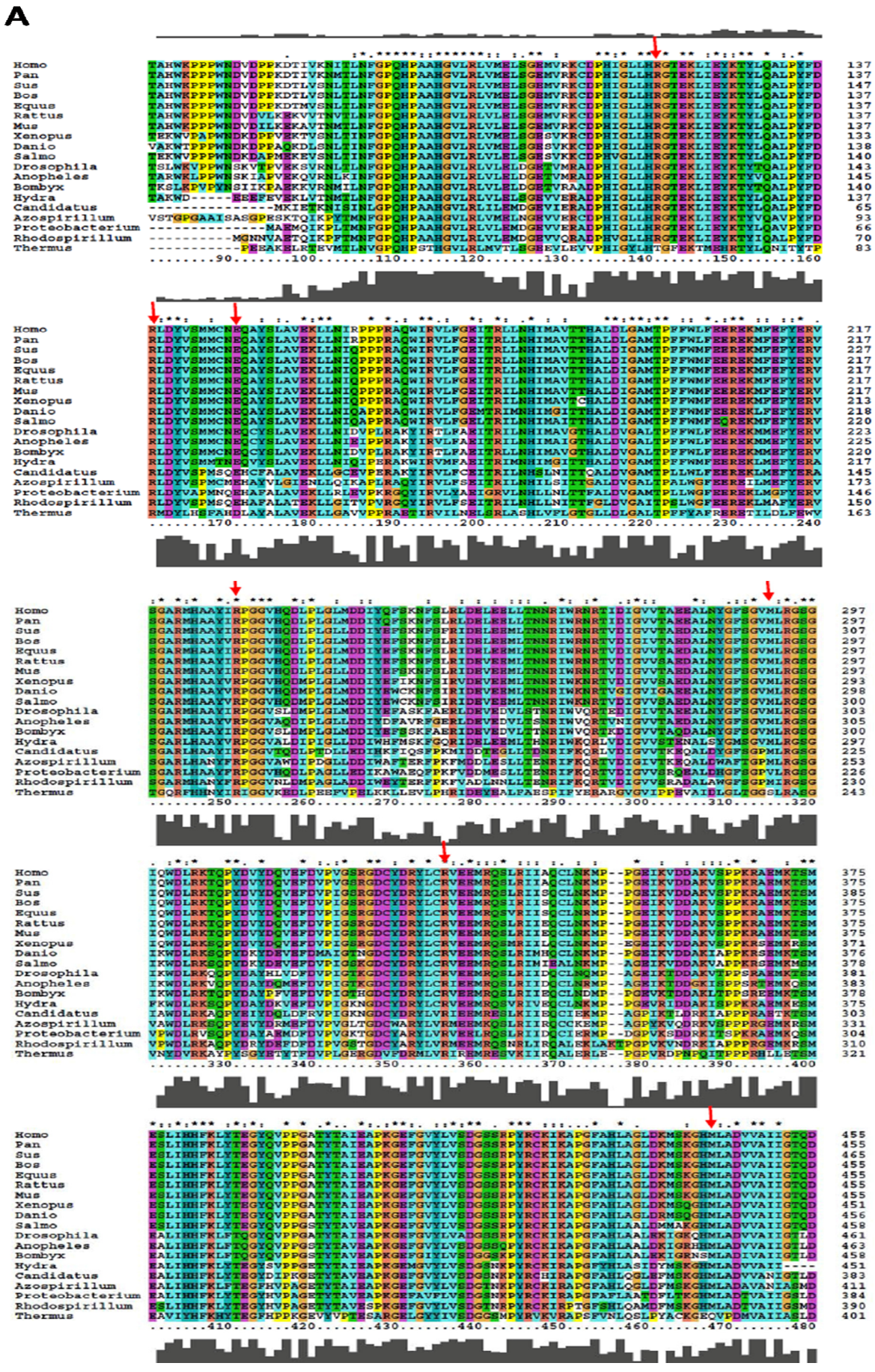
Figure 3.4: Subunit arrangement of the core subunits in the human Q module. Subunits *NDUFS2* (green), *NDUFS3* (magenta), *NDUFS7* (red) and *NDUFS8* (blue) are shown. The iron sulphur cluster N6a, N6b and N2 are shown as spacefill. The amino acid residues highlighted are the points of mutation considered. The red arrow indicates the entry point of n-DBQ (cyan).

The positions of the mutations were noted with respect to the disposition of other subunits, binding site and iron-sulphur clusters. On the basis of these observations the mutations were classified into those present on the surface, those in the interior and those at the interface or in the proximity of the ubiquinone binding site or near the iron sulphur clusters (Table 3.4)

Table 3.4: Classification of Leigh syndrome mutations based on their location and effects. Mutations in *NDUFS3* & *8* influence the structure whereas those in *NDUFS2* & *7* have functional implications.

Subunit	Mutation	Location of the amino acid
NDUFS2	R118Q	Interface of *NDUFS2* and *3*
Functional	R138Q	Interface of *NDUFS2*, *7* and 8
+Structural		Close to iron sulphur cluster N2
	E148K	Interior
	R228Q	Interface of *NDUFS2* and *8*
	M292T	Interior
	R333Q	Interior
	M443K	Surface
NDUFS3	T145I	Interior
Structural	R199W	Surface
NDUFS7	V122M	Close to iron sulphur cluster N2
Functional	R145H	Surface
NDUFS8	P85L	Interface of *NDUFS7* and *8*
Structural	R138H	Interface of *NDUFS3* and *8*

Interestingly, the amino acid residues at the point of mutations were highly conserved from bacteria to mammals (Figure 3.5)


A
Homo
Pan
Sus
Bos
Equus
Rattus
Mus
Xenopus
Danio
Salmo
Drosophila
Anopheles
Bombyx
Hydra
Candidatus
Azospirillum
Proteobacterium
Rhodospirillum
Thermus

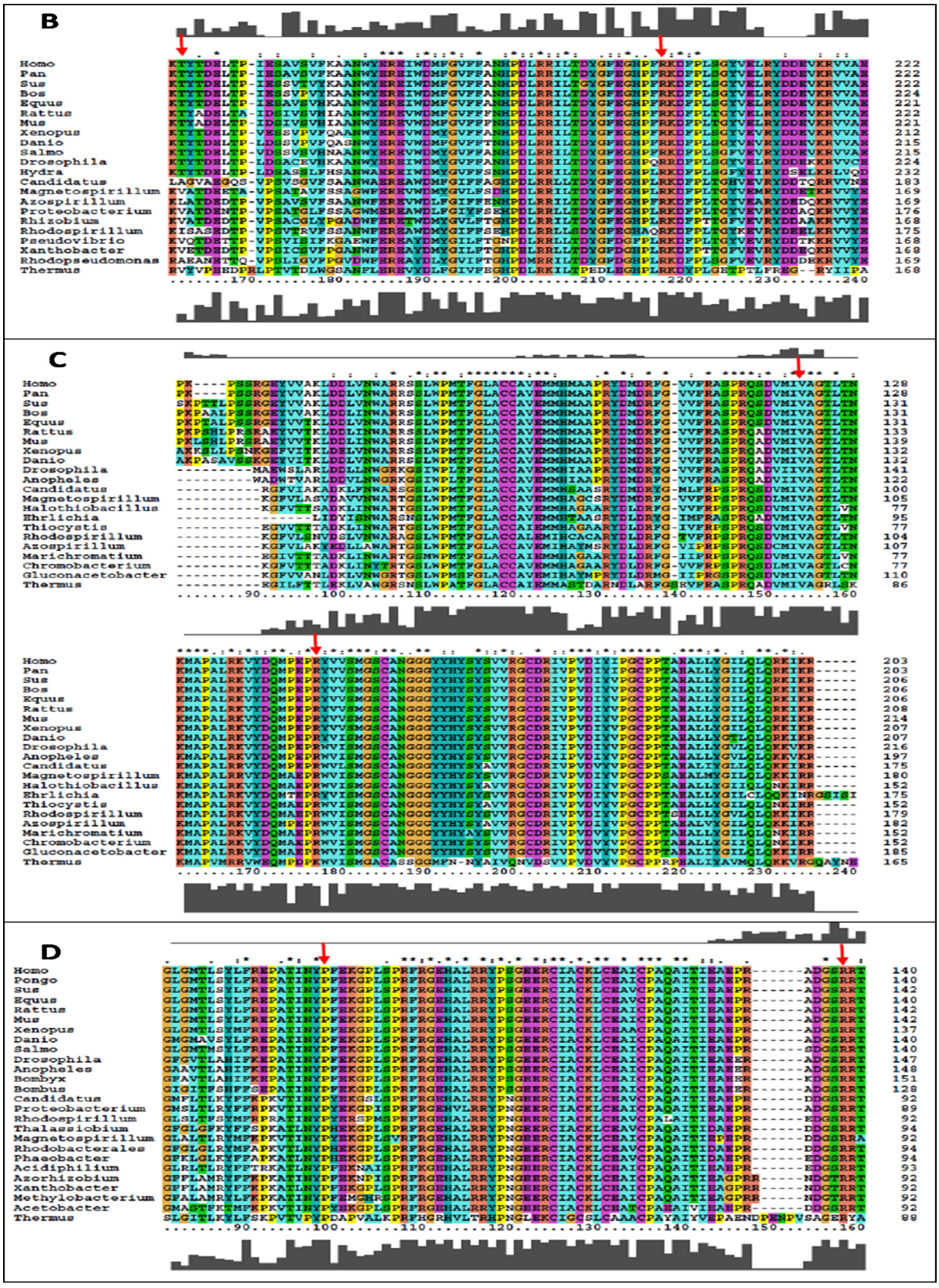


Figure 3.5:Multiple sequence alignment of A: NDUFS2, B: NDUFS3, C: NDUFS7 and D: NDUFS8 across species. The red arrow indicates the position of the mutations.

3.1.5 Docking of n-DBQ

Ten poses were generated in Glide5.8 for n-DBQ docking in Q module. The glide gscore and Emodel score (estimated conformational energy of the ligand) for the best pose was -6.08 and -31.07, respectively. The ligand interaction diagram for n-DBQ was plotted as a two-dimensional plot showing the interactions of the ligand bound to the protein (Figure 3.6B). The residue Y141 (equivalent to Y144 of yeast and Y87 of *T. thermophilus*) close to the iron-sulphur cluster N2 formed one hydrogen bond with the carbonyl head group of n-DBQ and the hydrophobic side chain interacted with the hydrophobic residues at the interface of *NDUFS2* and *NDUFS7* (Figure 3.6A and B).

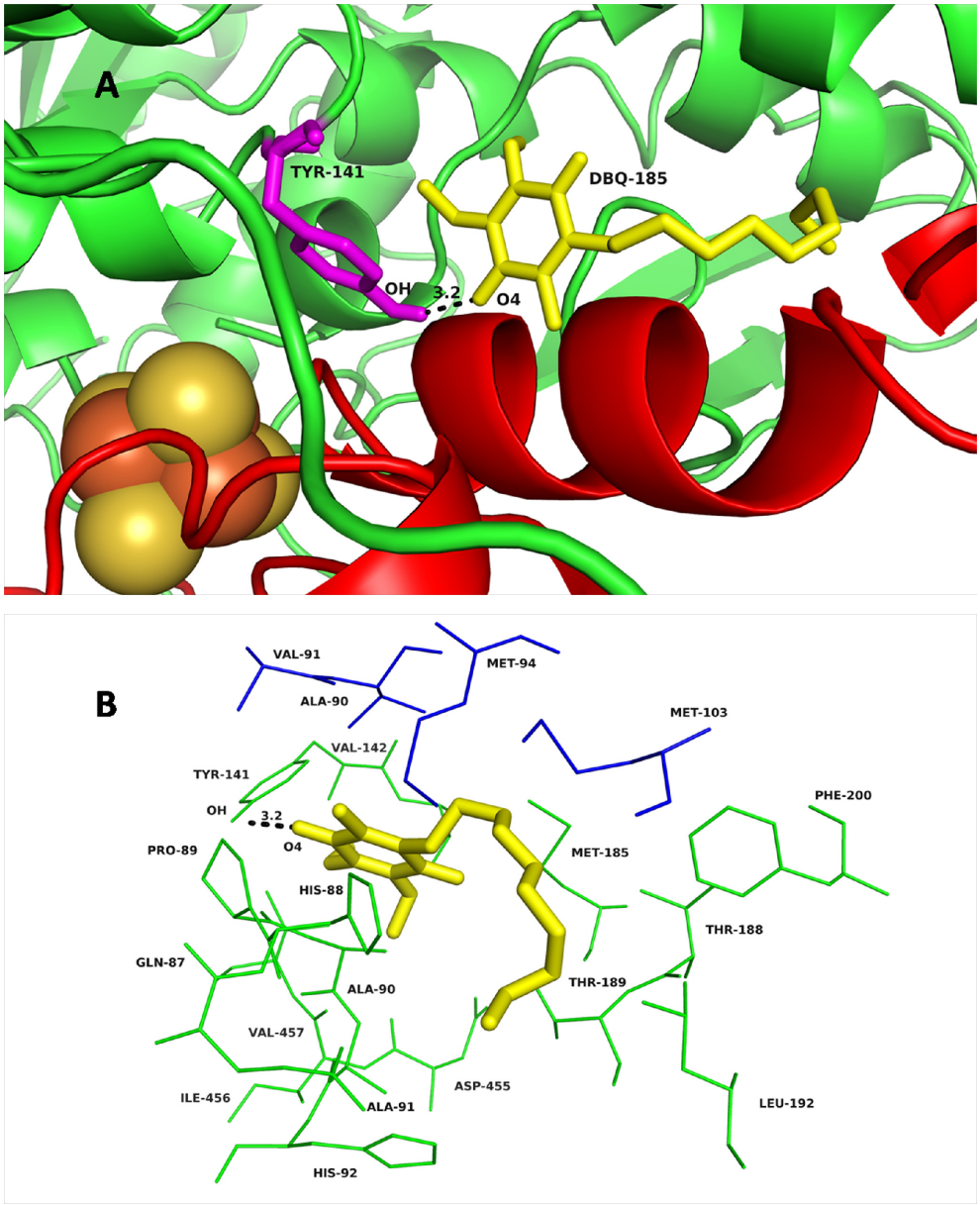


Figure 3.6: A: DBQ is docked at the interface of *NDUFS2* and *NDUFS7*. Iron-sulphur cluster N2 is in the close vicinity of residue Y141. The hydroxyl group of residue Y141 forms a hydrogen

bond with the carbonyl group atom O4 of DBQ. B: Ligand interaction diagram showing the interaction of DBQ (yellow) with the amino acid residues at the interface of *NDUFS2* (green) and *NDUFS7* (blue).

Amino acid residues T189, L192, F200 and V457 of *NDUFS2* and M94 and V91 of *NDUFS7* were in the proximity of n-DBQ. It is interesting to note that the importance of these residues in binding of n-DBQ has been shown previously in site-directed mutagenesis study in yeast [Angerer *et al*, 2012].

3.1.6 Energy calculations

The differences in energies between the w-t and mutant complexes of Q module were calculated in Bioluminate 1.0 residue scanning module. The contribution of individual mutations toward the energy and stability of the complex was estimated in terms of differences in energies of mutant and w-t protein structures (Table 3.5).

Table 3.5: Contribution from individual mutations to the total energy and stability of the complex along with the changes in affinity towards iron-sulphur clusters, DBQ and other subunits.

Subunit	Mutation	Δ Stability kcal/mol	Δ Affinity kcal/mol		
			Fe-S clusters	DBQ	Neighbouring subunits
NDUFS2	R118Q	-8.51	-17.86	0.28	6.73
	R138Q	17.38	-22.79	-0.17	21.09
	E148K	47.59	6.72	-0.04	-2.75
	R228Q	17.91	-7.20	0.02	25.27
	M292T	10.13	0.09	0.03	0.66
	R333Q	16.08	-1.68	0.12	-0.95
	M443K	4.63	10.46	-0.01	1.02
NDUFS3	T145I	7.17	0.00	-	0.02
	R199W	39.60	1.84	-	7.39
NDUFS7	V122M	5.25	-0.69	0.01	-0.56
	R145H	35.35	-2.20	0.01	-0.09
NDUFS8	P85L	26.40	3.53	-	-
	R138H	12.54	-15.91	-	-0.22

The Δ stability calculations measure the instability of mutant structures compared to w-t. All mutations except R118Q showed positive values of Δ stability indicating the role mutations play as structure destabilizing factor. The *NDUFS2* mutation S413P could not be tested because Bioluminate program had no provision to mutate from non-proline residue to proline. Estimate of the binding affinities of mutants towards iron-sulphur clusters, DBQ and neighbouring subunits indicated a reduction in affinity of certain mutants such as E148K of *NDUFS2* and P85L of *NDUFS8* towards iron-sulphur clusters. Similarly, a reduction in affinity towards n-DBQ was observed for most mutations of *NDUFS7* and mutations of *NDUFS2* except R138Q, E148K and M443K. Majority of the mutations were affecting the inter-subunit affinity and stability of the Complex (Table 3.5).

3.1.7 Molecular dynamic simulation

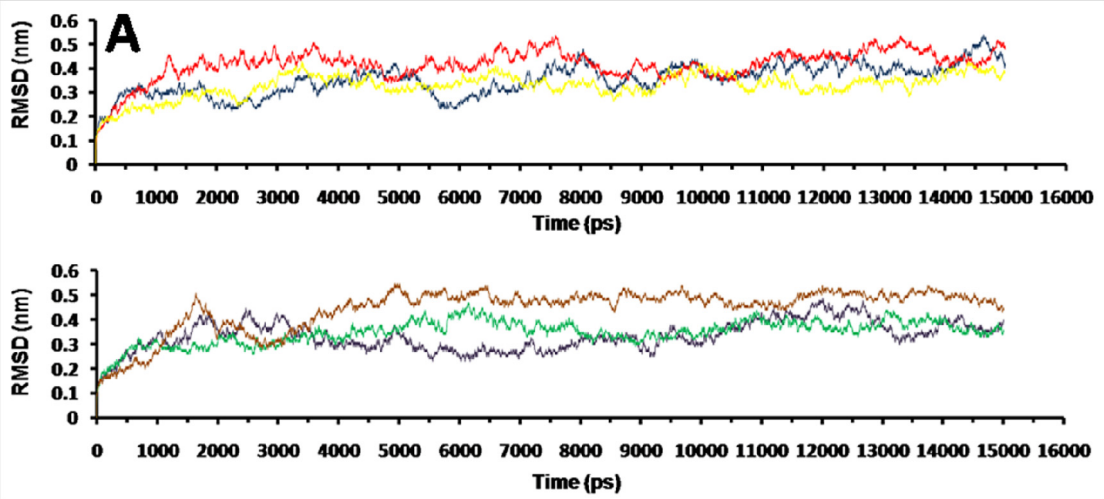
In order to highlight the structural changes in the subunits on mutation, molecular dynamic simulations were run as described in section 2.5. A 15 ns time scale trajectory was analyzed after suitable stabilization of the structures.

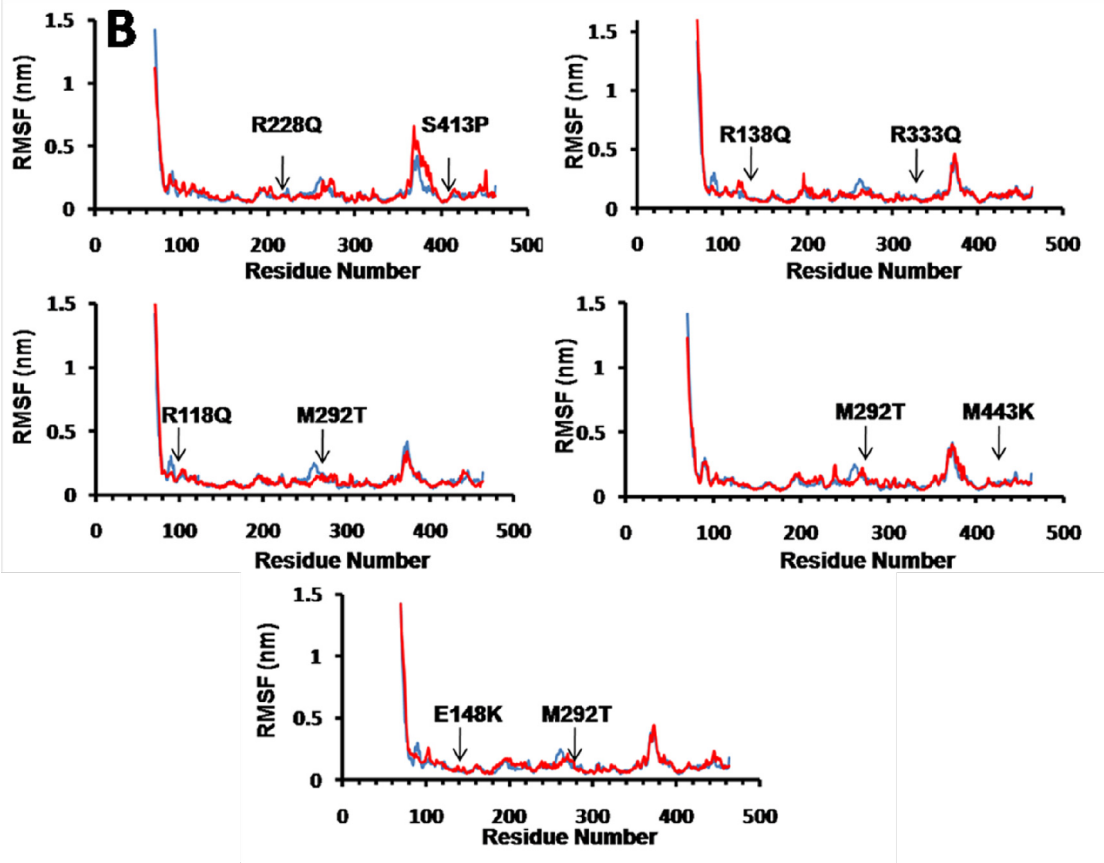
The RMSD corresponding to the Cα atom positions of each mutant with respect to the wild-type subunit were compared. Any changes in the hydrogen bonded interactions, on mutating each subunit followed by energy minimization, were identified. The changes in the solvent accessible area as well as average intra- and inter-molecular hydrogen bonds are listed in Table 3.6.

Table 3.6: Changes in solvent accessible area (nm/S2/N) and average number of hydrogen bonds within the four subunits when compared with their corresponding mutant structures.

	Solvent Accessible Area (nm/S^2/N)				Average hydrogen bonding			
	Hydrophobic	Hydrophilic	Total	D Gsolv	Intra-molecular		Inter-molecular or with water	
					H- bonds	Pairs within 0.35 nm	H- bonds	Pairs within 0.35 nm
NDUFS2								
Wild-Type	110.5	104.71	215.22	-540.28	257	1321	804	867
R228Q+S413P	112.13	107.99	220.12	-559.59	259	1307	811	881
R138Q+R333Q	107.96	104.17	212.14	-532.54	267	1321	784	843
M292T+R118Q	109.16	105.03	214.19	-537.71	258	1322	792	876
M292T+M443K	112.21	106.83	219.04	-549.87	260	1333	809	894
M292T+E148K	108.89	103.6	212.49	-533.44	270	1319	771	826
NDUFS3								
Wild-Type	95.91	77.44	173.36	-435.19	115	718	405	487
T145I+R199W	94.66	78.82	173.48	-435.50	108	703	415	495
NDUFS7								
Wild-Type	66.75	54.44	121.19	-304.23	75	529	383	488
V122M	65.67	53.10	118.77	-298.16	82	528	371	481
R145H	67.37	54.52	121.89	-305.99	75	531	390	481
NDUFS8								
Wild-Type	45.27	48.85	94.12	-236.27	61	390	358	372
P85L+R138H	44.00	48.08	92.07	-231.14	64	392	351	362

***NDUFS2*:** The root mean square fluctuations (RMSF) indicated minor changes in the positions of amino acid residues present at the mutation sites (Figure 3.7.1B). However, rest of the structure was almost stable. The overall compactness of the structure as indicated by the Radius of gyration (Rg) (Figure 3.7.1C) remains fairly constant at the end of the 15 ns simulation for the wild-type protein and mutants.





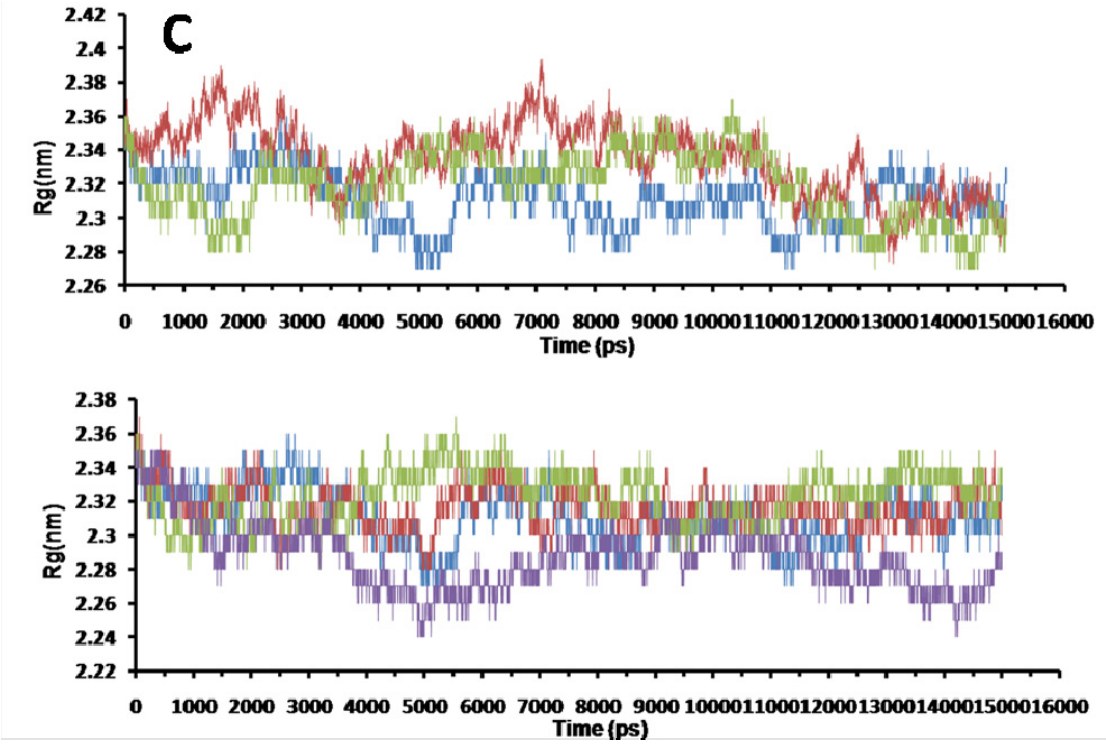


Figure 3.7.1: A: RMSD of the Cα atom positions of the w-t NDUFS2 (blue) and its mutants R228+S413P (red), R138+R333Q (yellow), M292T+ R118Q (brown), M292T+M443K (green) and M292T+ E148K (purple) over a 15 ns trajectory, B: RMSF of the amino acid residues of w-t (blue) and the mutants (red) at the end of the 15 ns trajectory, C: Radius of gyration (Rg) change of the w-t (blue), R228+S413P (green), R138+R333Q (brown), M292T+ R118Q (red), M292T+M443K (dark green) and M292T+ E148K (purple) over a 15 ns trajectory.

Of the five mutants studied, two mutants influenced the hydrogen bonding with neighbouring residues. In the mutant R228Q+S413P, the hydrogen bond formed between amino acid residues Arg228 and His223 was lost with R228Q mutation, similarly the hydrogen bonded interaction between Ser413 and Gly394 was lost by S413P mutation (Figure 3.7.2A). In mutant (R138Q+R333Q) the first mutation destroyed side chain ionic interaction between amino acids Arg138 and Asp137 (Figure 3.7.2B). The R333Q mutation disrupted the side chain interaction of Arg333 with Thr453. No change in hydrogen bonds were found in the remaining three mutants: (M292T+R118Q), (M292T+M443K), (M292T+E148K).

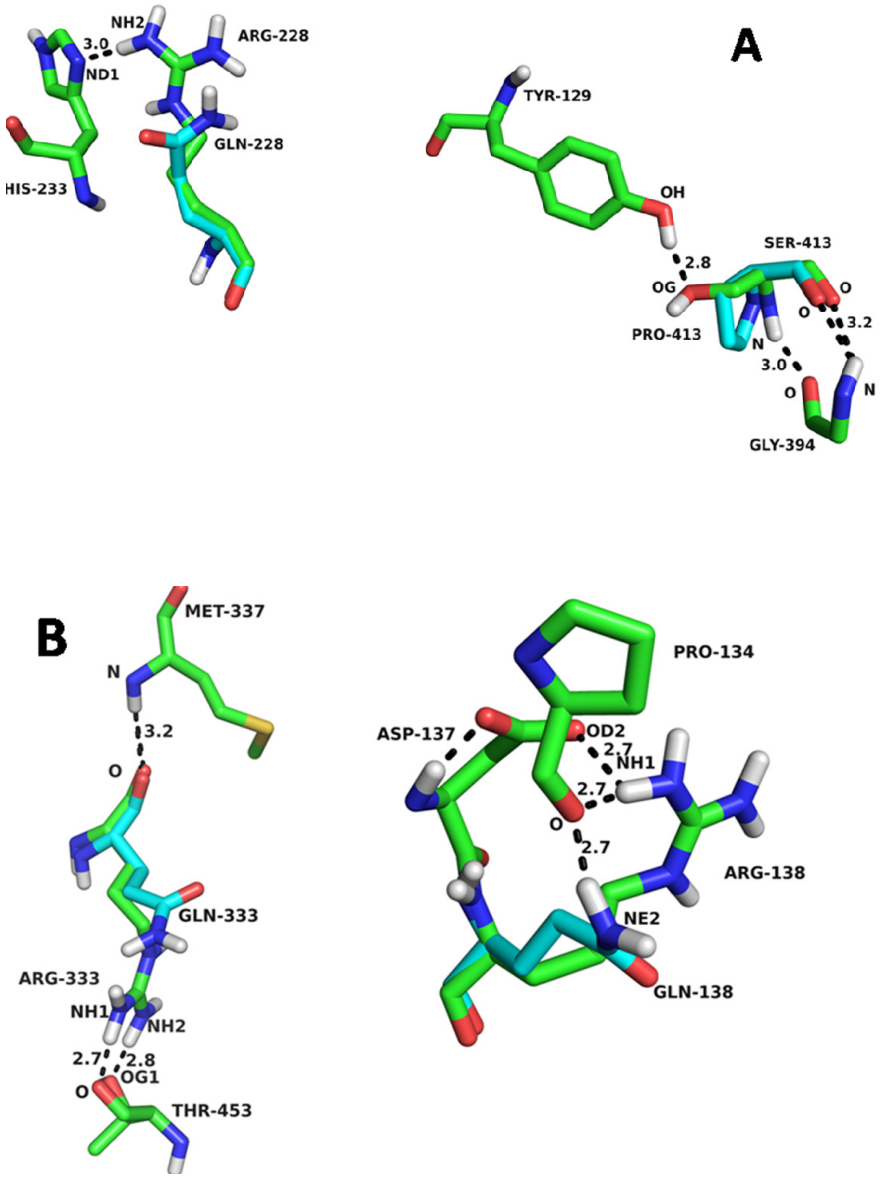
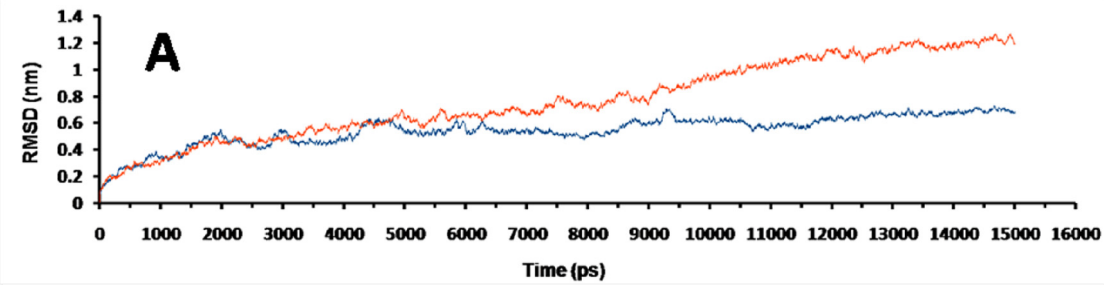


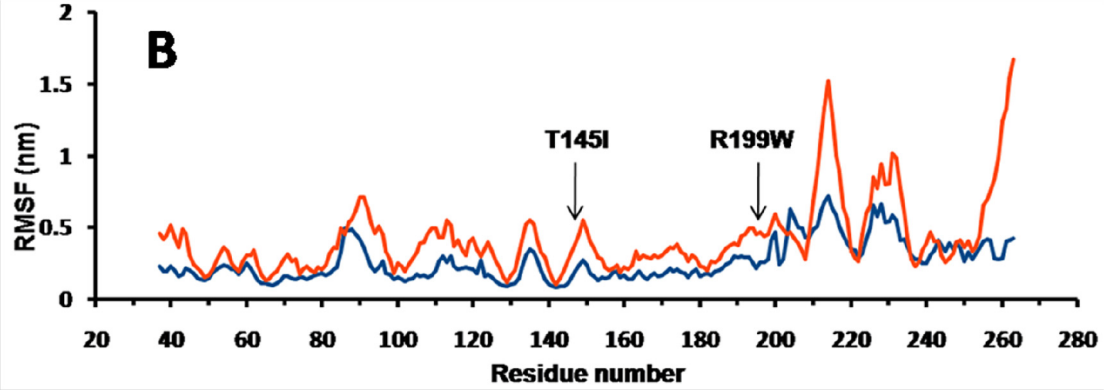
Figure 3.7.2: A: Hydrogen bond broken due to R228Q+S413P mutation is shown, B: Loss of hydrogen bond due to R138Q+R333Q mutation is shown by overlapping the amino acids involved in hydrogen bond in w-t and mutant.

The three combined mutations with the interior residue mutation M292T were reported in four unrelated individuals of Caucasian families suffering from Leigh syndrome [Tuppen *et al*, 2010]. This mutation did not affect hydrogen bond with neighbouring residue G290. However, it introduced an additional protein kinase C phosphorylation site at threonine, potentially affecting the regulatory function of the Q module [Pagniez-Mammeri *et al*, 2012]. The R118 amino acid residue is known to

be involved in a post-translational methylation (R85 in bovine). The methylated residue probably influences the assembly of the complex or the redox potential of the complex [Carroll *et al*, 2013]. Hence, though the mutation showed a better stability (negative Δ stability value); the post translation modification would not occur and thereby causing a negative impact on the functioning of the Complex-I. Thus, a combined mutation would have a deleterious effect. Two *NDUFS2* mutations, interior residue E148K and surface residue M443K mutations, although did not influence the hydrogen bonds, affected the affinity for the iron-sulphur clusters (Table 3.5).

***NDUFS3*:** The values of RMSF indicated that the mutant (T145I+R199W) structure had a higher number of fluctuating residues as compared to the wild-type (w-t) protein (Figure 3.8 B).





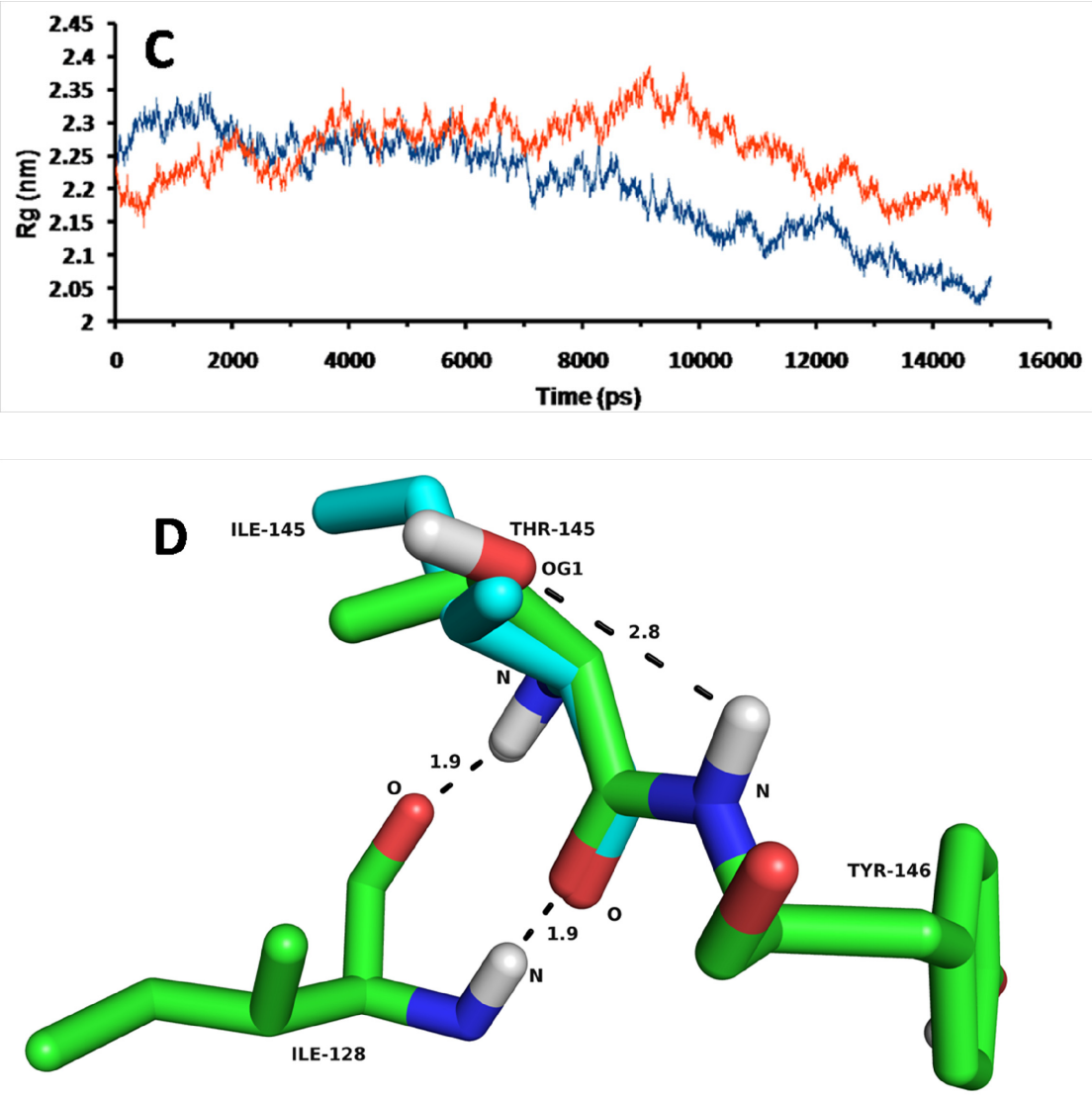
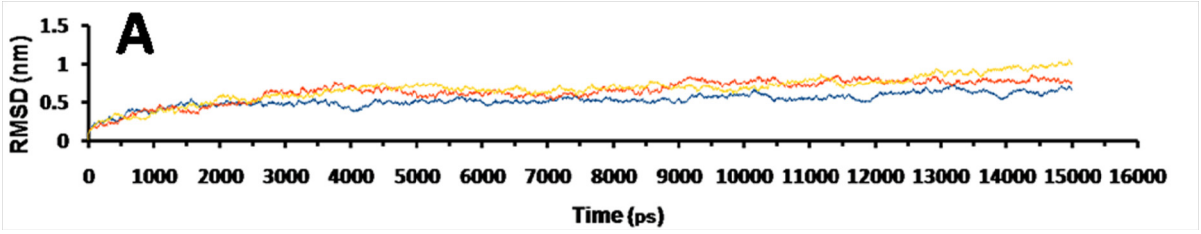


Figure 3.8: A: RMSD of the Cα atom positions of the w-t NDUFS3 (blue) and its mutant (T145I+R199W), B: RMSF of the amino acid residues of the w-t (blue) and its mutant (red), C: Rg of the w-t (blue) and its mutant (red), D: The loss of hydrogen bond due to T145I mutation is shown.

This may be accounted for by the large number of loops present in *NDUFS3* structure. At the end of the simulation, w-t protein showed a more compact structure compared to the mutant as measured by the radius of gyration (Rg) (Figure 3.8 C). The hydrogen bond between amino acid residue Thr145 and Tyr146 was lost by the mutation T145I however no changes were observed for R199W mutation (Figure 3.8 D).

The T145I interior residue mutation in *NDUFS3* was at a putative casein II phosphorylation site [Pagniez-Mammeri *et al*, 2012] and caused loss of a critical hydrogen bond thereby decreasing the protein stability. The surface mutation R199W occurring in combination with interior residue mutation T145I decreased the affinity of the *NDUFS3* for other subunits thereby decreasing the overall stability of the Q module. The interesting observation is that the patient inherited individual mutations from each healthy parent thereby showing that the combination only caused disease [Benit *et al*, 2004]. Analysis of the post simulation structures indicated that the T145I mutation resulted in a shift in the secondary structure. When it occurred as a combination with the other mutation R199W, it resulted in the shifting of the loop containing amino acid residues 196-206 towards the *NDUFS2* subunit causing a major deviation of the loops and β−sheets along the structure. Severe steric clashes with subunits *NDUFS2* and *NDUFS8* of assembly were also observed. Thus, when T145I mutation occurs together with R199W, the *NDUFS3* subunit becomes highly unstable, evidenced by the increased fluctuations of residues and decreased compactness. The loops, by turning highly dynamic, affect the interaction between *NDUFS3* and *NDUFS2* subunits preventing the initiation of assembly process.

***NDUFS7*:** Both the mutant structures (V122M and R145H) show difference in fluctuation of residues compared to w-t protein throughout the structure (Figure 3.9 B). The w-t protein and mutants have similar compactness as measured by Rg (Figure 3.9 C). V122M mutation did not influence hydrogen bonding, whereas the R145H mutation disrupted hydrogen bonds of Arg residue with Met141 and Asp139 (Figure 3.9 D).



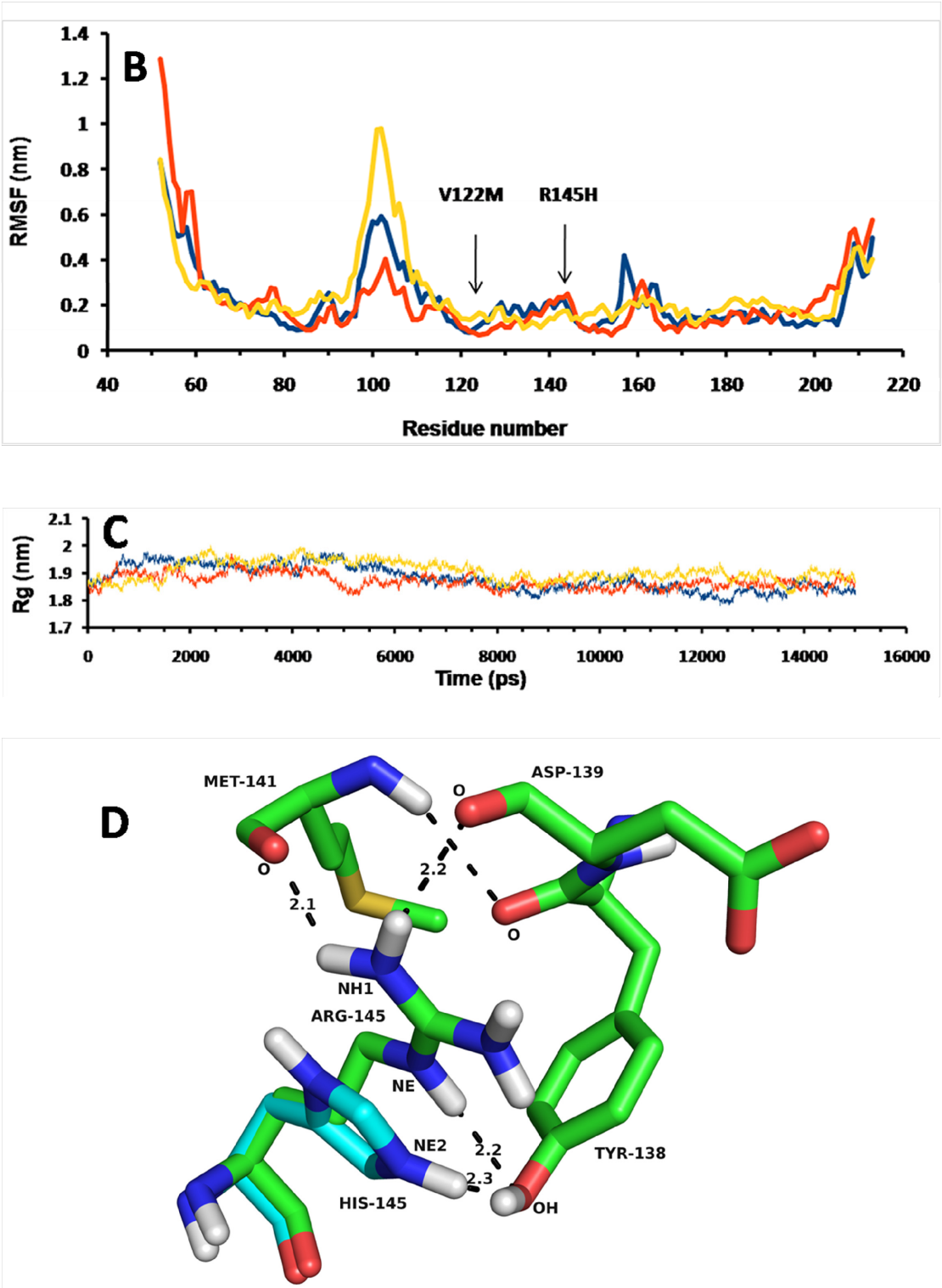
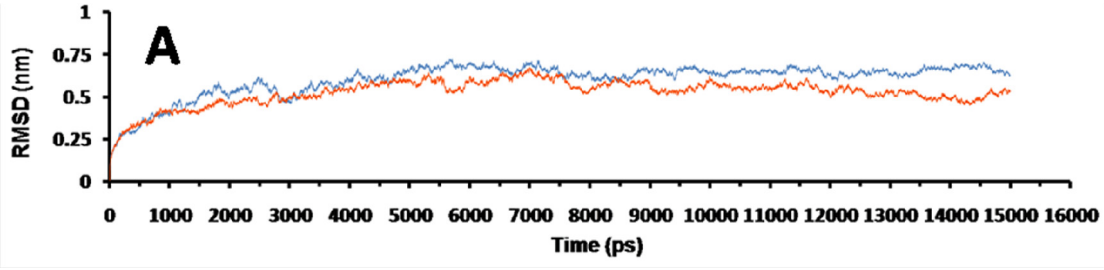


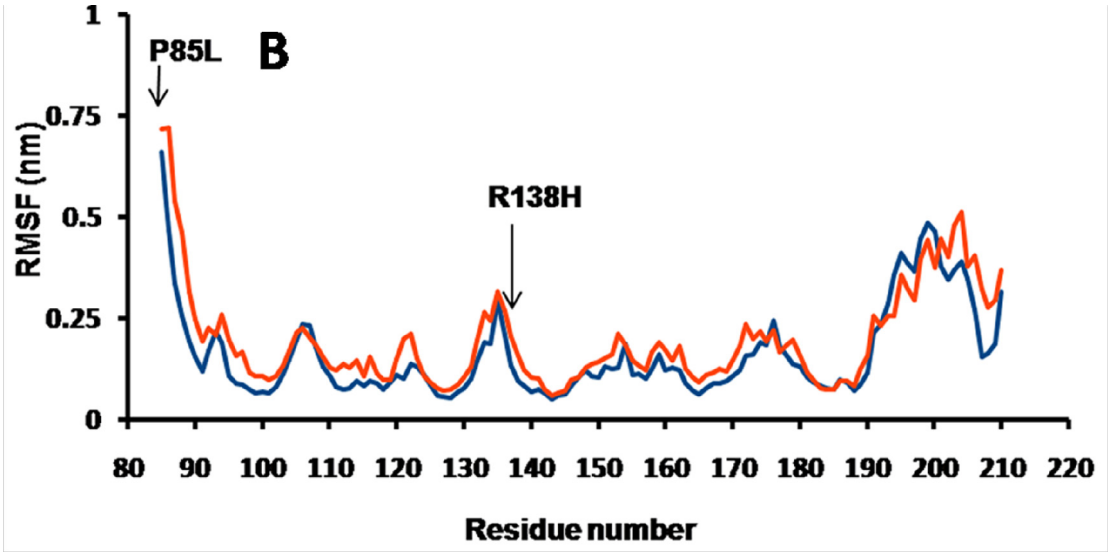
Figure 3.9: A: RMSD of the Cα atom positions of the w-t NDUFS7 (blue) and V122M (red), R145H (yellow), B: RMSF of the amino acid residues of the w-t (blue) and mutants V122M (red) and R145H (yellow), C: Rg of the w-t (blue) and V122M (red) and R145H (yellow) D: The loss of hydrogen bond due to R145H mutation is shown.

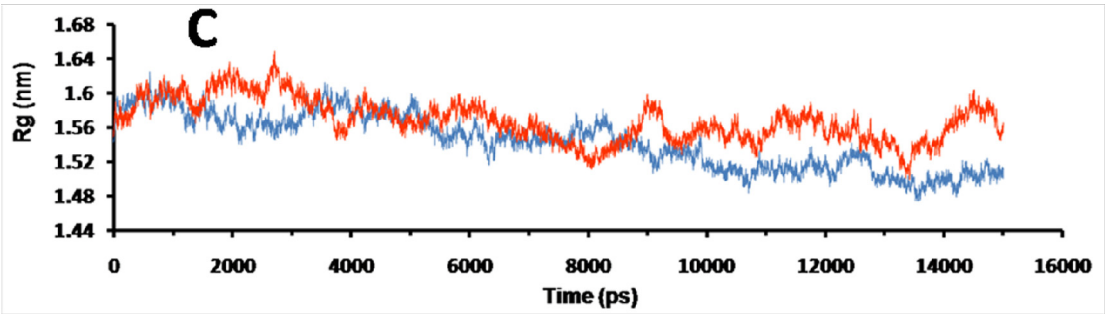
The mutation V122M in the functionally important subunit *NDUFS7* had no effect on any hydrogen bond. The mutated residue lies within a 4A° radius of the iron-sulphur cluster N2, located within the cysteine motif -CxxE-$(X)_{60}$-C-$(X)_{30}$-CP- and involved in the formation of N2 cluster [Triepels *et al*, 1999]. Irrespective of its proximity the mutation did not change affinity for the iron-sulphur clusters. Surprisingly the mutation slightly decreased the affinity for n-DBQ (Table 3.6), although not lying in the immediate vicinity of DBQ binding site, indicating the importance of this residue for the function of subunit. The surface mutation R145H not only caused loss of a hydrogen bond but also reduced the affinity towards n-DBQ. Thus, the point mutations in *NDUFS7* severely affect the activity of Complex-I by reducing affinity towards DBQ.

***NDUFS8*:** The residue-fluctuations were more in mutant (P85L+R138H) compared to the w-t protein across the structure (Figure 3.10 B). At the end of simulation the values of Rg pointed to a more compact w-t structure than mutant (Figure 3.10 C). The hydrogen bond between residues Arg138 and Glu131 was lost as a result of R138H mutation (Figure 3.10 D).

The interface mutation P85L of *NDUFS8* reduced the affinity of the module towards iron-sulphur clusters. Similarly, the other surface mutation R138H resulted in removal of hydrogen bond with residue E131 thereby decreasing the stability of subunit. The patient bearing these two mutations was homozygous for both mutations, which had been inherited one each from healthy heterozygous parents, indicating lethality of their combination to cause disease [Loeffen *et al*, 1998].







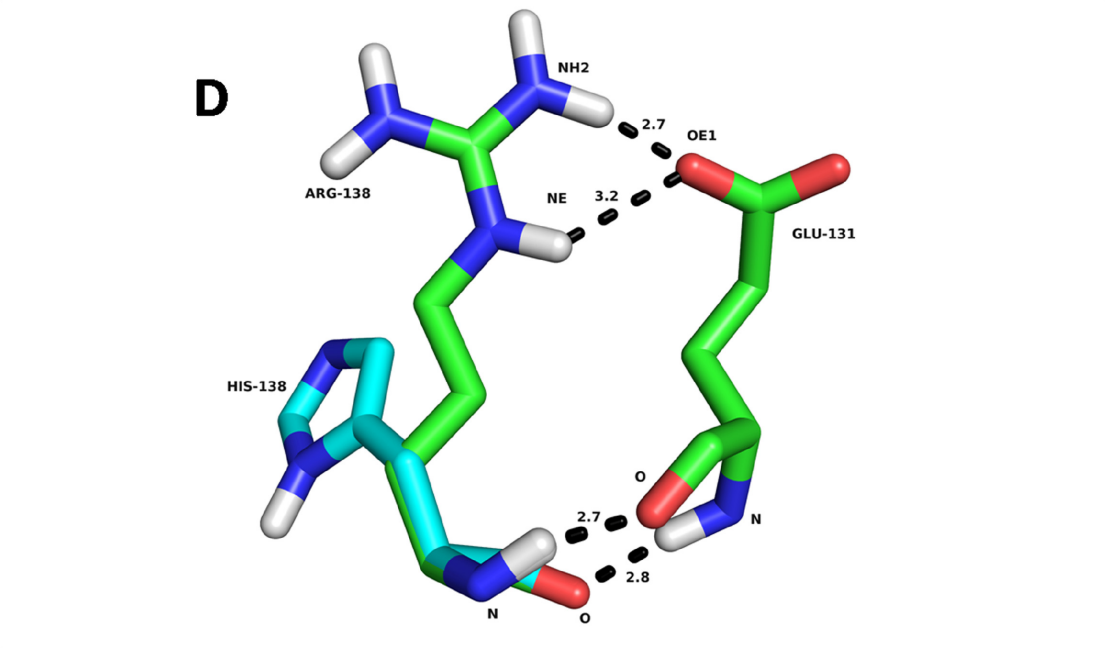


Figure 3.10: A: RMSD of the Cα atom positions of the w-t NDUFS8 (blue) and its mutant P85L+R138H (red), B: RMSF of the amino acid residues of the w-t (blue) and its mutants

P85L+R138H (red), C: Rg of the w-t (blue) and its mutant P85L+R138H, D: The loss of hydrogen bond due to R138H mutation is shown.

3.1.7 Overall effect of mutations on structure

Several structural and energetic changes were observed in the study revealing the unstable and deleterious effect of the mutations on the overall structure of the complex.

3.1.7.1 Fluctuation of residues and compactness

NDUFS2 and *NDUFS7* act as functional subunits of the Q module whereas the role of *NDUFS3* and *NDUFS8* are more of a structural nature. Hence, during the MD simulation drastic structural changes were observed in the *NDUFS3* and *8* whereas *NDUFS2* and *7* showed instability contributing from energy changes associated with function. The structures of *NDUFS3* and *NDUFS8* displayed an overall increase in fluctuation of residues in their mutant structures, probably owing to the large number of loops present in these structures. Also the structures of the w-t protein of both subunits remained more compact compared to the corresponding mutants. In *NDUFS7* the residue fluctuations at the points of mutation were accompanied with large fluctuations in the loop regions whereas *NDUFS2* had fluctuations only at the points of mutations. Thus, the influence of the mutations on the flexibility of polypeptide chain was confined to the points of mutations in *NDUFS2*, but extended to neighbourhood in others.

3.1.7.2 Effect on solvent accessibility, hydrogen bonding and energetic instability

The change in the intra- and inter-molecular hydrogen bonds of w-t and mutants during the 15 ns simulation and their average values at the end of the trajectory are listed in Table 3.6. In each case different changes were observed in either the number of intra- or inter-molecular hydrogen bonds formed by w-t compared to mutants. The changes in the number and nature of hydrogen bonds could be considered an indication of the changes to atomic positions in w-t and mutant structures taking place through the 15 ns simulation. On examining the time evolution of structures during the 15 ns trajectory no major changes in the secondary structure between mutants and w-t could be detected. In all the mutants of *NDUFS2* and *NDUFS8* a distinct increase

or decrease in the solvent accessible area of mutant compared to w-t was seen. Although there was no change in solvent accessible area of mutant compared to w-t of *NDUFS3*, w-t was more compact than mutant. Similarly, in the case of *NDUFS7*, the R145H mutant did not show any change in the solvent accessibility.

The residues R118, R138 and R228 of *NDUFS2* and P85 and R138 of *NDUFS8* were present at the interface of the 4 subunits (Table 3.4). In accordance with their positions the mutations had a deleterious effect on the inter-subunit affinity as shown by the energy calculations in Bioluminate 1.0 except for *NDUFS8* mutants (Table 3.6). The positions of the mutations with respect to the iron-sulphur clusters and docked DBQ indicated that amino acid residues R118, R138 and R228 of *NDUFS2*, V122 of *NDUFS7* and R138 of *NDUFS8* were fairly close to ligands (Figure 3.4). The energy calculations in Bioluminate 1.0 reflect destabilizing effect of these mutations.

3.1.8 Conclusion

Large changes in RMSF, Rg and observed differences in the hydrogen bond formation showed the lethal effect of these point mutations on the structure of the individual subunits. Thus, the mutant molecule, with initial secondary structure similar to wild-type protein, tends to be distinctly different during MD simulation. The effect of these point mutations on the assembly and function of the Q module studied gave further insight on the consequences of these mutations on the structure assembly as a whole. However, it would be interesting to note that some of these mutations, especially in those residues which are structurally interior, need to occur in combination to show up as disease phenotype, confirming that their cumulative effect that is causing disease condition. The stability values also show that most of these mutants are energetically highly unstable. In *NDUFS3* and *NDUFS8* some single mutations alone were not lethal as they were present in healthy parents, also reflected in the Δ affinity values that were not significant (Table 3.5). Hence, their effects are more pronounced only when occur in combination.

In silico structural and functional analysis of the Q module showed that the mutations affected the system structurally by the mutated region tending to become more flexible, changing the compactness of the structure and resulting in loss of some hydrogen bonded interactions accompanied by an increase in energy; the mutations

functionally affect the system by reducing affinities towards the iron-sulphur clusters, DBQ and with the other core subunits. The accumulation of the assembly intermediates observed in patients having such mutations is the net effect of all the factors described here with respect to the structure and function of the Q module.

CHAPTER 4

Cloning, expression, purification and preparation of site-directed mutants of *NDUFS3* and *NDUFS7* subunits of human mitochondrial Complex-I Q module

NDUFS2, 3, 7 and *8* form the core subunits of the ubiquinone (Q) reduction module. They form the minimal enzyme assembly required for the smooth functioning of the Q module. These subunits are encoded by the nuclear genome and are transferred to the mitochondria. All of them possess fairly long signal sequences, followed by additional amino acid sequence which binds to the accessory subunits, and next to that is a highly conserved C-terminal domain. These subunits not only possess the binding site for ubiquinone but also for Complex-I inhibitors like rotenone, piericidin A etc [Darrouzet *et al*, 1998].

During the assembly of Complex-I, these subunits form the initial soluble peripheral arm which binds to the mitochondrially encoded ND1 subunit in the mitochondrial membrane [Mimaki *et al*, 2012]. Thus, these subunits are not only functionally important but are also playing an important role in the biogenesis and assembly of Complex-I.

In the present study, cloning and expression of the four core subunits of human mitochondrial Q module was tried. However, due to problems of expression and solubilisation of *NDUFS2, 8* and mitochondrially expressed ND1 (Appendix), purification and further characterization of only *NDUFS3* and *7* was carried out.

4.1 *NDUFS3*

The human *NDUFS3* gene (Entrez gene 4722) is present on chromosome 11p11.11 [Emahazion *et al*, 1998]. This gene consists of seven exons ranging in size from 66 to 248 bp. Intron sizes vary from 110 and 1845 bp. CCAAT and TATA boxes are absent in the sequence in the promoter region however, five sites for transcription factor Sp1 are present. Sites for transcriptional factors Ets, nuclear respiratory transcription factor NRF-2, transcriptional activator/repressor YY1 are also present [Procaccio *et al,* 2000]. This gene encodes for the protein NADH dehydrogenase (ubiquinone) iron sulphur protein 3. It is composed of 264 amino acids and has a molecular weight of ~30 kDa. The first 38 amino acid residues code for the mitochondrial targeting sequence. It shares 90% identity with the bovine protein. The C-terminal part is highly conserved from bacteria to mammals.

4.1.1 Cloning of full length human *NDUFS3* gene

4.1.1.1 RNA extraction, cDNA preparation and amplification of *NDUFS3* gene

RNA was isolated from *HT 29* cell line at a concentration of 415 ng μl^{-1} (Figure 4.1A). The cDNA library prepared was utilized to amplify the *NDUFS3* gene of size 808 bp with the primers described in materials and methods chapter (Figure 4.1B).

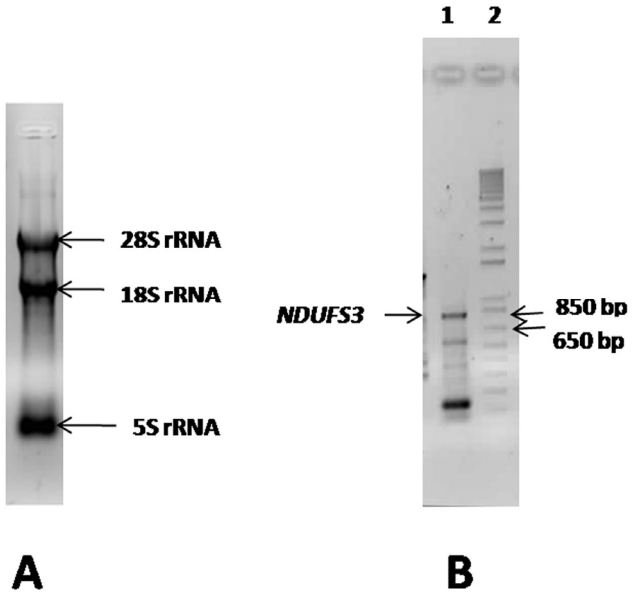


Figure 4.1: Image of A: Denaturing agarose gel electrophoresis showing the *HT 29* RNA used for preparing cDNA library and amplification of w-t *NDUFS3*. B: 1% Agarose gel electrophoresis, lane 1: amplified full length *NDUFS3* (808bp) gene, lane 2: 1 kb plus DNA ladder.

4.1.1.2 Cloning in pGEM-T vector

The A tailed *NDUFS3* PCR product was ligated with the pGEM-T vector. The ligation mixture was transformed in *E coli* DH5α cells. The colonies obtained on incubation were subjected to colony PCR and the plasmids isolated from them were sequenced to confirm the *NDUFS7* clones.

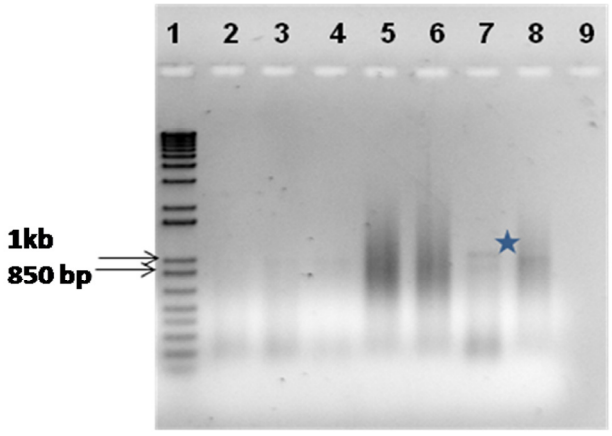


Figure 4.2: Image of 1% agarose gel electrophoresis. Lane 1: 1kb plus DNA Ladder Lane 2-9: Colony PCR products for colony number 1-8 for *NDUFS3* gene cloned in pGEMT vector and transformed in *E. coli* DH5α cells. Star indicates the presence of amplified *NDUFS3* Colony PCR product.

Colonies in lane number 7 and 8 showed amplified band of size between 850 bp -1 kb (Figure 4.2). The expected band for *NDUFS3* is 855 bp. The plasmid from colony number 7 was isolated, sequenced and further used for subcloning in pET-28b(+) expression vector.

4.1.1.3 Cloning in pET-28b(+) vector

The plasmid from colony number 7 was confirmed by sequencing. It was digested with restriction enzymes, NdeI and XhoI. The *NDUFS3* gene was thus, sub-cloned in pET-28b(+) vector between NdeI and XhoI sites. The ligation mixture was transformed and the colonies obtained on incubation were subjected to colony PCR.

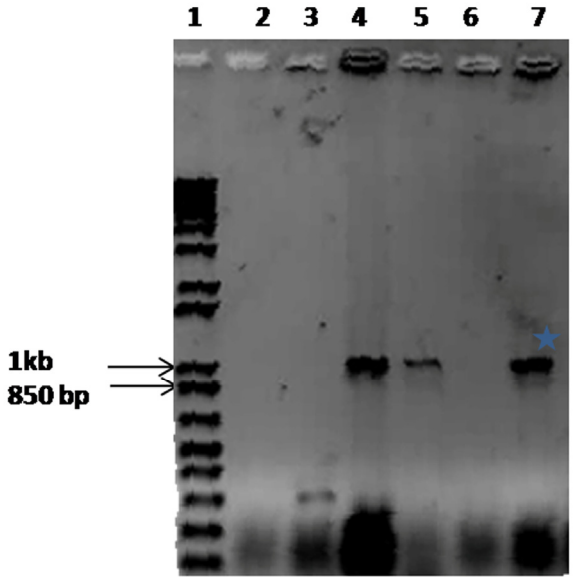


Figure 4.3: Image of 1% agarose gel electrophoresis. Lane 1: 1 kb plus DNA ladder, Lane 2-8: Colony PCR products for colony number 1-7 for NDUFS3 gene cloned in pET-28b(+) vector and transformed in *E. coli* DH5α cells. Star indicates the presence of *NDUFS3* colony PCR product.

Colonies in lane number 4, 5 and 7 showed bands of appropriate sizes for *NDUFS3* (Figure 4.3). The plasmids from these respective colonies were isolated and subjected to sequencing to confirm the positive clones. Blastp was run for the sequences obtained from plasmids isolated from colonies with a positive colony PCR. The plasmids obtained from all the 3 colonies showed 100% sequence identity with the human NADH dehydrogenase (ubiquinone) Fe-S protein 3(Figure 4.4). These colonies were maintained by preparing 15% glycerol stocks stored at -80°C.

```
>gb|AAD40386.1|  NADH-Ubiquinone reductase [Homo sapiens]
Length=263

 GENE ID: 4722 NDUFS3 | NADH dehydrogenase (ubiquinone) Fe-S protein 3, 30kDa
(NADH-coenzyme Q reductase) [Homo sapiens] (Over 10 PubMed links)

 Score =  526 bits (1355),  Expect = 2e-147
 Identities = 257/257 (100%), Positives = 257/257 (100%), Gaps = 0/257 (0%)
 Frame = +3

Query  225  RLWWRGILGASALTRGTGRPSVLLLPVRRESAGADTRPTVRPRNDVAHKQLSAFGEYVAE  404
            RLWWRGILGASALTRGTGRPSVLLLPVRRESAGADTRPTVRPRNDVAHKQLSAFGEYVAE
Sbjct  7    RLWWRGILGASALTRGTGRPSVLLLPVRRESAGADTRPTVRPRNDVAHKQLSAFGEYVAE  66

Query  405  ILPKYVQQVQVSCFNELEVCIHPDGVIPVLTFLRDHTNAQFKSLVDLTAVDVPTRQNRFE  584
            ILPKYVQQVQVSCFNELEVCIHPDGVIPVLTFLRDHTNAQFKSLVDLTAVDVPTRQNRFE
Sbjct  67   ILPKYVQQVQVSCFNELEVCIHPDGVIPVLTFLRDHTNAQFKSLVDLTAVDVPTRQNRFE  126

Query  585  IVYNLLSLRFNSRIRVKTYTDELTPIESAVSVFKAANWYEREIWDMFGVFFANHPDLRRI  764
            IVYNLLSLRFNSRIRVKTYTDELTPIESAVSVFKAANWYEREIWDMFGVFFANHPDLRRI
Sbjct  127  IVYNLLSLRFNSRIRVKTYTDELTPIESAVSVFKAANWYEREIWDMFGVFFANHPDLRRI  186

Query  765  LTDYGFEGHPFRKDFPLSGYVELRYDDEVKRVVAEPVELAQEFRKFDLNSPWEAFPVYRQ  944
            LTDYGFEGHPFRKDFPLSGYVELRYDDEVKRVVAEPVELAQEFRKFDLNSPWEAFPVYRQ
Sbjct  187  LTDYGFEGHPFRKDFPLSGYVELRYDDEVKRVVAEPVELAQEFRKFDLNSPWEAFPVYRQ  246

Query  945  PPESLKLEAGDKKPDAK  995
            PPESLKLEAGDKKPDAK
Sbjct  247  PPESLKLEAGDKKPDAK  263
```

Figure 4.4: Blastp results to show the matching of cloned sequence with that of human *NDUFS3*.

4.1.2 Reported mutations

In *NDUFS3;* a compound heterozygous mutation, C434T and C595T resulting in a transition in exon 5 and 6, respectively, is reported to cause Leigh syndrome [Benit *et al*, 2004]. These mutations result in the transition of a conserved Thr145 to Ile and Arg199 to Trp (Figure 4.5). A single homozygous mutation resulting in a transition of Arg199 to Trp has also been identified to cause Complex-I deficiency. The patient is reported to have suffered from encephalopathy, myopathy, developmental delay and lactic acidosis [Haack *et al*, 2012].

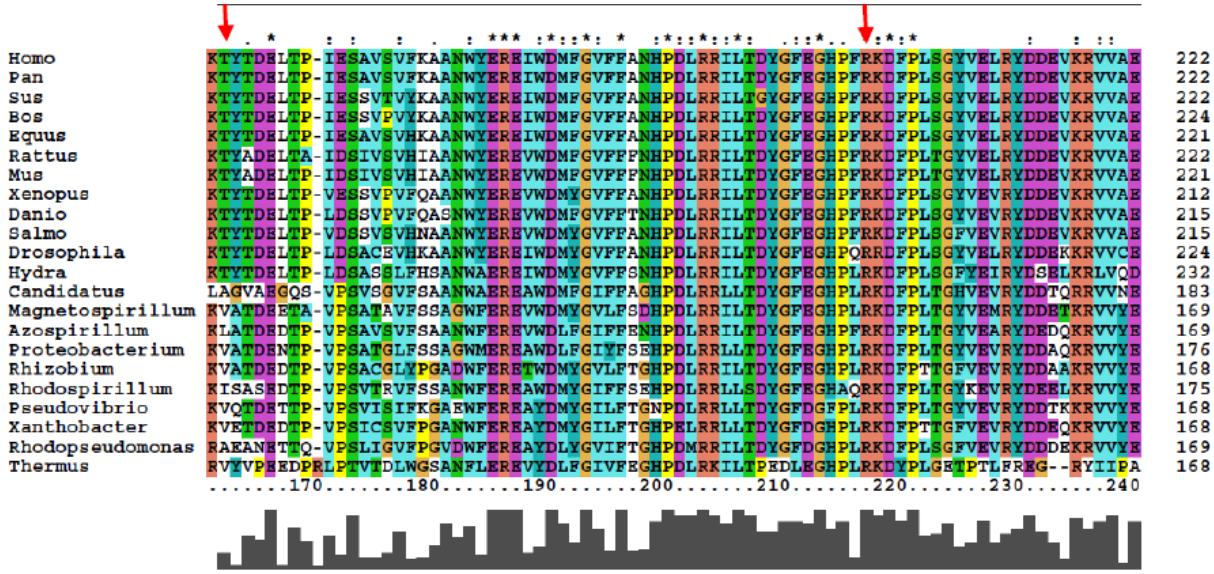


Figure 4.5: Multiple sequence alignment of *NDUFS3* from different species. The red arrow highlights the points of mutations.

4.1.2.1 Site-directed mutagenesis to engineer T145I+R199W double mutant

A two step PCR protocol described in section 2.11.1 was used to engineer the double mutant (T145I+R199W). The protocol of Phusion™ Site-directed mutagenesis kit (Finnzymes Cat#F541) was used to generate the first (T145I) point mutant with suitably designed primers described in section 2.11.1. A second PCR protocol was utilized with the plasmid having T145I mutation and the second set of primers to generate the double mutant.

In short, the entire plasmid was amplified with primers having suitable mutant base. The PCR product was phosphorylated and ligated. Two steps of PCR were performed to obtain the suitable double mutant (Figure 4.6A). The ligated plasmid was then transformed into *E coli* BL21(DE3) cells. The plasmids were sequenced to confirm the mutation (Figure 4.6B).

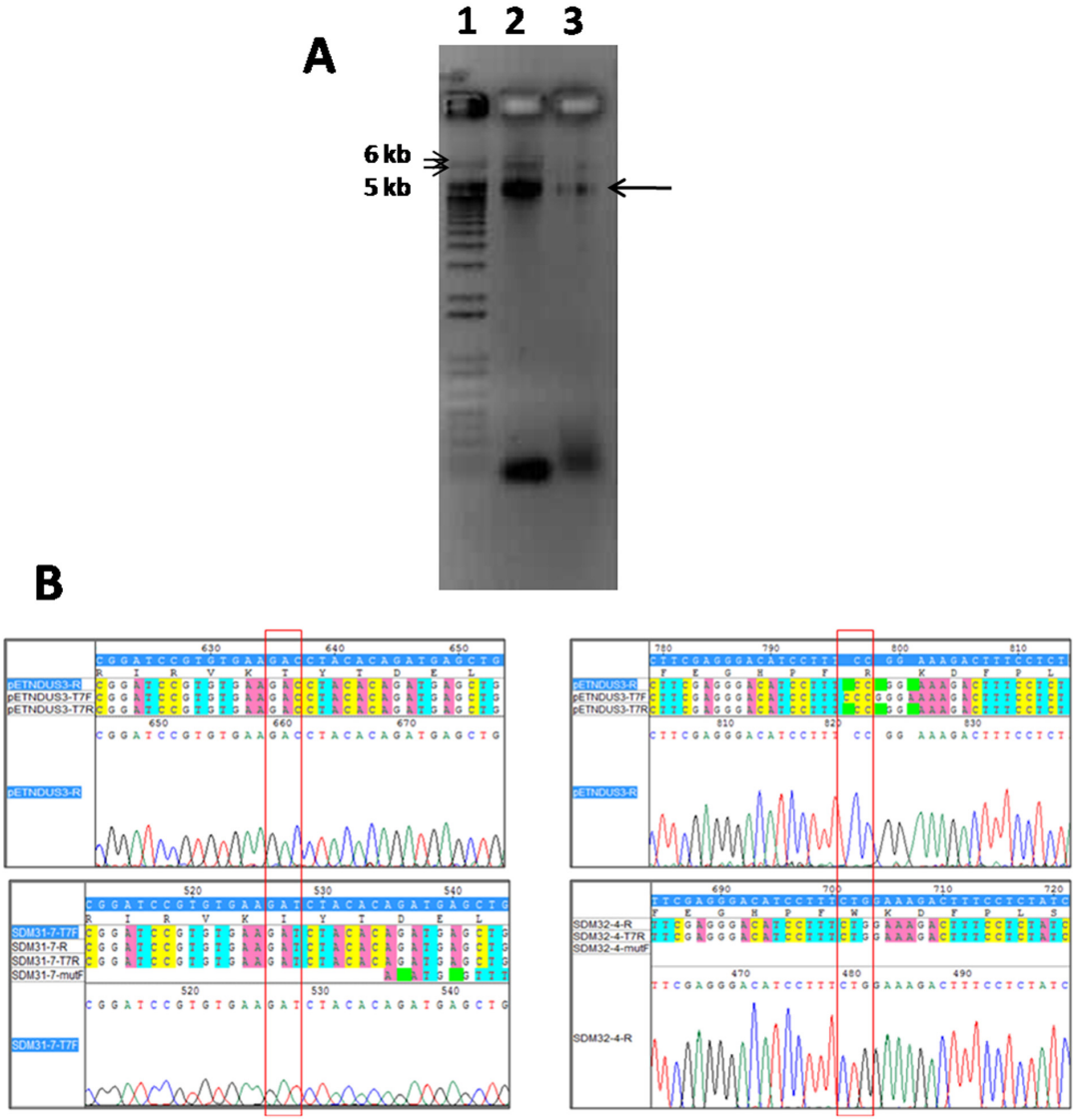


Figure 4.6: A: Image of 1% Agarose gel electrophoresis run; lane 1: 1 kb plus DNA ladder, lane2: PCR product with first set of primers (SDMF1 & SDMR1), lane 3: PCR product with second set of primers (SDMF2 & SDMR2), B: Sequence electropherogram. Upper panel is the w-t sequence and lower panel is the mutant sequence. The red box highlights the point of mutation.

4.1.2.2 Expression and purification of the w-t and double mutant (T145I+R199W)

The w-t and mutant recombinant proteins were expressed in the soluble form by the autoinduction method described in section 2.12. Both the proteins were purified by a 2 step column chromatography: Anion exchange (Q Sepharose column) followed by size exclusion (Sephadex G-200 column).

In short, the supernatant obtained after cell lysis by sonication was loaded onto the Q Sepharose column. The protein eluted out with 100mM NaCl. The fractions were pooled, dialyzed and loaded onto to Sephadex G-200 column to obtain a homogenous protein preparation (Figure 4.7). The yield of the w-t protein was 2 mg per litre and that of mutant was 1.5 mg protein per litre of culture.

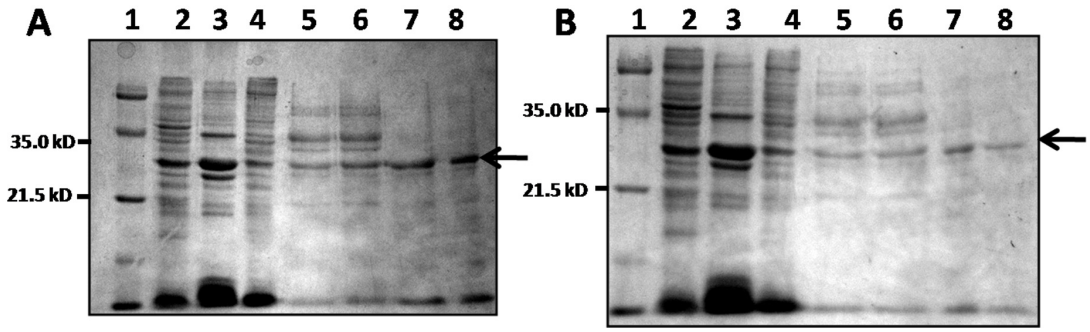


Figure 4.7: Purification of A: w-t *NDUFS3* and B: T145I+ R199W mutant. 12% SDS-PAGE; lane 1: Bio-rad low molecular weight protein markers, lane 2: uninduced cells, lane 3: autoinduced pellet, lane 4: autoinduced supernatant, lane 5-6: 100 mM NaCl fraction, lane 7-8: Sephadex G-200 fraction.

The purified proteins were confirmed by Western blot using monoclonal anti-His antibody and the molecular weights were confirmed by MALDI-TOF/TOF, to be 35.35 kDa and 35.43 kDa for the w-t and mutant (T145I+R199W) proteins, respectively (Figure 4.8).

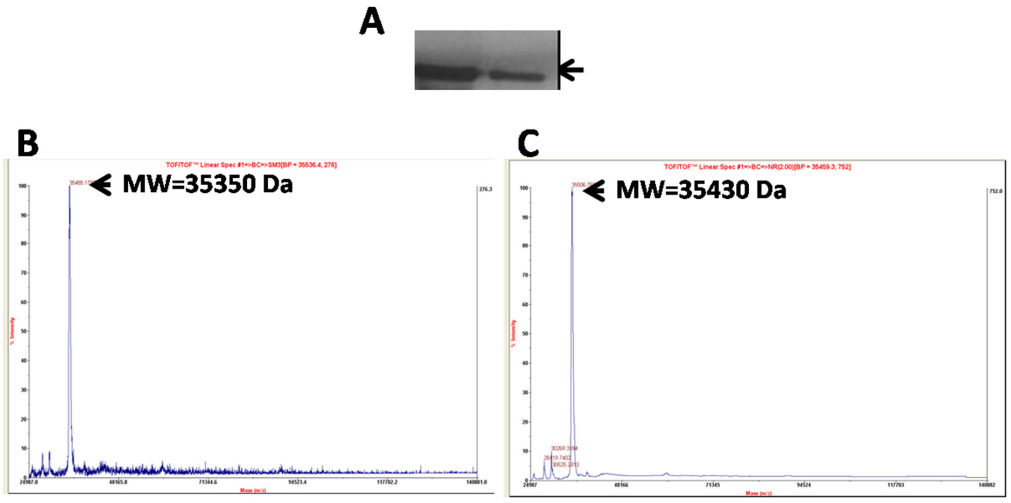


Figure 4.8: A: Western blot of 1: w-t, 2: T145I+R199W mutant with monoclonal anti-His antibody, MALDI TOF/TOF of B: w-t *NDUFS3* and C: T145I+R199W mutant protein.

4.2 *NDUFS7*

The *NDUFS7* gene (Entrez gene 374291) is located on the chromosome 19p13.3 [Hyslop *et al*, 1996] and was found to be expressed in all tissues. This gene is composed of 15 distinct introns. Several regulatory transcription factor binding sites are defined.

NADH dehydrogenase (ubiquinone) iron sulphur protein 7 expressed by this gene is composed of 213 amino acids and has a molecular weight of ~20 kDa. It is one of the most conserved subunits in the mitochondrial respiratory chain complex I and plays a central role in interaction with electron acceptor ubiquinone and in the proton translocation mechanism. The first 52 amino acid residues code for the mitochondrial targeting sequence. It shares 93% identity with the bovine protein.

4.2.1 Cloning of full length human *NDUFS7* gene

4.2.1.1 RNA extraction, cDNA preparation and amplification of *NDUFS7* gene

RNA extracted from the *HT29* cell line at a concentration of 415 ng μl^{-1} was utilized to prepare the cDNA library. The cDNA library was used to amplify *NDUFS7* gene of 656 bp with the primers described in materials and methods (Figure 4.9).

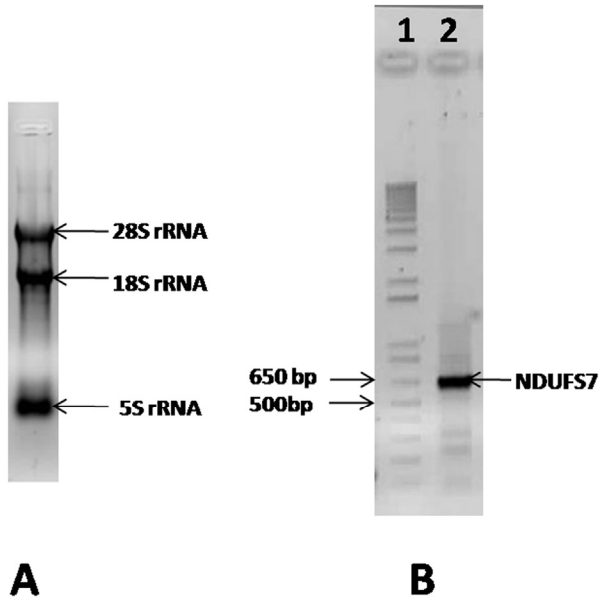


Figure 4.9: Image of A: Denaturing agarose gel electrophoresis showing the *HT 29* RNA used for preparing cDNA library and amplification of w-t *NDUFS7*. B: 1% Agarose gel electrophoresis, lane 1: 1 kb plus DNA ladder, lane 2: amplified full length *NDUFS7* (656 bp) gene.

4.2.1.2 Cloning in pGEM-T vector

The A tailed *NDUFS7* PCR product was ligated with the pGEM-T vector. The ligation mixture transformed in *E coli* DH5α cells. The colonies obtained on incubation were subjected to colony PCR and the plasmids isolated from them were sequenced to confirm the *NDUFS7* clones.

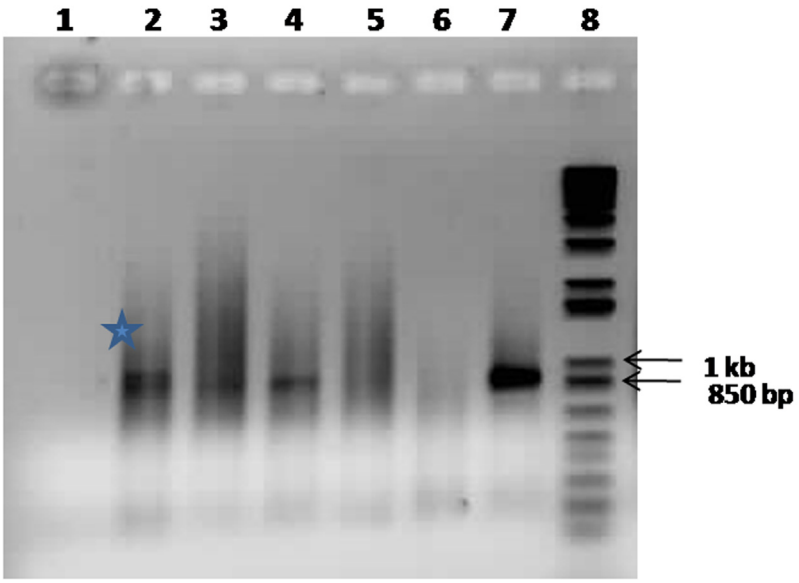


Figure 4.10: 1% agarose gel electrophoresis Lane 1-7: Colony PCR products for colony number 1-7 for *NDUFS7* gene cloned in pGEMT vector and transformed in *E. coli* DH5α cells, Lane 8: 1kb plus DNA Ladder. Star indicates the presence of amplified *NDUFS3* colony PCR product.

Colony number 2, 3, 4 and 7 showed an amplified product of appropriate size (Figure 4.10). The plasmids from these colonies were isolated and sequenced. All the 5 plasmids showed a 100 % sequence identity with the human *NDUFS7* gene. These were further used for subcloning into the expression vector.

4.2.1.3 Cloning in pET-28b(+) vector

The plasmids confirmed by sequencing were restricted digested with restriction enzymes NdeI and XhoI. The *NDUFS7* gene was ligated with a pre-digested pET-28b(+) vector in between the NdeI and XhoI sites.

The colonies obtained on transformation of the ligation mixture in *E coli* were subjected to colony PCR.

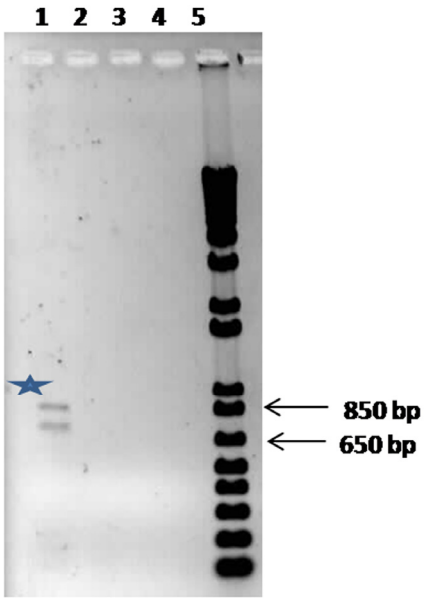


Figure 4.11: 1% agarose gel electrophoresis, Lane 1-4: Colony PCR products for colony number 1-4 for *NDUFS7* gene cloned in pET-28b(+) vector and transformed in *E. coli* DH5α cells, lane 5: 1 kb plus DNA ladder. Star indicates the presence of *NDUFS7* colony PCR product.

Colony number 1 showed an amplified PCR product of appropriate size (Figure 4.11). The plasmid from this colony was isolated and sequenced for confirmation. The sequence obtained was subjected to Blastp search (Figure 4.12). The sequence showed 100% sequence identity with the human NADH dehydrogenase (ubiquinone) Fe-S protein 7.

GENE ID: 374291 NDUFS7 | NADH dehydrogenase (ubiquinone) Fe-S protein 7, 20kDa (NADH-coenzyme Q reductase) [Homo sapiens] (Over 10 PubMed links)

Score = 435 bits (1118), Expect = 2e-120, Method: Compositional matrix adjust.
Identities = 213/213 (100%), Positives = 213/213 (100%), Gaps = 0/213 (0%)

```
Query  1    MAVLSAPGLRGFRILGLRSSVGLAVQARGVHQSVATDGPSSTQPALPKARAVAPKPSSRG  60
            MAVLSAPGLRGFRILGLRSSVGLAVQARGVHQSVATDGPSSTQPALPKARAVAPKPSSRG
Sbjct  1    MAVLSAPGLRGFRILGLRSSVGLAVQARGVHQSVATDGPSSTQPALPKARAVAPKPSSRG  60

Query  61   EYVVAKLDDLVNWARRSSLWPMTFGLACCAVEMMHMAAPRYDMDRFGVVFRASPRQSDVM  120
            EYVVAKLDDLVNWARRSSLWPMTFGLACCAVEMMHMAAPRYDMDRFGVVFRASPRQSDVM
Sbjct  61   EYVVAKLDDLVNWARRSSLWPMTFGLACCAVEMMHMAAPRYDMDRFGVVFRASPRQSDVM  120

Query  121  IVAGTLTNKMAPALRKVYDQMPEPRYVVSMGSCANGGGYYHYSYSVVRGCDRIVPVDIYI  180
            IVAGTLTNKMAPALRKVYDQMPEPRYVVSMGSCANGGGYYHYSYSVVRGCDRIVPVDIYI
Sbjct  121  IVAGTLTNKMAPALRKVYDQMPEPRYVVSMGSCANGGGYYHYSYSVVRGCDRIVPVDIYI  180

Query  181  PGCPPTAEALLYGILQLQRKIKRERRLQIWYRR  213
            PGCPPTAEALLYGILQLQRKIKRERRLQIWYRR
Sbjct  181  PGCPPTAEALLYGILQLQRKIKRERRLQIWYRR  213
```

Figure 4.12: Blastp results to show the matching of cloned sequence with that of human *NDUFS7*.

The colony was maintained by preparing a 15% glycerol stock and stored at -80°C.

4.2.2 Reported mutations

Several reports of mutations causing Complex-I deficiency from *NDUFS7* have highlighted the functional importance of this subunit. A V122M mutation was described by Smeitink *et al* [Smeitink *et al*, 1999] in two siblings with Leigh syndrome. Visch *et al* [Visch *et al*, 2004] showed that the V122M mutation also causes a defect in calcium homeostasis leading to increased calcium levels in the cell that may be toxic to the cell. Another mutation reported was R145H; causing Leigh syndrome and severe Complex-I deficiency [Lebon *et al*, 2007b]. Both the residues, namely, Val 122 and Arg 145 are highly conserved from bacteria to mammals (Figure 4.13).

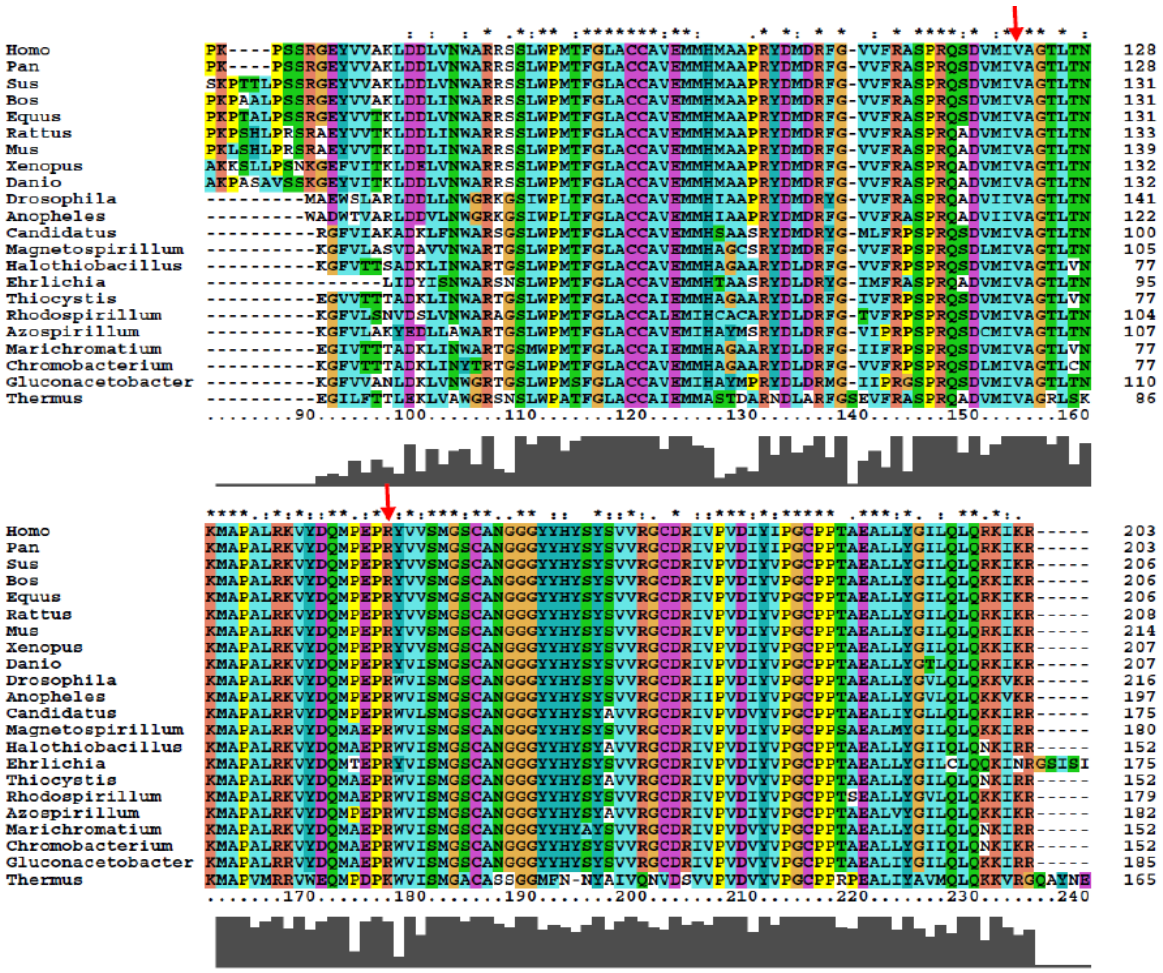


Figure 4.13: Multiple sequence alignment of *NDUFS7* from different species. The red arrow highlights the points of mutations.

Apart from point mutations in the exons, Lebon *et al* [Lebon *et al*, 2007a] identified a mutation in the intron 1 of the gene, resulting in the creation of an alternative splice

site and the generation of a 122 bp cryptic exon. This resulted in the generation of a shortened mutant protein of 41 amino acids. Also in most cases of mutation a fully assembled Complex-I was absent indicating the structural effects that these mutations may have on the subunit resulting in lack of assembly.

4.2.2.1 Site-directed mutagenesis to engineer V122M and R145H point mutants

The protocol of Phusion™ Site-directed mutagenesis kit (Finnzymes Cat#F541) was utilized to generate the point mutants with suitably designed primers described in section 2.11.2. In short, the entire plasmid was amplified with primers having the suitable mutant base. The PCR product was phosphorylated and ligated. The ligated plasmid was then transformed into *E coli* BL21(DE3) cells (Figure 4.14A). The plasmids were sequenced to confirm the mutation (Figure 4.14B).

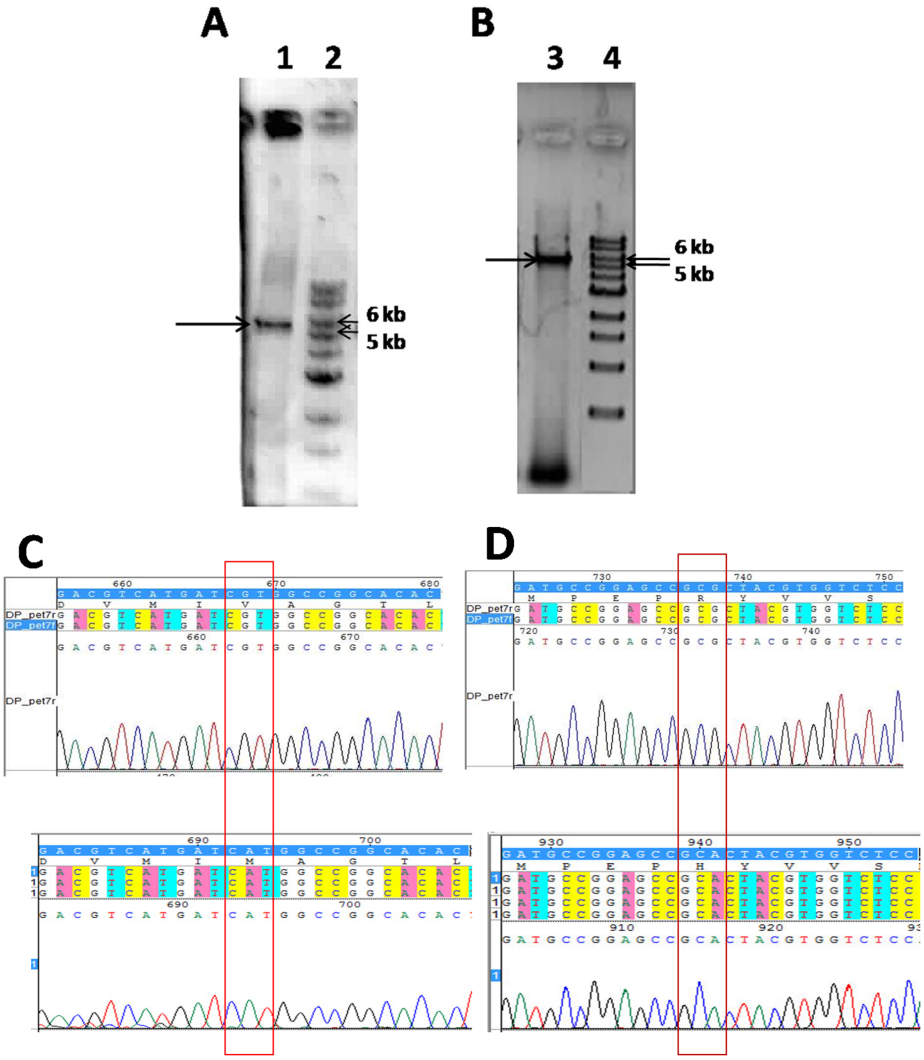


Figure 4.14: A: 1% agarose gel; lane 1: V122M mutant PCR product, lane 2: 1 kb plus NEB DNA ladder, B: 1% agarose gel lane 1: R145H mutant PCR product, lane 2:1 kb plus NEB DNA

ladder, C & D: Sequence electropherogram. Upper panel is the w-t sequence and lower panel is the mutant sequence of V122M and R145H respectively. The red box highlights the point of mutation.

4.2.2.2 Expression and purification of the w-t *NDUFS7,* V122M and R145H mutant proteins

The w-t *NDUFS7* and its mutants V122M and R145H were produced in soluble form by the autoinduction method. In short, the primary culture of the clones was subjected to prolonged incubation at low temperature in the rich ZY autoinduction medium. The recombinant proteins were purified from the supernatant by a 2 step column chromatography; Affinity chromatography (Ni-NTA) followed by size exclusion chromatography (Superose 12) to obtain a homogenous protein preparation. The proteins eluted from the Ni-NTA column with 100mM imidazole gradient fraction. The fractions were pooled, dialyzed and concentrated and the protein was purified to homogeneity on the Superose 12 column (Figure 4.15).

The w-t protein was obtained at a concentration of 2.5 mg per litre and both V122M and R145H at 1.8 mg of protein per litre of culture.

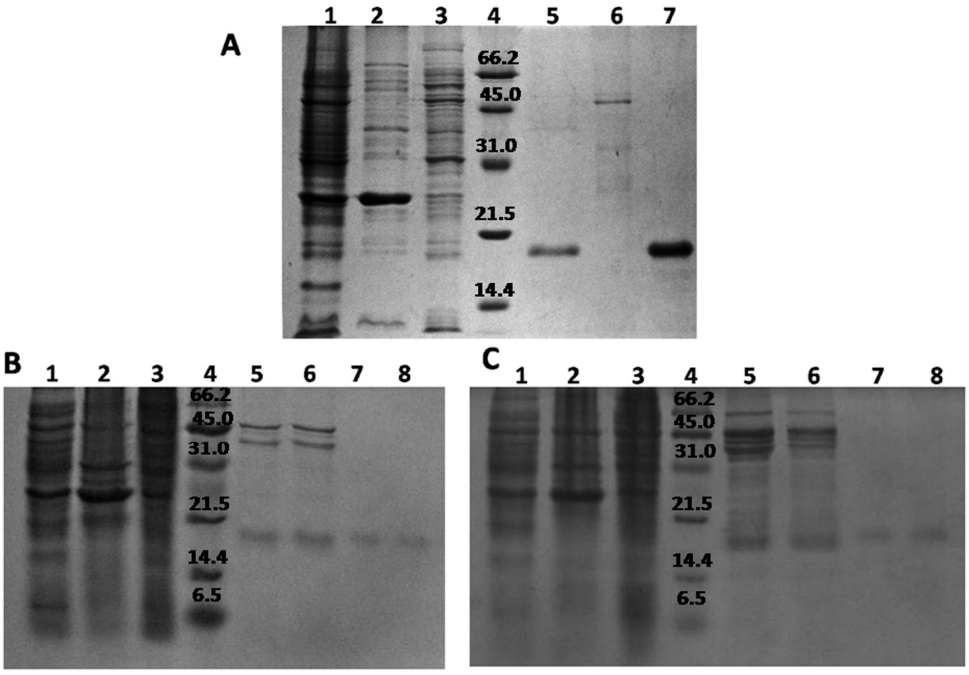


Figure 4.15: 12% SDS PAGE. Purification of A: w-t *NDUFS7,* lane 1: uninduced cells, lane 2: autoinduced pellet, lane 3: autoinduced supernatant, lane 4: Bio-rad low range protein marker, lane 5: 100 mM imidazole Ni-NTA fraction, lane 6: Superose 12 peak 1, lane 7: Superose 12 peak 2 (purified w-t *NDUFS7*) and B: V122M mutant, C: R145H mutant; lane 1: uninduced cells, lane

2: autoinduced pellet, lane 3: autoinduced supernatant, lane 4: Bio-rad broad range molecular weight protein markers, lane 5-6: 100 mM imidazole Ni-NTA fraction, lane 7-8: Superose 12 (purified mutant protein).

The purified proteins were confirmed by Western blot using monoclonal anti-His antibody and the molecular weights were confirmed by MALDI-TOF/TOF, to be 20.71, 20.72 and 20.68 for w-t, V122M and R145H, respectively (Figure 4.16).

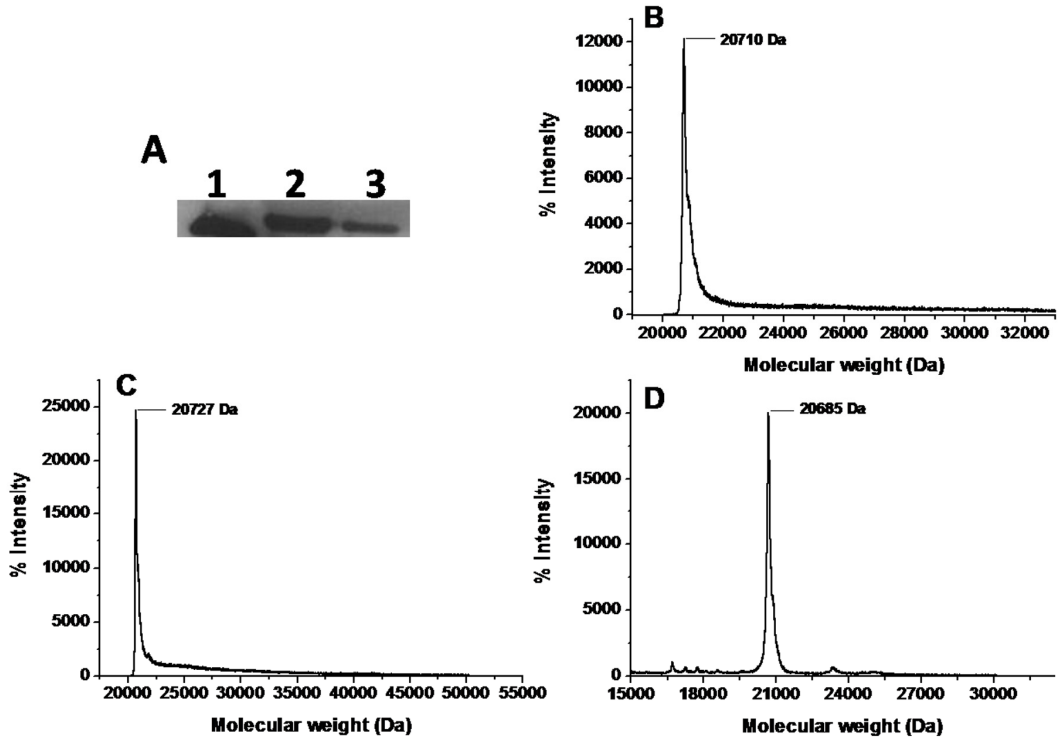


Figure 4.16: A: Western blot of 1: w-t, 2:V122M and 3: R145H proteins with monoclonal anti-His antibody, MALDI TOF/TOF of B: w-t *NDUFS7* and C:V122M and D: R145H proteins.

4.3 Conclusion

The core subunits of the Q module along with mitochondrially expressed ND1 were cloned and expressed. As there were problems in expression or solubility of the recombinant *NDUFS2, 8* and ND1, only *NDUFS3* and *7* were further purified and characterized.

NDUFS3 and *7* full length genes were isolated from the RNA *HT29* cell line. Double mutant (T145I+R199W) of *NDUFS3* was engineered with a 2 step PCR protocol. Single point mutants of *NDUFS7* (V122M and R145H) were prepared in a single step

PCR protocol. Expression of protein into soluble fraction was achieved by autoinduction method. The yield of the mutant proteins was less than the w-t protein in each case.

The wild-type form of *NDUFS3* and *NDUFS7* recombinant proteins along with their mutants T145I+R199W of *NDUFS3* and V122M and R145H of *NDUFS7* were purified to homogeneity. The w-t and mutant proteins did not show any significant difference in their molecular weights.

CHAPTER 5

Comparative studies using biophysical techniques of the human mitochondrial *NDUFS3* subunit and its Leigh syndrome causing mutant

NADH dehydrogenase iron sulphur protein 3 or *NDUFS3* is one of the core subunits of the ubiquinone reduction module of the Complex-I. It is one of the eight core subunits of Complex-I encoded by the nuclear genome (n-DNA). The *NDUFS3* gene consists of seven exons of varying sizes. FISH analysis and radioactive *in situ* hybridization showed that the *NDUFS3* gene was located in the 11p11-p12 region of the chromosome [Procaccio *et al*, 2000]. The protein (UniProtKB O75489) encoded by the gene comprises of 263 amino acids and has a predicted signal sequence of 38 amino acids (Table 2.1). It is also known as NqoC in bacteria or 30 kDa subunit in mammals. The structural role of *NDUFS3* protein and changes due to mutation in the Q module studied by using *in silico* approaches were extensively discussed in Chapter 3. This chapter highlights the differences in the w-t and mutant *NDUFS3* protein behaviour *in vitro* or in solution.

5.1 *NDUFS3* and its function

The *NDUFS3* protein is involved in the early assembly of the peripheral arm of the Complex-I in the mitochondrial matrix. Inhibiting the mitochondrial translation pathway leads to the accumulation of *NDUFS3* in two distinct subassemblies reflecting its pivotal role in initiating the assembly of Complex-I [Vogel *et al*, 2007; Dieteren *et al*, 2008]. Its role has also been implicated as a biomarker for breast carcinoma [Suhane *et al*, 2011]. Also, aberrant expression of *NDUFS3* is correlated with hypoxic or necrotic regions of the tumour relating it with tumour aggressiveness [Suhane *et al*, 2013]. Thus, apart from being a structural subunit of the Q module, the functional role of this protein in the cell has been highlighted.

The presence of *NDUFS3* protein is thus essential for the cell as it is a part of the ROS minimizing and energy generating cycle. A double mutant of *NDUFS3*; T145I+R199W has been reported to cause optic atrophy, Complex-I deficiency and delayed onset case of Leigh syndrome [Benit *et al*, 2004]. Mutant *NDUFS3* proteins are reported to express at normal levels in the patient cell line. This highlights the fact that mutations affect the functioning of the subunit or its assembly in the Complex-I. Accumulation of assembly intermediates in the patient cell lines points to the possible structural changes caused by the double mutation affecting the formation of correct assembly of the Complex-I.

Thus, for understanding the pathological consequences of Complex-I defective mitochondria, in depth characterization of proteins involved in Complex-I biogenesis is essential. *NDUFS3* has been considered in the present study to highlight the differences between the w-t and double mutant causing Leigh syndrome. The differences give insight into the changes in behaviour of the w-t *in vitro* caused by mutation. Finally, the broad spectrum of symptoms in Leigh syndrome, caused by the double mutant have been correlated to the various biophysical changes observed at molecular level *in vitro* in the present study.

5.2 W-t and T145I+R199W mutant *NDUFS3* protein

The w-t and double mutant proteins were cloned, expressed and purified by methods already described in chapter 2 and 4. Both Thr145 and Arg199 are conserved amino acid residues in the *NDUFS3* protein (Figure 5.1).

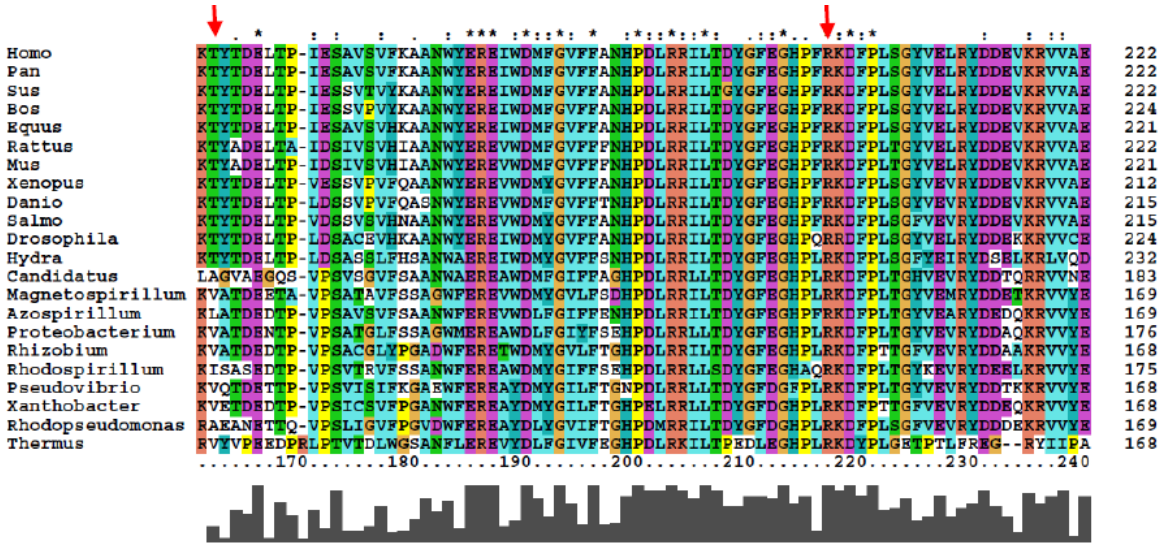


Figure 5.1: Multiple sequence alignment of *NDUFS3* subunit from mammals to bacteria. The red arrows indicate the points of mutations.

The T145I mutation changes a neutral amino acid residue to a polar residue in a putative casein II phosphorylation site. The mutation is absent in healthy individuals except in one case where the individual was heterozygous for the mutation with the absence of any other mutation. The R199W mutation changed a conserved amino acid and was absent in healthy controls. These two mutations together are the disease causing mutations in the *NDUFS3* gene [Benit *et al*, 2004].

The homology modelling study of *NDUFS3* protein described in Chapter 3 based on the crystal structure of *Thermus thermophilus* Complex-I subunits has helped to locate the positions of the mutation sites. The T145I mutation took place in a β-sheet while the R199W mutation was present in a highly flexible loop (Figure 5.2).

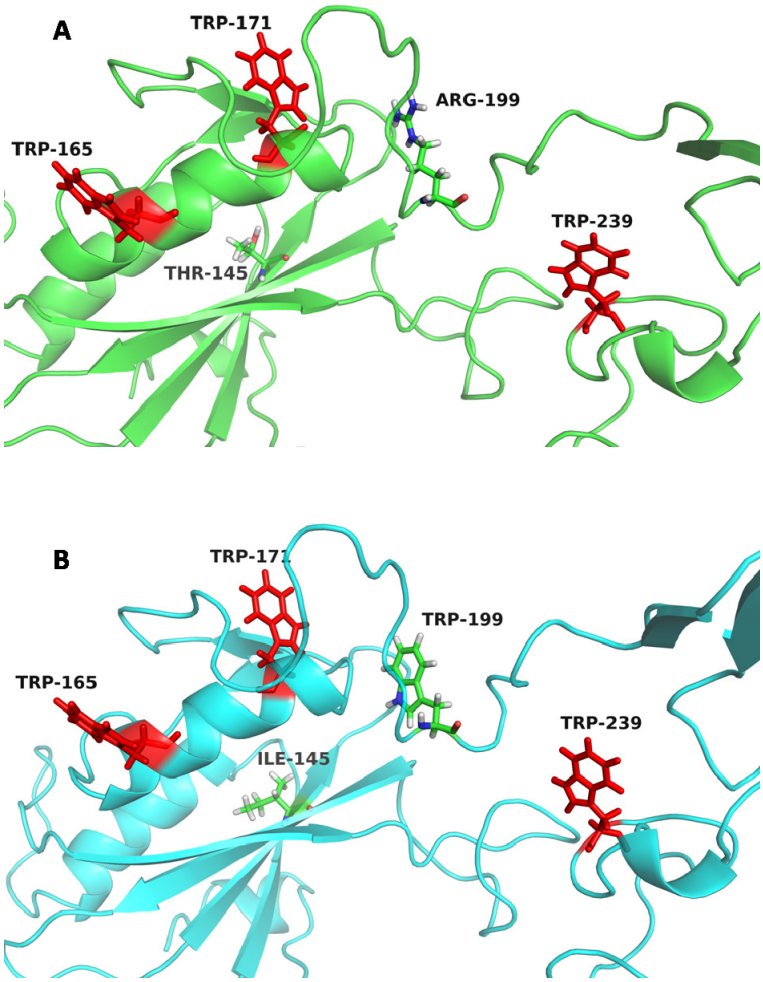


Figure 5.2: A: Homology model of NDUFS3 w-t protein showing positions of 3 tryptophans, B: Homology model of NDUFS3 mutant protein showing positions of the 3 tryptophans and the mutated T145I and R199W residues.

A combined effect of the two mutations was also seen to affect the root mean square fluctuations, radius of gyration and hydrogen bonding of the protein according to the *in silico* studies described in Chapter 3. The purified w-t and double mutant *NDUFS3* were compared for changes in their biophysical properties.

5.3 Steady State Fluorescence spectroscopy and Circular Dichroism

The intrinsic fluorescence maximum (λ_{max}) of the full length w-t *NDUFS3* protein was at 342 nm indicating tryptophans to be in a partially polar environment. A red shift in case of the mutant protein to 346 nm was observed indicating increased polarity of environment of the tryptophans, which could be due to the additional tryptophan (W199) present in the mutant (Figure 5.3.1). W-t *NDUFS3* has 5 tryptophan residues and decomposition analysis of the steady state fluorescence spectra revealed Class A (11%) and II (89%) of tryptophans. The mutant protein has 6 tryptophan residues, owing to the additional tryptophan at position 199; the classes have slightly differed as S (21%) and II (79%). Class A represents fluorophores which can't form hydrogen bond, Class S represents buried fluorophores or tryptophans in 1:1 ratio and Class II indicates tryptophan associated with water molecules and having low dipole relaxation time [Burstein *et al*, 2001; Reshetnyak *et al*, 2001a; 2001b]

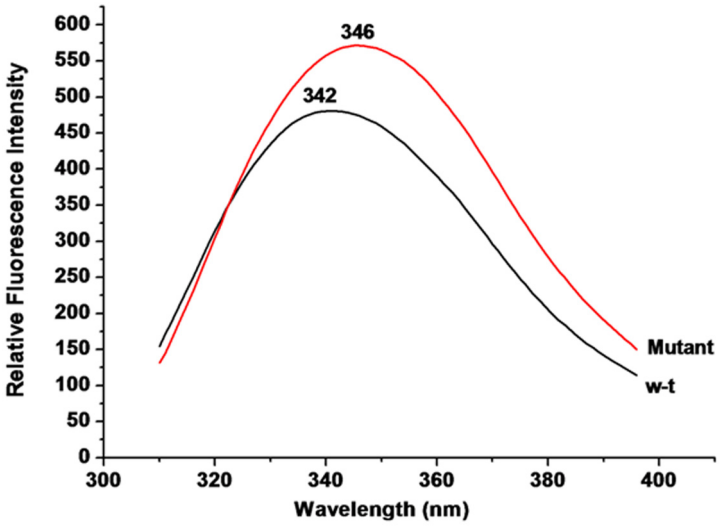


Figure 5.3.1: Steady-state fluorescence spectra of the native (342 nm) and mutant (346 nm) *NDUFS3* protein. Values marked on the spectrum represent λ_{max} at respective conditions.

The far UV CD spectra of both w-t and mutant *NDUFS3* proteins are shown in Figure 5.3.2. The composition of secondary structure elements as calculated from CDPro are; for w-t protein: 56% α-helices, 10.2% β-sheets, 10.5% turns and 23.3% unordered structure. The mutant on the other hand showed 50.3% α-helices, 10.3% β-sheets, 17.1% turns and 22.3% unordered structure, thus indicating that both the full length w-t and mutant proteins were predominantly α-helical. The decrease in the α-helical content from 56% to 50.3 % with an increase in the turns from 10.5 to 17.1 % in the mutant protein as compared to the w-t are due to the substitutions T145I and R199W

(Table 5.1). . Though the partial model predicted by MODELLER was dominated by unordered structure (Table 5.1), the full length purified proteins showed a high helical content. This discrepancy is because the model lacks the signal sequence which may contribute to the secondary structure content in the full length purified proteins.

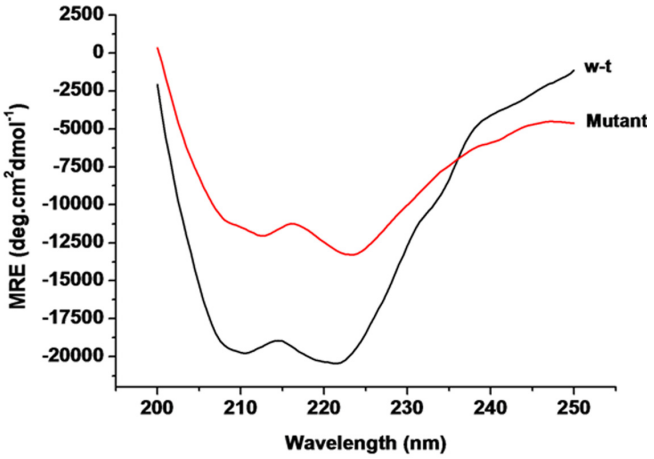


Figure 5.3.2: Far UV CD spectroscopic scan for the w-t (black) and mutant (red) protein.

Table 5.1: Secondary structure composition of the w-t and mutant (T145I+R199W) proteins. Calculated with the CDSSTR of the CDPro suite. A part of the sequence could not be modelled due to absence of template structure. So, the values for model are only indicative and not actual.

	% Helix	% Sheet	% Turn	% Unordered	NRMSD
w-t	56	10.2	10.5	23.3	0.049
Mutant (T145I+R199W)	50.3	10.3	17.1	22.3	0.043
Partial *NDUFS3* model	32	8	11	49	NA

A noticeably different tertiary structure was observed for the mutant protein in the near UV CD analysis with prominent signal in the tyrosine and tryptophan regions as compared to the w-t (Figure 5.3.3).

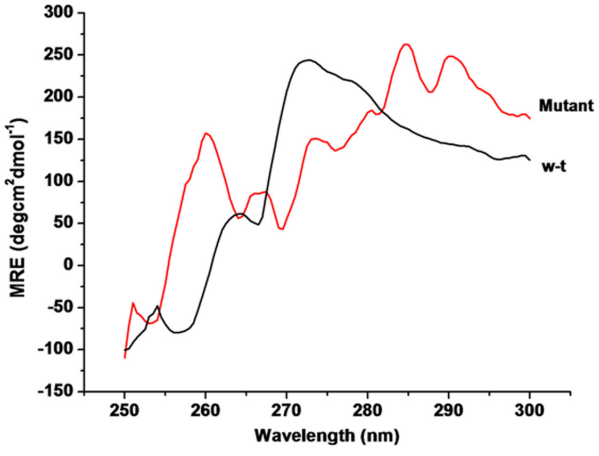


Figure 5.3.3: Near-UV CD spectroscopic scans for the w-t (black) and mutant (red) protein.

5.4 Time resolved fluorescence studies

The lifetimes of intrinsic fluorescence of the w-t and mutant protein were deduced from the fluorescence decay curve by fitting it to a bi-exponential function. Two decay times τ_1 and τ_2 with their corresponding weight factors α_1 and α_2 and their relative contributions were estimated (Table 5.2). In the w-t protein, the shorter life-time τ1, 2.07 ns showed 55% and the longer one τ2, 5.50 ns showed 45% contribution as against the mutant protein where the shorter component τ1, 1.51 ns showed 33% and longer one τ2, 6.62 ns showed 67% contribution. The tryptophan population/ conformer with a shorter lifetime, decaying faster is supposed to be on the surface of the protein, while one with a longer life-time, decaying slowly is buried in the interior. In the w-t protein, both the components had an equal contribution, while in the mutant protein the longer lifetime contributed much more in the total fluorescence. However, the red shift observed in the mutant protein steady state fluorescence indicated a more polar environment. Thus, a distinct change in the tryptophan microenvironment and populations of tryptophan residues was observed for w-t and mutant protein.

Table 5.2: The fluorescence decay lifetimes of *NDUFS3* and their pre-exponential factors for the w-t and mutant (T145I+R199W). Figures in the bracket indicate percent contribution to the total fluorescence.

	τ1(ns)	τ2(ns)	α1	α2	χ2	τ	<τ>
W-t	2.07	5.50	0.058	0.018	1.015	2.88	3.62
	(55)	(45)					
Mutant (T145I+R199W)	1.51	6.62	0.061	0.028	1.038	3.12	4.92
	(33)	(67)					

5.5 Solute quenching studies

Solute quenching studies for the w-t protein as well as T145I+R199Q mutant protein and their denatured forms (using 6M Gdn-HCl) were carried out. Three quenchers were utilized for the same: Acrylamide (Neutral), KI (Negative) and CsCl (Positive) to characterize the change in the tryptophan environments of the proteins.

5.5.1 Native *NDUFS3* and mutant T145I+R199W

The rate of quenching K_{sv}, calculated from the Stern-Volmer plots (Figure 5.4), for all the quenchers was higher for mutant compared to the w-t protein, indicating subtle changes in the tryptophan microenvironment (Table 5.3)

Table 5.3: Quenching of tryptophan fluorescence of w-t, mutant (T145I+R199W) and denatured w-t and mutant (in 6M Gdn-HCl), respectively. The two values of *Ksv* in quenching by CsCl and mutant by KI and CsCl indicate *Ksv* 1 and *Ksv* 2 reflecting the two parts of the curve linearly fitted.

	Acrylamide		**KI**		**CsCl**	
	K_{sv} (M^{-1})	fa	K_{sv} (M^{-1})	fa	K_{sv}(M^{-1})	fa
w-t	2.24	0.76	1.00	0.41	0.67	0.12
					0.24	
Mutant (T145I-R199W)	3.86	0.65	3.37	0.36	1.6	0.12
			1.17		0.18	
Denatured w-t	6.4 (Ks 2.0)	1.00	4.85	0.65	0.63	0.10
Denatured mutant	8.47	1.00	4.74	0.69	0.82	0.22

In the w-t protein 76% of the fluorescence was accessible to acrylamide while for the mutant protein the accessibility had decreased to 65% (Table 5.3). This change in accessibility could be due to the additional tryptophan in the highly flexible loop of the mutant protein.

The positive and negative charge density around the surface tryptophan/s is equal for the w-t as seen from the K_{sv} values (Table 5.3), while the positive charge density is higher in case of the mutant protein. Also, in the mutant protein, K_{sv} for iodide was much higher than that for caesium indicating more positively charged environment of surface tryptophans. This reflects the increased polarity resulted due to rearrangement of atoms/residues in the vicinity of tryptophans, which also correlates with the red shift in the λ_{max}.

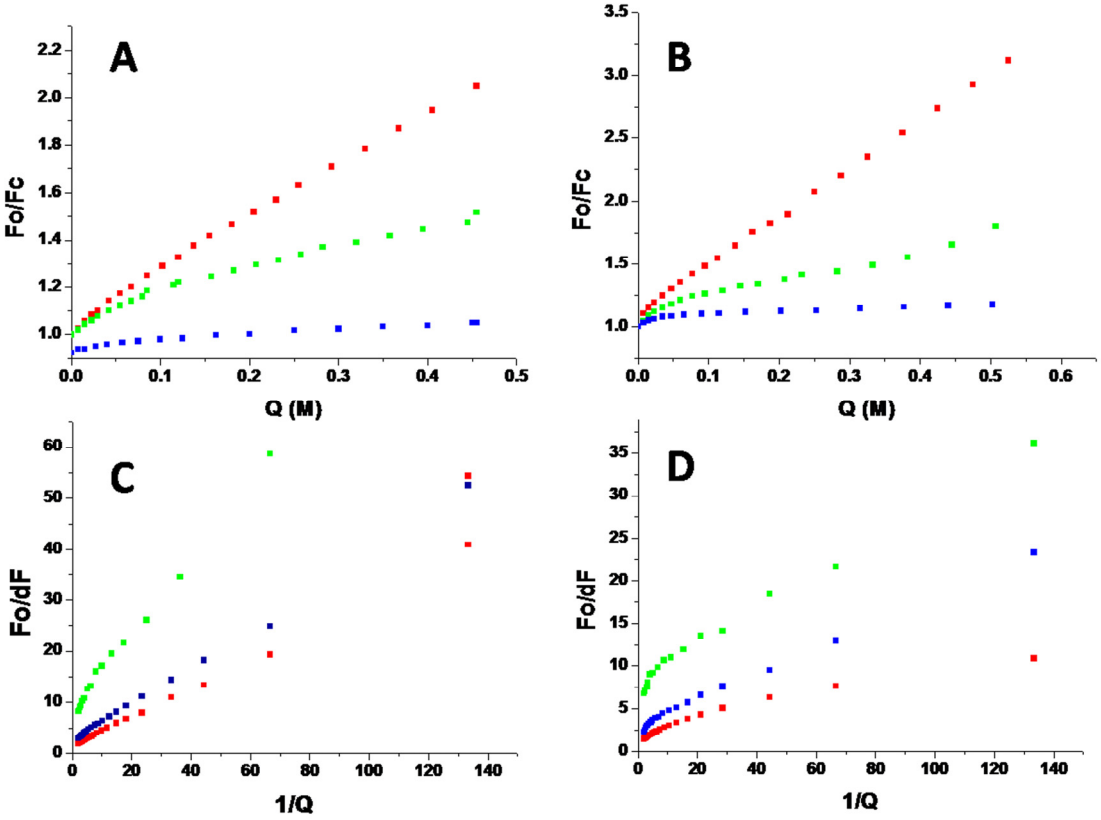


Figure 5.4: Stern-Volmer plots for A: w-t , B: mutant (T145I+R199W); Modified Stern-Volmer plots for C: w-t and D: mutant (T145I+R199W) proteins. Red squares indicate quenching with acrylamide, blue with KI and green with CsCl.

5.5.2 Denatured *NDUFS3* and mutant T145I+R199W

On denaturing the proteins using 6M Gdn-HCl, both the w-t and mutant (T145I+R199W) proteins showed an increase in the K_{sv} values (Figure 5.5, Table 5.3). The positive charge density around the tryptophan was higher in the case of both the w-t and mutant proteins. Both the w-t and mutant proteins showed 100 % accessibility for acrylamide on denaturation. Accessibility of KI was almost the same for both, however, in the case of caesium chloride the accessibility in the case of mutant was slightly higher than for the w-t.

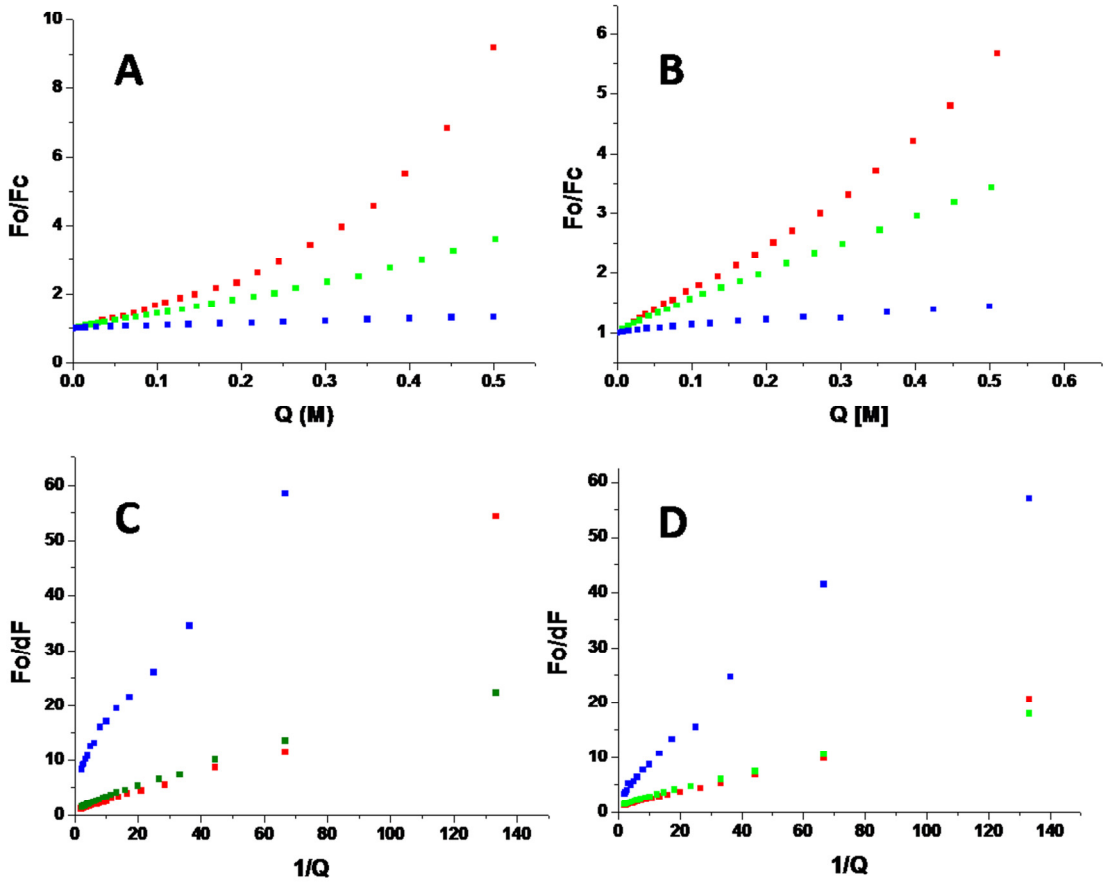


Figure 5.5: Stern-Volmer plots for denatured (6M Gdn-HCl) proteins A: w-t, B: mutant (T145I+R199W); Modified Stern-Volmer plots for C: w-t and D: mutant (T145I+R199W) proteins. Red squares indicate quenching with acrylamide, blue with KI and green with CsCl.

The w-t protein denatured using 6M Gdn-HCl for 16 hours showed a positive curvature during quenching with acrylamide indicating that the quenching has both dynamic and static components (Figure 5.5A). The static component involves complex formation whereas the dynamic mechanism involves collisions with acrylamide during the lifetime of tryptophan in the excited state. These components were resolved and the K_{sv} obtained for acrylamide quenching of the denatured protein was 6.55 M^{-1} (Table 5.4, Figure 5.6A and B).

Table 5.4: The lifetimes of fluorescence decay of the denatured w-t *NDUFS3* (in 6M Gdn-HCl) and the corresponding pre-exponential factors along with calculated average lifetimes τ /<τ> for acrylamide quenching.

Q	τ1 (ns)	α1	τ2 (ns)	α2	τ (ns)	<τ> (ns)	χ2
0.000	1.15	0.069	3.59	0.036	1.99	2.66	1.00
0.015	0.91	0.075	3.16	0.041	1.71	2.38	1.13
0.030	0.97	0.078	2.92	0.037	1.60	2.12	1.08
0.045	0.96	0.082	2.63	0.037	1.48	1.88	1.14
0.060	0.82	0.094	2.42	0.035	1.25	1.66	1.05
0.085	0.74	0.100	2.19	0.038	1.14	1.51	1.08
0.110	0.81	0.108	2.19	0.026	1.08	1.35	1.05
0.135	0.72	0.126	2.09	0.025	0.95	1.22	1.07
0.160	0.75	0.127	2.02	0.021	0.93	1.14	1.04
0.195	0.66	0.142	1.85	0.023	0.83	1.03	1.11
0.230	0.73	0.141	2.08	0.017	0.88	1.08	1.07
0.280	0.71	0.158	2.63	0.005	0.77	0.91	1.01
0.330	0.6	0.173	1.96	0.011	0.68	0.83	1.06
0.380	0.58	0.179	1.9	0.011	0.59	0.60	1.07
0.455	0.51	0.196	1.76	0.010	0.57	0.70	1.06
0.530	0.54	0.204	2.92	0.003	0.57	0.72	1.09

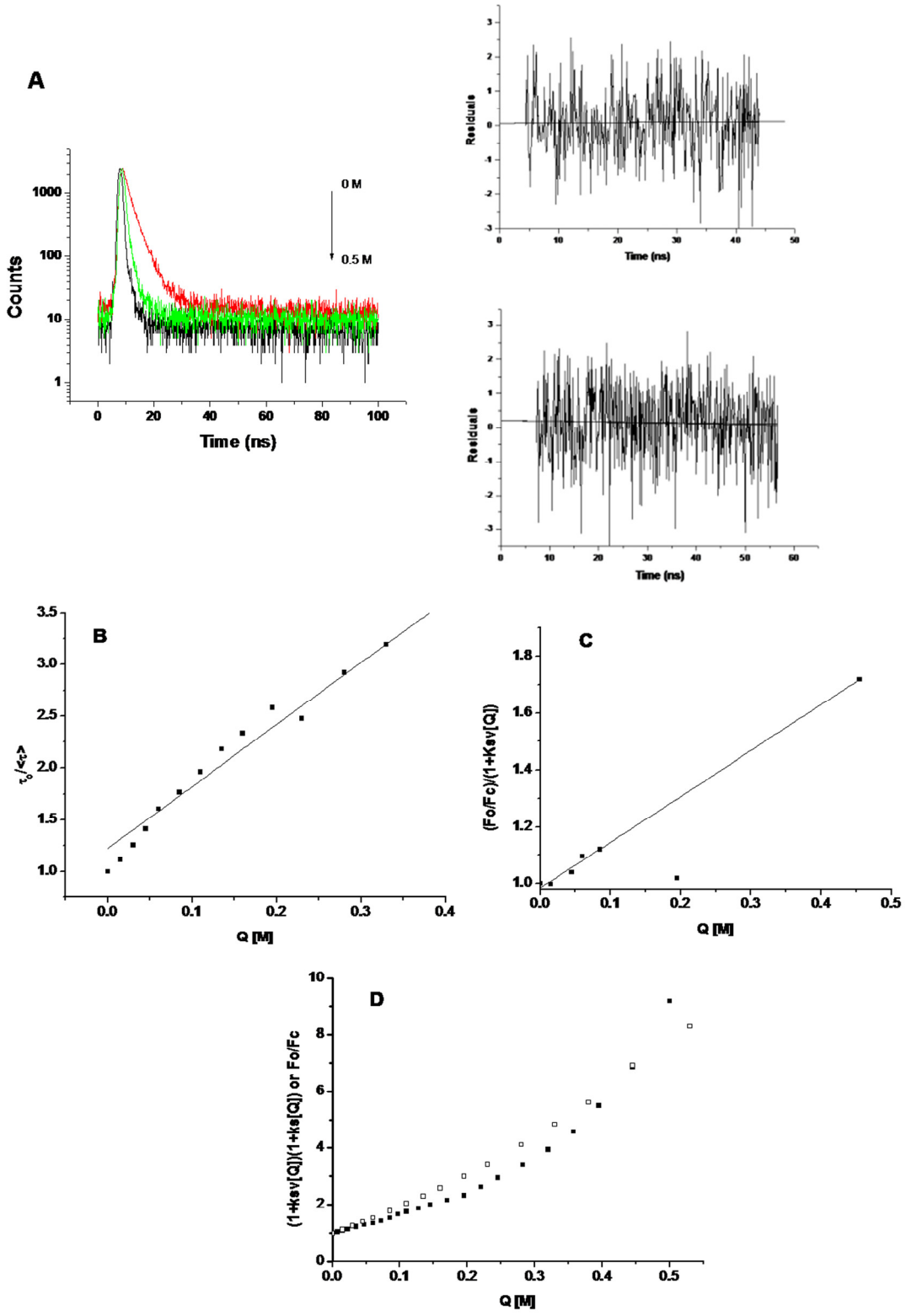


Figure 5.6: A: Quenching of denatured w-t NDUFS3 fluorescence with acrylamide on time resolved spectrofluorimeter. B: Plot of τo/τ versus [Q], C: Plot of (Fo/Fc)/(1+Ksv[Q]) versus [Q] D: Plot of ((1+K_{sv}[Q])(1+K_s[Q]) or Fo/Fc versus [Q].

Plotting a graph of (Fo/Fc)/(1+K_{sv}[Q]) against [Q] gave the value of the static quenching constant to be 2.0 M^{-1} (Figure 5.5C). The bimolecular quenching constant kq was calculated to be 2.47*10^9 $M^{-1}s^{-1}$. The mutant protein on denaturation did not show any upward curvature.

5.6 Aggregation studies

5.6.1 Rayleigh light scattering: The mutant protein showed higher scattering intensity as compared to the w-t protein at the same concentration, both at 25°C and 37°C (Figure 5.7.1) when measured in the time drive module.

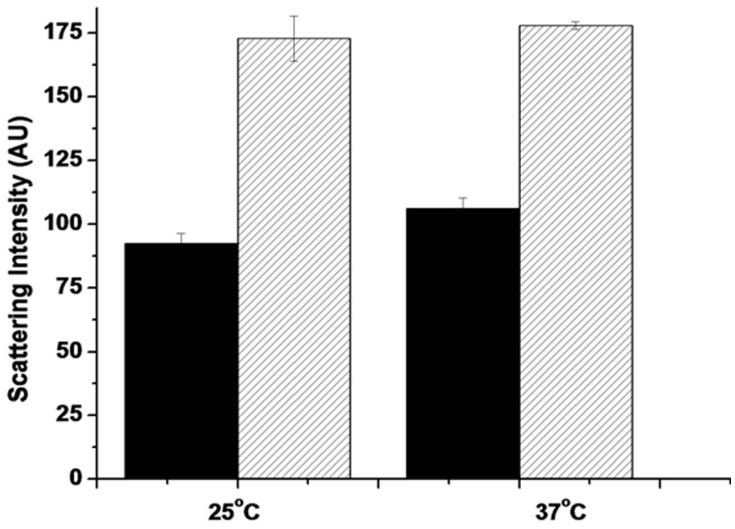


Figure 5.7.1: Rayleigh scattering bar graph, scattering intensity (AU) plotted for w-t (filled) and mutant (stripped) protein at 25°C and 37°C.

5.6.2 Th-T dye binding assay: The amyloidophilic Th-T dye shows a characteristic increase in fluorescence upon binding to amyloid fibrils [Khurana *et al*, 2005; Chang *et al*, 2009; Groenning, 2010]. The Th-T fluorescence of amyloid results from the restricted rotation of the two rings in the Th-T by the hydrophobic pocket formed by the amino acid side chains of the amyloid fibrils. The w-t *NDUFS3* protein didn't show an increase in fluorescence emission at 37°C in the concentration range of 30-100 µg ml^{-1}. However, the mutant protein showed a characteristic increase in fluorescence intensity at 25°C as well as 37°C in the same concentration range (Figure 5.7.2).

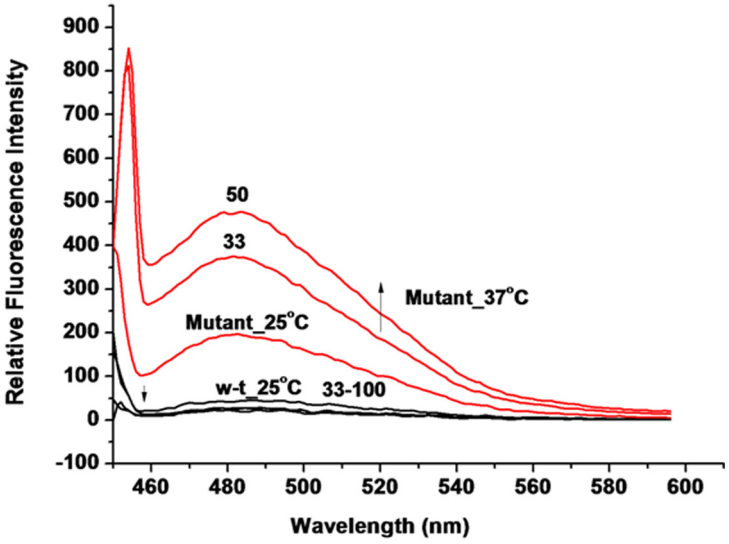


Figure 5.7.2: Thioflavin-T dye binding assay, for w-t (black) and mutant (red) protein at 25°C and 37°C. The numbers on the spectra represent the concentration of the protein in µg ml^{-1}.

5.6.3 Congo red (CR) dye binding assay: CR is similar to Th-T and binds to β-rich structures in a similar manner, inducing an increase in absorption and a red shit in the CR absorption band from 490 to 540 nm [Klunk *et al*, 1999, Chang *et al*, 2009]. The w-t protein (0.1 mg ml^{-1}) showed a spectrum similar to the CR dye with neither increase in absorption nor a shift in the absorption band. However, the mutant protein (0.1 mg ml^{-1}) showed an increase in absorption along with a red shift to 540 nm indicating a propensity for aggregation resulting in β-sheet formed from intermolecular strands (Figure 5.7.3).

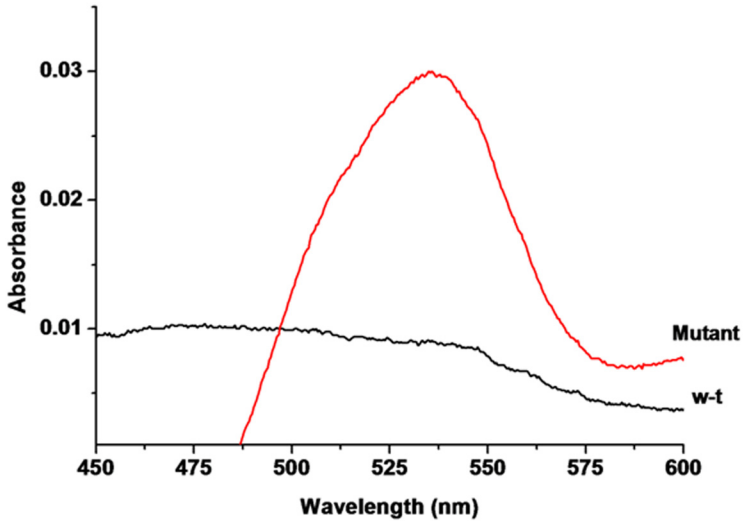


Figure 5.7.3: Congo red binding assay, for w-t (black) and mutant (red) protein at 37°C.

Thus, the three assays performed for checking aggregation indicate that the mutant protein aggregates with a higher propensity than the w-t.

5.7 Stability studies

The effect of varying pH, temperature and concentration of Gdn-HCl on the proteins were studied to assess the stability differences between w-t and mutant (T145I+R199W) proteins.

5.7.1 pH dependence

5.7.1.1: *Intrinsic fluorescence:* Under highly protonated or deprotonated conditions of *NDUFS3* a decrease in the intrinsic fluorescence intensities was observed in the w-t protein as against the mutant, which showed a significant decrease only in protonated conditions (Figure 5.8.1A and B). Also the positive environment of surface tryptophans was observed to be higher in mutant during quenching studies.

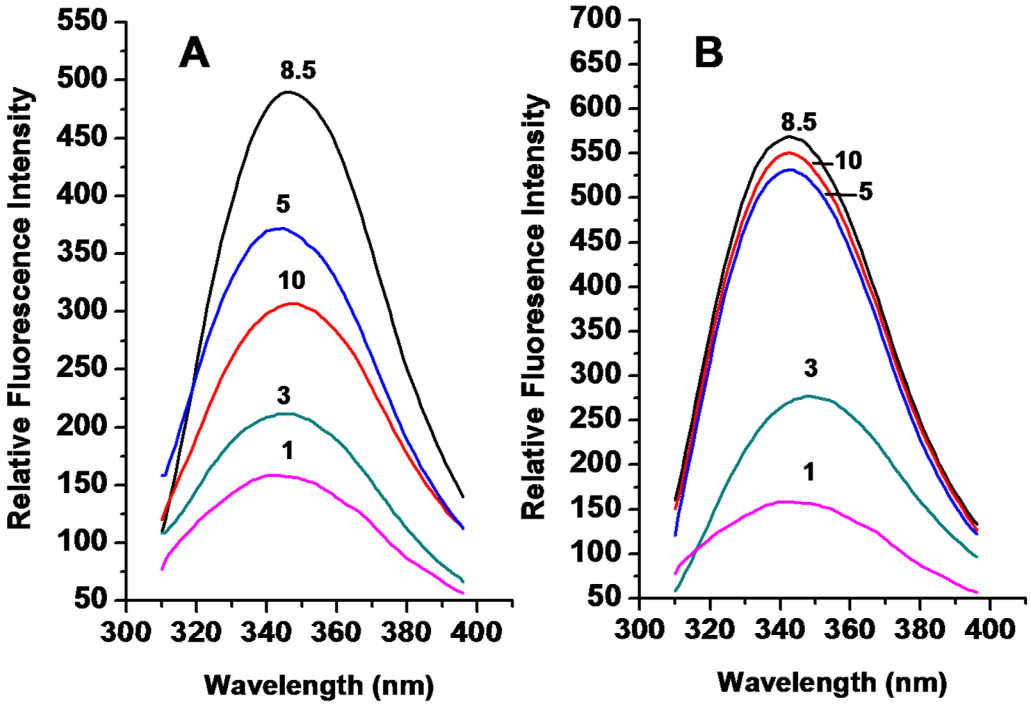


Figure 5.8.1: Intrinsic fluorescence intensity changes at different pH. A: w-t, B: T145I+R199W mutant proteins. The numbers on the spectra indicate the pH values at which the spectra were recorded.

5.7.1.2 *Extrinsic fluorescence:* The extrinsic fluorescence by ANS binding to both w-t and mutant increased significantly at pH 1, 2 and 3 indicating an exposure of the hydrophobic patches at extreme acidic pH (Figure 5.8.2A and B).

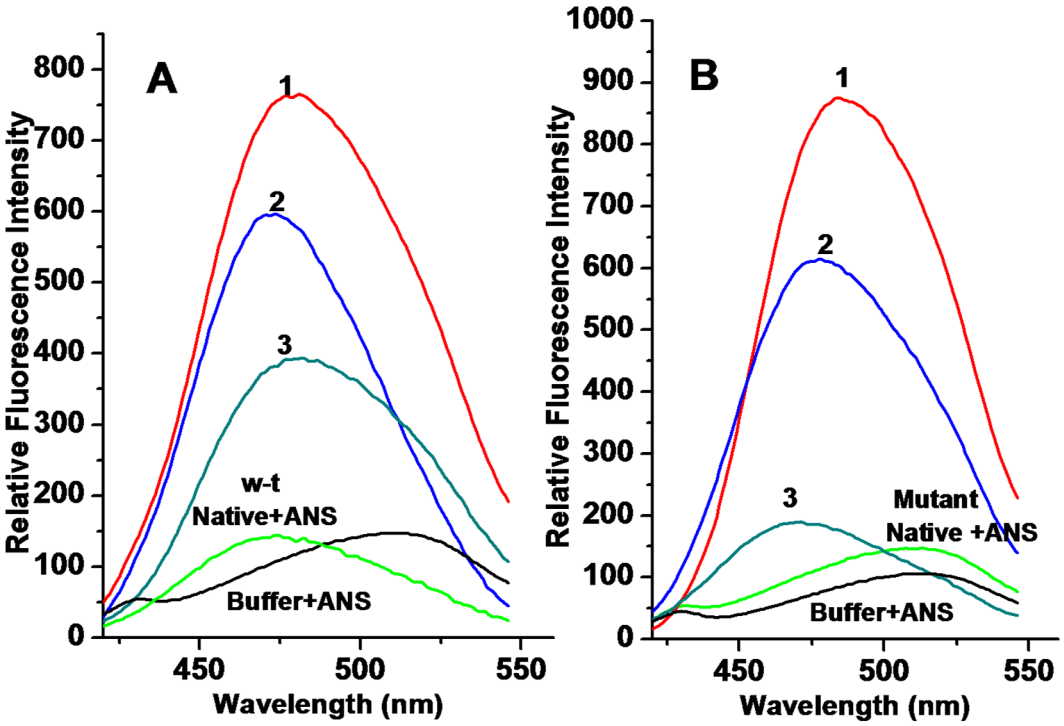


Figure 5.8.2: ANS binding spectra at different pH, A: w-t B: mutant (T145I+R199W) proteins. The numbers on the spectra indicate the pH values.

5.7.1.3 *CD analysis*: At pH 8.5 both the w-t and mutant proteins showed primarily α-helical structure with a component of β-sheets as indicated by minima at 208 and 222 nm and by the estimated values using CDPro CDSSTR. As the pH was gradually reduced from neutral to acidic (pH7 to 3) there was a decrease in ellipticity in the case of w-t which was not observed in mutant protein. Also, a gradual loss of α-helix content and an increase in β-sheets was observed with decrease in pH from 7 to 3 or with an increase in pH from 10 to 12 in the case of both w-t and mutant proteins (Table 5.5)

Table 5.5: Secondary structure composition of w-t and mutant (T145I+R199W) proteins at varying pH. All the calculations were done in CDSSTR of the CDPro suite.

pH	% Helix	% Sheet	% Turn	% Unord	NRMSD	% Helix	% Sheet	% Turn	% Unord	NRMSD
	w-t protein					Mutant (T145I+R199W) protein				
1	39.6	19.9	15.2	25.3	0.053	8.1	32.3	24.3	35.3	0.057
2	14.7	31.1	24.1	30.1	0.063	20.9	27.8	21.7	29.6	0.062
3	15.5	32.2	22.2	30.1	0.100	22.6	25.1	21.7	30.6	0.044
4	20.1	28.0	23.1	28.8	0.067	20.2	27.5	22.6	29.1	0.056
5	15.4	31.8	23.1	29.2	0.087	20.4	28.4	22.6	28.7	0.029
6	24.2	36.0	20.5	28.6	0.071	20.6	28.6	24.4	28.3	0.080
7	48.1	18.3	16.0	17.4	0.095	21.0	28.6	21.7	28.3	0.049
8	56	10.2	10.5	23.3	0.049	50.3	10.3	17.1	22.3	0.043
9	43.6	29.9	11.9	15.4	0.033	22.3	34.4	21.7	21.8	0.061
10	18.6	33.5	19.5	28.1	0.062	22.3	34.4	20	23.5	0.021
11	17.8	19.0	22.8	40.4	0.036	4.8	36.9	26.7	30.8	0.087
12	22.1	33.3	22.9	21.5	0.066	4.5	37.2	26.7	30.8	0.021

The w-t protein structure opened up at pH 2 and interestingly, refolded at pH 1 assuming a native like secondary structure. However, this was not observed in the mutant protein where the structure was stable at pH 2 and opened up at pH 1 with an increase in the unordered structure (Figure 5.8.3A and B). Thus, they behaved differently at extreme low pH.

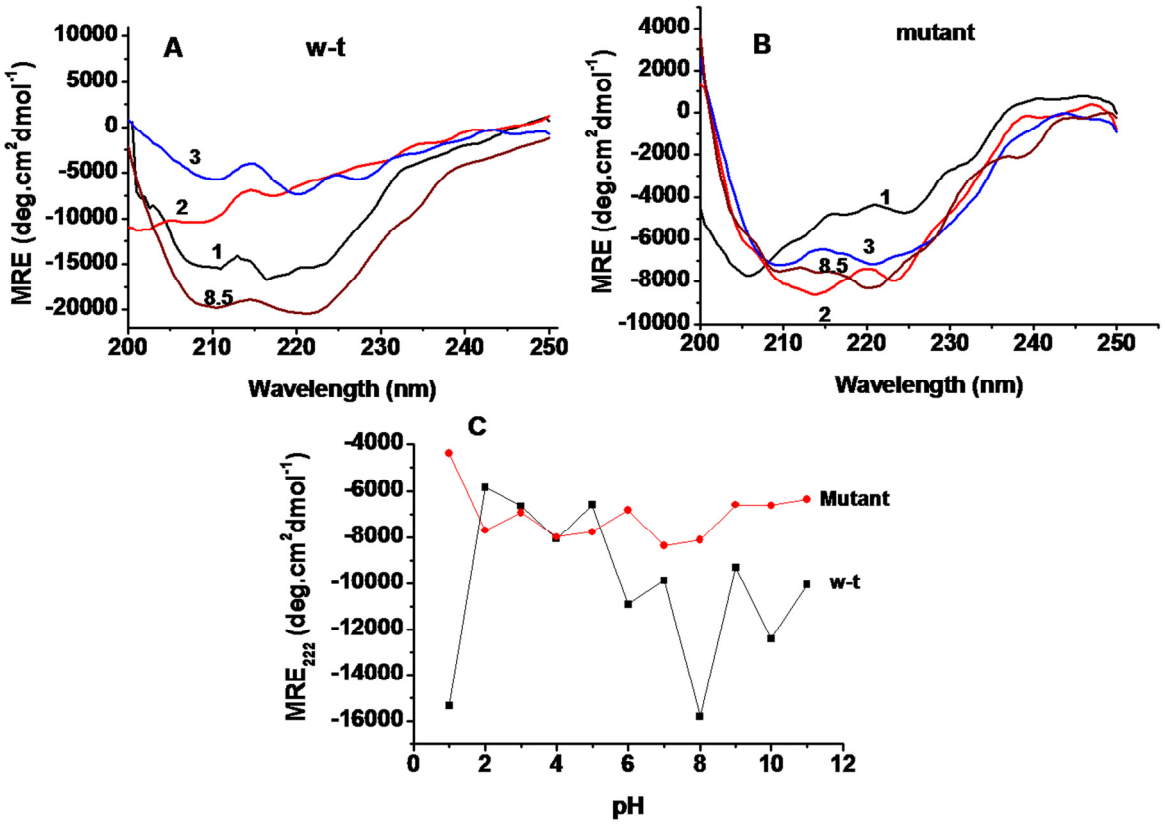


Figure 5.8.3: Far UV MRE spectra of the A: w-t and B: mutant protein at different pH, C: plot of MRE_{222} of w-t and mutant protein at different pH. The numbers on the spectra indicate the pH values.

5.7.2 Thermal unfolding

5.7.2.1 *Intrinsic fluorescence and scattering*: A gradual shift in λ_{max} accompanied by decrease in relative fluorescence intensity was observed as the temperature of the w-t or mutant protein was increased from 25° to 85°C (Figure 5.9.1.1).

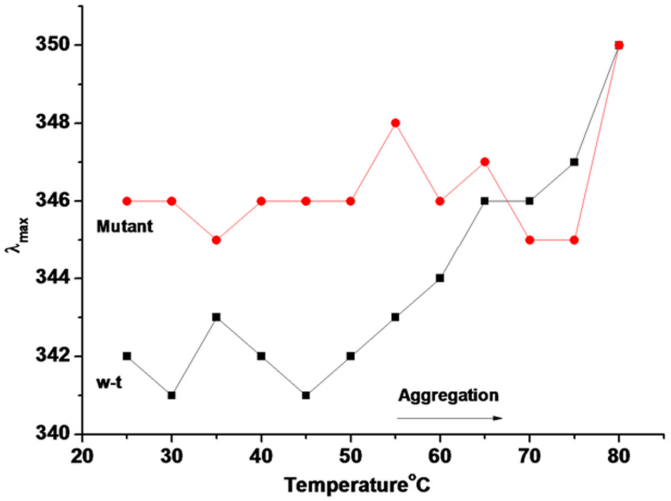


Figure 5.9.1.1: Thermal unfolding of w-t and mutant *NDUFS3* protein, plot of λ_{max} vs temperature in °C, w-t (black) and mutant (red).

Both the w-t and mutant *NDUFS3* proteins showed tendency to aggregate (Figure 5.9.1.2) at 55°C as observed in scattering measurements. The aggregation process was directly proportional to the protein concentration and inversely proportional to the concentration of salt (NaCl) and detergent (IGEPAL®).

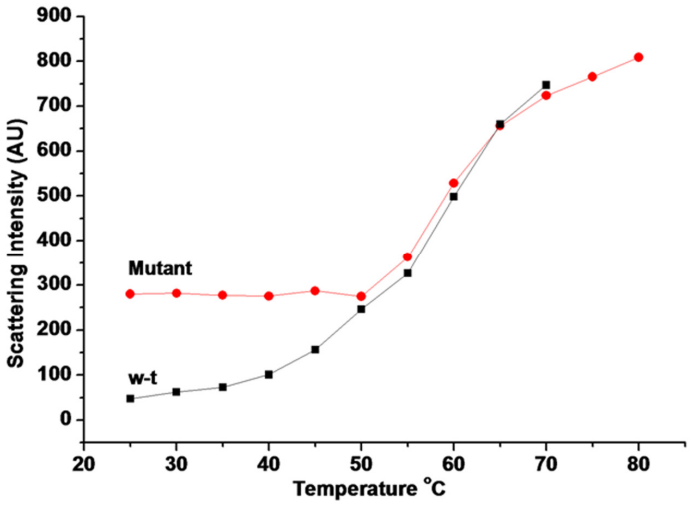


Figure 5.9.1.2 Thermal unfolding of w-t and mutant *NDUFS3* protein, plot of Rayleigh scattering intensity vs temperature, w-t (black) and mutant (red).

5.7.2.2 *Extrinsic fluorescence*: No ANS binding was observed on heating the protein at all temperatures, indicating the aggregation to be a fast process.

5.7.2.3 *CD analysis:* Till 50°C, rapid structural rearrangements of α-helices and β-sheets were observed in far UV CD spectra of the w-t protein (Table 5.6, Figure 5.9.2A, B and C). Above 50°C, simultaneous alteration of protein fold into a β-sheet like structure was observed, which lead to aggregation along with an increase in the content of loops and unordered structure. Both proteins remained α-helix dominated structures till 50°C above which they aggregated and altered to a β-sheet structure.

Table 5.6: Secondary structure composition of w-t and mutant (T145I+R199W) proteins with increasing temperature (^{0}C). All the calculations were done in CDSSTR of the CDPro suite.

Temp ^{0}C	% Helix	% Sheet	% Turn	% Unord	NRMSD	% Helix	% Sheet	% Turn	% Unord	NRMSD
	w-t protein					Mutant (T145I+R199W) protein				
25	56	10.2	10.5	23.3	0.049	50.3	10.3	17.1	22.3	0.043
30	33.3	18.3	19.3	29.1	0.057	38.6	16.9	18.6	25.9	0.037
35	30.3	21.8	19.3	29.1	0.091	39.5	14.8	18.7	26.4	0.059
40	38.4	20.8	17.9	22.9	0.068	43.1	15.8	18.4	22.7	0.055
45	33.9	12.0	18.9	28.3	0.075	41.4	17.1	19.9	21.3	0.048
50	38.1	17.4	18.7	25.8	0.058	38.2	17.5	20.1	24.2	0.057
55	17.2	31.9	21.7	21.9	0.067	34.8	18.8	23.1	23.2	0.070
60	21.1	37.2	21.7	20.0	0.059	18.9	28.1	23.0	30.0	0.072
65	22.5	27.1	24.7	26.4	0.060	15.4	34.2	25.4	24.2	0.093
70	17.4	32.1	24.5	26.0	0.028	6.2	36.6	25.8	30.0	0.097
75	4.3	37.0	30.4	28.3	0.084	4.2	32.5	27.7	34.6	0.104
80	7.3	36.8	27.4	27.7	0.081	3.4	42.0	22.0	33.9	0.112
85	6.0	32.5	29.3	32.6	0.122	4.6	42.7	28.0	24.1	0.114

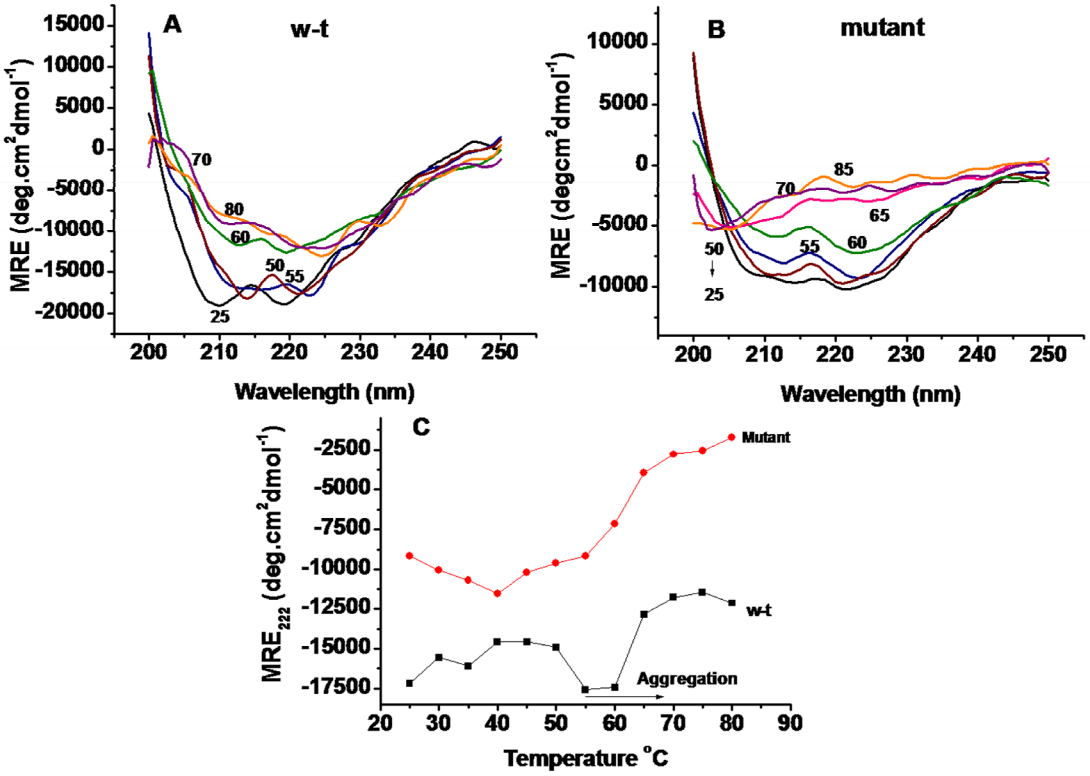


Figure 5.9.2: Far UV MRE spectra of the A: w-t and B: mutant protein at different temperatures (^{0}C), C: plot of MRE_{222} vs temperature for the wt versus mutant protein. The numbers on the spectra indicate the temperature.

5.7.3 Gdn-HCl induced unfolding

5.7.3.1 *Intrinsic fluorescence:* With increasing concentrations of Gdn-HCl a decrease in the intrinsic fluorescence intensity was observed for both w-t and mutant protein. The λ_{max} 342 nm (w-t) and 346 nm (mutant) showed a red shift to 352 nm close to concentration 2M of Gdn-HCl indicating a loss of structure. Parameter A, which is the ratio of intrinsic fluorescence intensity at 320 nm to that at 365 nm, the characteristic of the shape and position of the spectrum, measures the sensitivity of w-t and mutant protein in different Gdn-HCl concentrations [Marthino *et al*, 2003; He *et al*, 2005 ;Su *et al*, 2007]. Parameter A analysis and λ_{max} changes at different concentrations of Gdn-HCl for the w-t and mutant proteins indicate that the unfolding caused by Gdn-HCl was a multi-step process for both. The w-t (A, II) and mutant protein (S, II) tryptophan classes slowly shift to Class III type on complete unfolding (Figure 5.10.1).

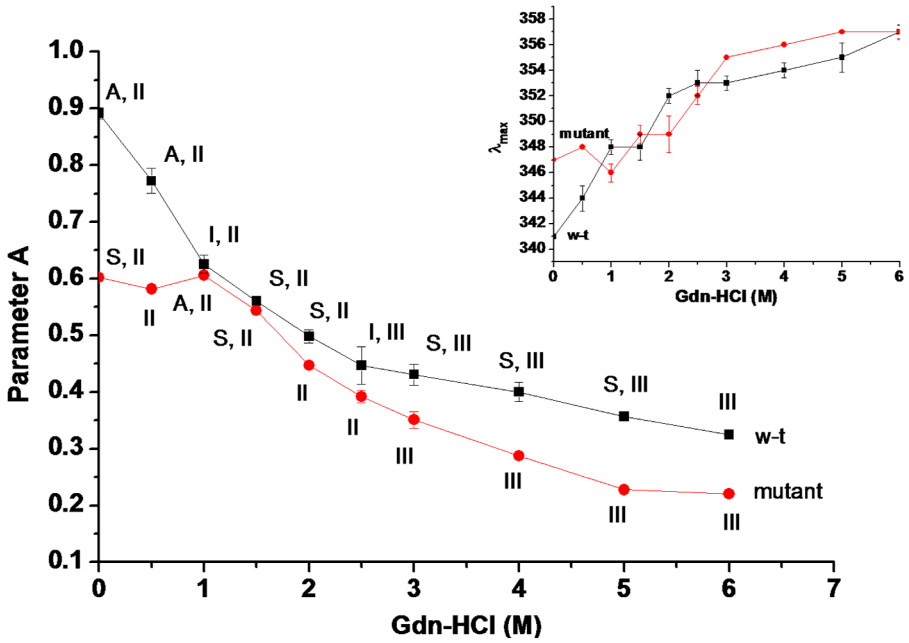


Figure 5.10.1: Parameter A analysis for the Gdn-HCl induced unfolding of w-t and mutant (T145I+R199W) proteins, showing the tryptophan conformer classes, inset shows the graph for λ_{max} vs Gdn-HCl concentrations [M].

5.7.3.2 *Extrinsic fluorescence*: Significant ANS binding with w-t *NDUFS3* was observed in the presence of low concentrations of Gdn-HCl (<0.5M) indicating exposure of hydrophobic surface patches at the beginning of unfolding pathway, however, no ANS binding was observed in mutant protein under any condition (Figure 5.10.2).

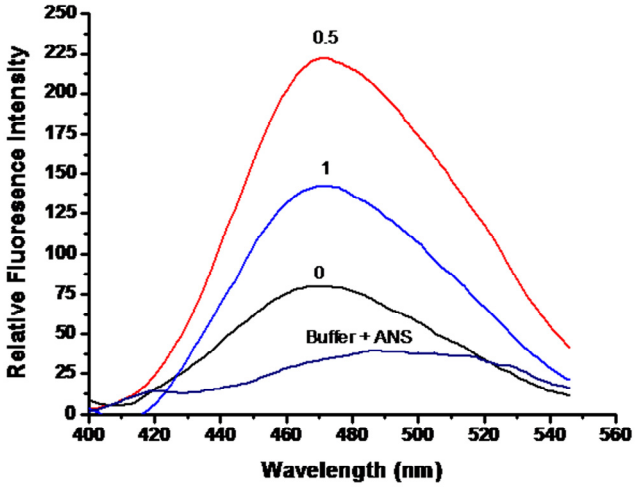


Figure 5.10.2: ANS binding spectra for w-t *NDUFS3* at different Gdn-HCl concentrations. The numbers on the spectra indicate the different concentrations of Gdn-HCl [M].

5.7.3.3 *CD analysis*: Far UV CD analysis (Table 5.7) showed the complete unfolding of the w-t and mutant protein in the presence of 2M Gdn-HCl correlating well with the fluorescence data. The MRE_{222} of the w-t and mutant proteins are distinctly different in the absence of Gdn-HCl, showing reduced ellipticity for mutant, while in the presence of Gdn-HCl, they follow a similar pattern (Figure 5.10.3) reflecting the different native structures but fairly similar sensitivity to Gdn-HCl mediated denaturation.

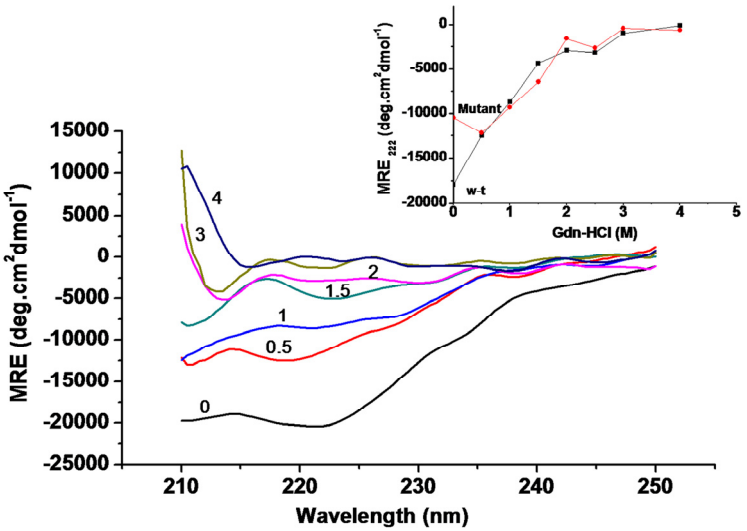


Figure 5.10.3: Far UV MRE spectra of the w-t *NDUFS3* protein, inset shows a graph of MRE_{222} vs Gdn-HCl concentrations for the w-t (black) and mutant (red) protein. The numbers on the spectra indicate the Gdn-HCl concentration.

Table 5.7: Secondary structure composition of w-t and mutant (T145I+R199W) proteins with increasing Gdn-HCl concentration [M]. All the calculations were done in CDSSTR of the CDPro suite.

Gdn-HCl [M]	% Helix	% Sheet	% Turn	% Unord	NRMSD	% Helix	% Sheet	% Turn	% Unord	NRMSD
	w-t protein					Mutant (T145I+R199W) protein				
0	56	10.2	10.5	23.3	0.049	50.3	10.3	17.1	22.3	0.043
0.5	43.8	22.8	14.1	19.4	0.024	45.4	14.9	16.4	22.5	0.015
1	41.5	11.8	16.0	29.5	0.061	41.5	20.3	14.5	23.7	0.024
1.5	24.3	30.5	18.5	26.6	0.089	27.3	28.5	17.5	26.6	0.041
2	6.2	48.8	22.7	22.4	0.237	5.0	50.0	22.7	22.4	0.145

5.8 Conclusion

This study highlights the differences between the wild-type protein and its Leigh syndrome causing mutant (T145I+R199W) protein. Although the expression of the defective protein is at normal levels in patients, it is the accumulation of assembly intermediates in patient cell lines that causes the disease, which highlights the need to look at this problem from a structural point of view. Biophysical studies of the proteins in solution have been utilized to understand the differences in their respective stability. The study showed that the behaviour and stability of w-t and mutant was almost same, however, the native states of the proteins were significantly different. This suggests the reasons for the lack of proper assembly of the mutant protein in the Complex-I. The data and the analysis presented here highlights, from a structural point of view, the differences in the behaviour of w-t and mutant of *NDUFS3* protein and demonstrates that the misfolding of mutant with the resultant aggregation is the reason for the aberrant behaviour of Leigh syndrome mutant during Complex-I assembly formation.

NDUFS3 protein is energetically unstable [Jaokar *et al*, 2013]. Experimental results presented here demonstrated the higher aggregation propensity of mutant protein and recognized its several modified structural and folding properties compared to w-t. These mutations would thus, affect the assembly of a viable Complex-I at the early stages itself. Lack of a properly assembled Q module would prevent the assembly of the N module, although the P module would get assembled independently. Lack of a fully functional complex not only creates energy deficiencies in the cells but also causes reactive oxygen species (ROS)-mediated cell damage; together they present a broad spectrum of symptoms in Leigh syndrome. The symptoms of dystonia, motor dysfunction, optic nerve atrophy etc. [Benit *et al*, 2004] are all indicative of muscle and nerve damage due to a non functional Complex-I and ROS-mediated damage.

CHAPTER 6

Characterization and binding studies of *NDUFS7* subunit and its mutants that cause Leigh syndrome

NADH dehydrogenase ubiquinone iron sulphur protein 7; *NDUFS7* is one of the core subunits of the ubiquinone reduction (Q) module [Fernández-Vizarra *et al*, 2009]. It is one of the 8 core subunits of Complex-I encoded by the nuclear genome (n-DNA). The *NDUFS7* gene (Entrez gene 374291) is located on the chromosome 19p13.3 [Hyslop *et al*, 1996] and is expressed in all tissues. This gene is composed of 15 distinct introns. The protein (UniProtKB O75251) encoded by the gene comprises of 213 amino acids and has a predicted signal sequence of 52 amino acids (Table 2.1). It is also known as PSST subunit in bacteria or 20 kDa *NDUFS7* subunit in mammals. The study by *in silico* methods on the structural role of w-t *NDUFS7* protein and the changes in the Q module due to pathogenic mutations were extensively discussed in Chapter 3. This chapter highlights the differences in the w-t and mutant *NDUFS7* protein behaviour *in vitro* or in the solution.

6.1 *NDUFS7* and its function

NDUFS7 is a functional subunit of the ubiquinone reduction domain or Q module. It possesses the iron sulphur cluster N2 [Hinchliffe *et al*, 2005]. The cluster N2 is the ultimate electron donor to ubiquinone for reduction [Tocilescu *et al*, 2010; Angerer *et al*, 2012]. The ubiquinone binding site is located at the interface of the two subunits; *NDUFS2* and *7* [Hunte *et al*, 2010]. Ubiquinone plays a dual role by acting as a substrate by accepting electrons from iron sulphur cluster N2 and as a tightly bound cofactor during recycling [Prieur *et al*, 2001]. Apart from catalytic function, *NDUFS7* or PSST subunit also plays a pivotal role in the biogenesis of the Complex-I. Two models highlight its role in the assembly of Complex-I. The first one involves the binding of *NDUFS7* to *NDUFB6, ND1* and *ND6* to initiate assembly of the P module [Ugalde *et al*, 2004b]. This partially assembled P module then joins the peripheral arm to form a fully assembled Complex-I. The second model states that *NDUFS7, NDUFS4* and *NDUFV2* assemble together prior to associating with other subunits [Antonicka *et al*, 2003]. In *Neurospora carassa*, PSST is involved in the assembly of the peripheral arm, highlighting its crucial role in biogenesis of Complex-I [Duarte *et al*, 2002].

6.2 *NDUFS7* and its mutants

The *NDUFS7* transcripts are present in all human tissues indicating their expression to be like housekeeping genes. It is the fourth most mutated nuclear gene in patients with Complex-I deficiency [Lebon *et al*, 2007a]. Tissues like muscles and nerves having high metabolic rates have higher expression of *NDUFS7* transcripts thereby making them most susceptible to energy deficit in case of mutations. Two point mutations, V122M and R145H in the *NDUFS7* gene and a mutation in one of the introns of *NDUFS7* leading to a cryptic exon are reported to cause Leigh syndrome [Lebon *et al* 2007a;b; Triepels *et al*, 1999]. The V122 and R145 are highly conserved positions in *NDUFS7* protein (Figure 6.1) during evolution.

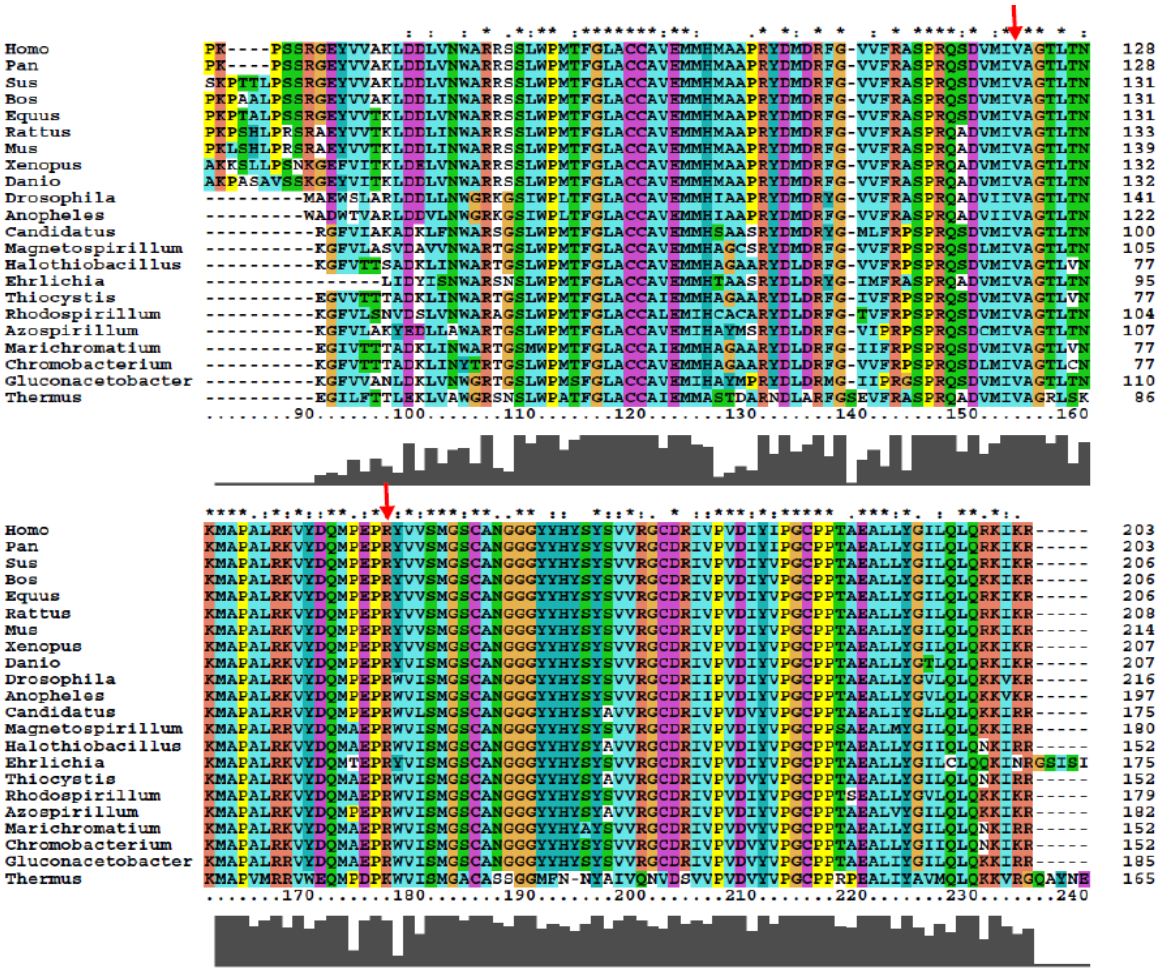


Figure 6.1: Multiple sequence alignment of the *NDUFS7* protein from bacteria to mammals. The red arrow highlights the conserved positions mutated in Leigh syndrome.

The V122M mutation is present in the cysteine motif -CxxE-$(x)_{60}$-C-$(x)_{30}$-CP- involved in the formation of the iron sulphur cluster N2. However, the R145H mutation is relatively away from the binding site of ubiquinone as well as from the iron sulphur N2 site (Figure 6.2). Thus, the reasons for these point mutations to show such disastrous consequences are relatively unknown.

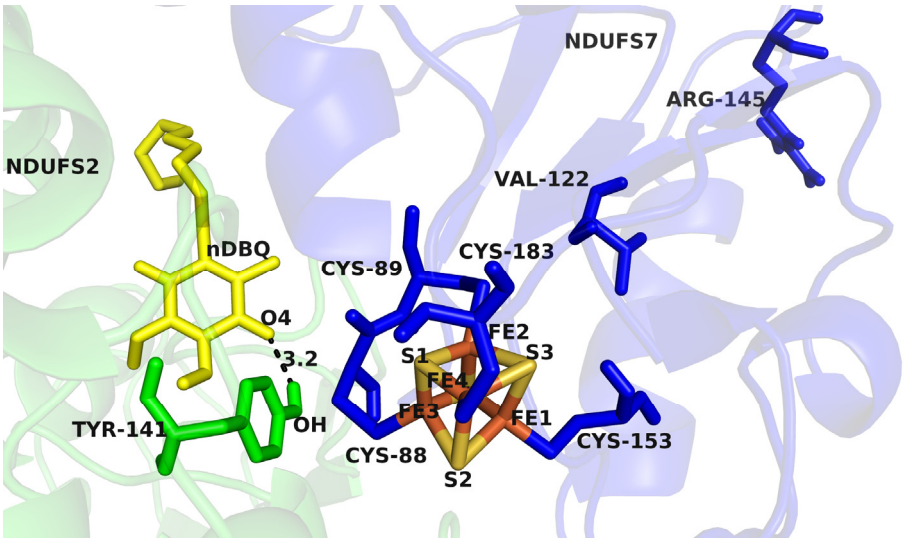


Figure 6.2: Positions of the amino acids mutated (V122 and R145) with respect to the Fe-S cluster N2 and n-DBQ (yellow).

In this chapter, comparative biophysical analysis was carried out to understand the differences between the w-t and mutant proteins in order to get further insights into the loss of function and disassembly of Complex-I. Attempts have been made to understand the broad spectrum of symptoms in the disease conditions and to correlate them with the effects of mutants on the structure and functioning of Complex-I.

6.3 Comparison of fluorescence spectra and secondary/tertiary structure of w-t and mutants

W-t *NDUFS7* and mutant proteins have three tryptophans belonging to the class I, corresponding to buried tryptophans. The intrinsic fluorescence maximum (λ_{max}) for w-t and mutant proteins was at 341 nm (Figure 6.3.1) indicating that the tryptophans were present in a buried environment.

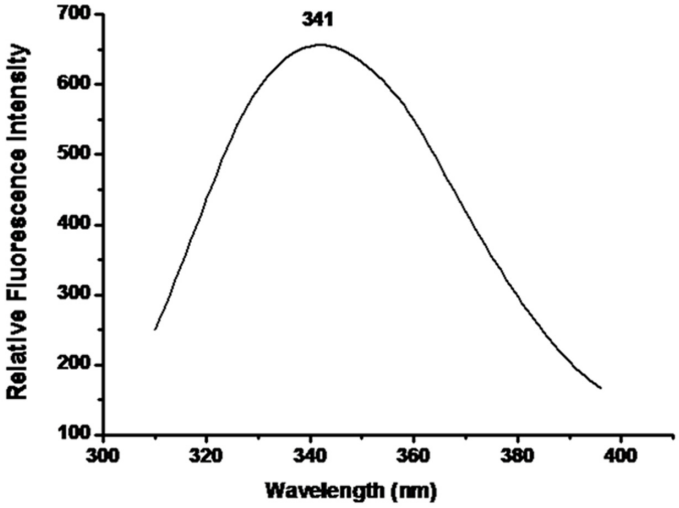


Figure 6.3.1 Steady state intrinsic fluorescence spectrum of *NDUFS7*: w-t, V122M and R145H mutants. The spectra overlap each other and λ_{max} for all the 3 proteins is 341 nm.

However, a variation in the secondary and tertiary structures was observed (Figure 6.3.2 and 6.3.3). The secondary structures of w-t as well as the mutants were primarily β-sheet (Table 6.1).

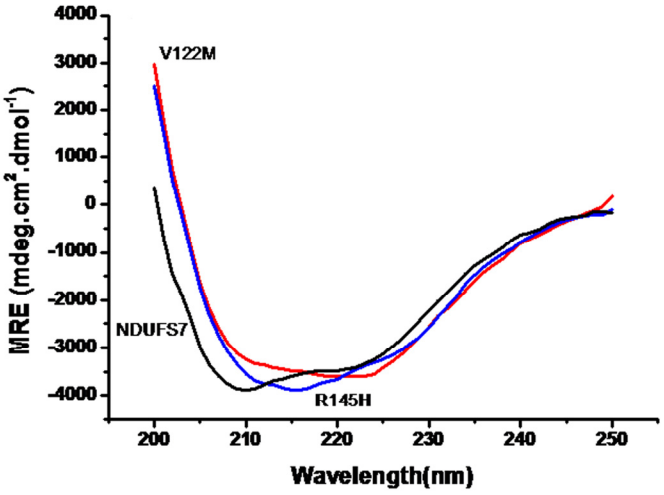


Figure 6.3.2: Far UV CD spectra of w-t (black) *NDUFS7* and its mutants V122M (red) and R145H (blue).

In the tertiary structure, a change was observed in the tryptophan and tyrosine regions for the mutants as compared to w-t (Figure 6.3.3).

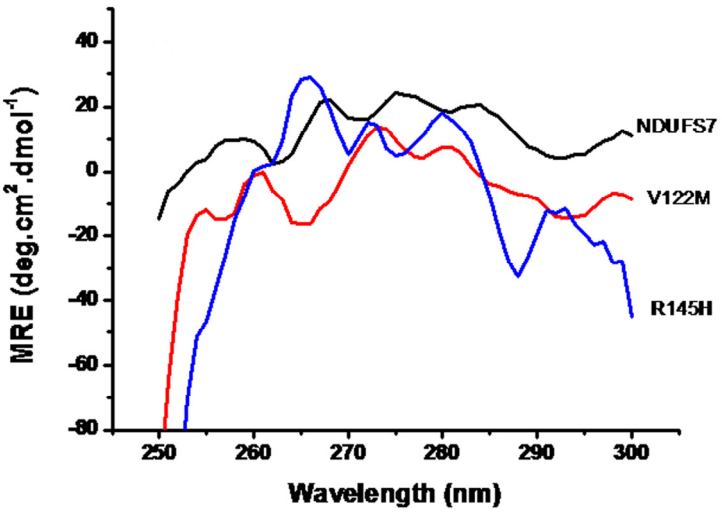


Figure 6.3.3: Near UV CD spectra of w-t (black) *NDUFS7* and its V122M (red), R145H (blue) mutants.

Table 6.1: Secondary structure composition of NDUFS7 w-t and mutant (V122M and R145H) proteins. All the calculations were done in CONTINLL of the CDPro suite. Since the full sequence of NDUFS7 could not be modelled due to absence of desired template the values shown for model are only indicative and not actual.

	% Helix	% Sheet	% Turn	% Unord	NRMSD
w-t	9.3	34.7	21.9	31.4	0.03
V122M	11	32.8	21.0	36.2	0.02
R145H	8.7	39.3	22.0	30.0	0.08
Partial model *NDUFS7*	12.0	40.0	30.0	18.0	NA

6.4 Steady state fluorescence quenching

The rate of quenching K_{sv} calculated from the Stern-Volmer plots was higher for the denatured w-t protein. In the w-t protein 75% of the tryptophan fluorescence was accessible to acrylamide which became fully accessible on complete denaturation. A high positive charge density is observed around the surface tryptophans in the w-t protein. On denaturation the negative charge density increased although the accessibility remained more or less same (Figure 6.4 and Table 6.2).

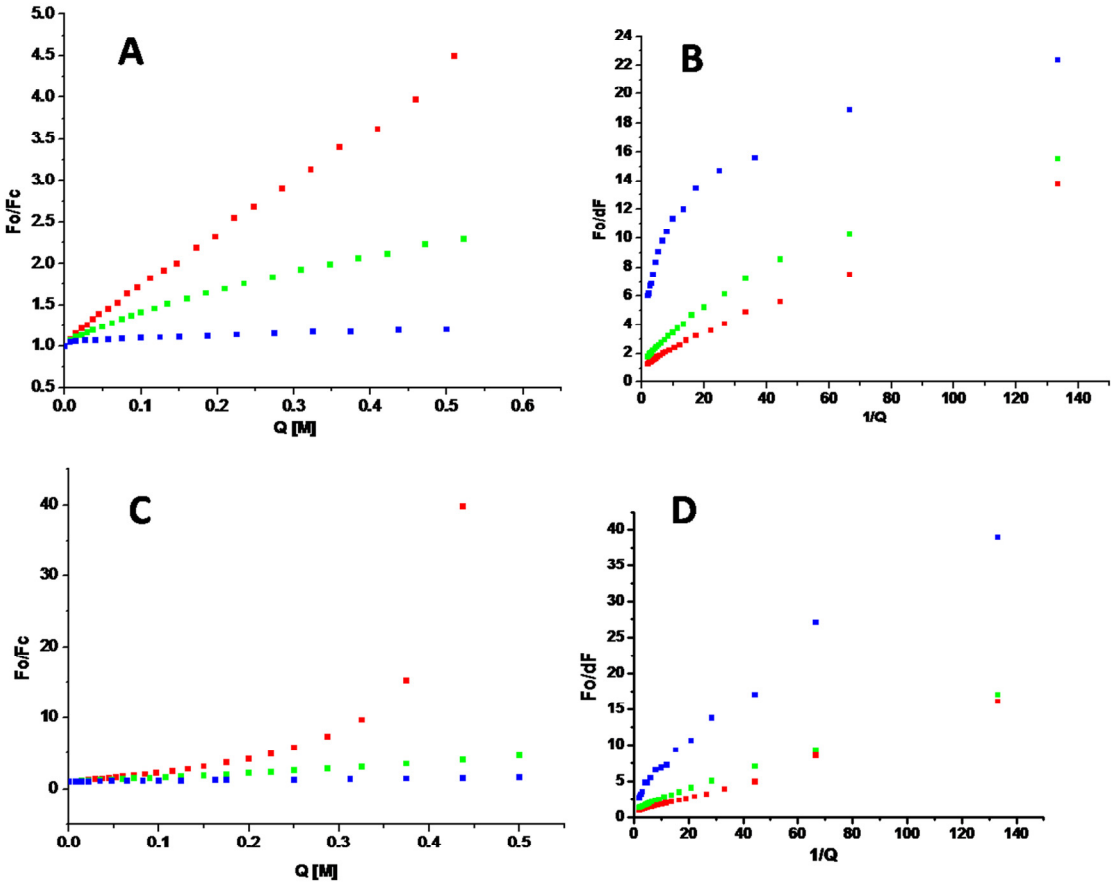


Figure 6.4A: Quenching of the intrinsic fluorescence of w-t *NDUFS7* native A: Stern-Volmer plot, B: Modified Stern-Volmer plot, denatured (in 6M Gdn-HCl), C: Stern-Volmer plot, D: Modified Stern-Volmer plot. Acrylamide (red), KI (green), CsCl (blue).

Table 6.2: Quenching of tryptophan fluorescence of w-t and that of corresponding denatured w-t (in 6M Gdn-HCl). The two values of *fa* in quenching by KI and CsCl are fractional accessibility; fa1 and fa2 reflecting the linear fitting of two parts of the curve.

	Acrylamide		KI		CsCl	
	K_{sv} (M^{-1})	fa	K_{sv} (M^{-1})	fa	K_{sv}(M^{-1})	Fa
w-t	6.46	0.75	2.443	0.606	0.26	0.080
				0.280	0.30	0.210
Denatured w-t	--	1.63	6.81	0.775	1.11	0.264

6.5 Conformational transitions

6.5.1 pH dependence

6.5.1.1: *Intrinsic fluorescence:* Under highly protonated or deprotonated conditions of *NDUFS7* a decrease in the intrinsic fluorescence intensities was observed in the w-t protein (Figure 6.5.1.1).

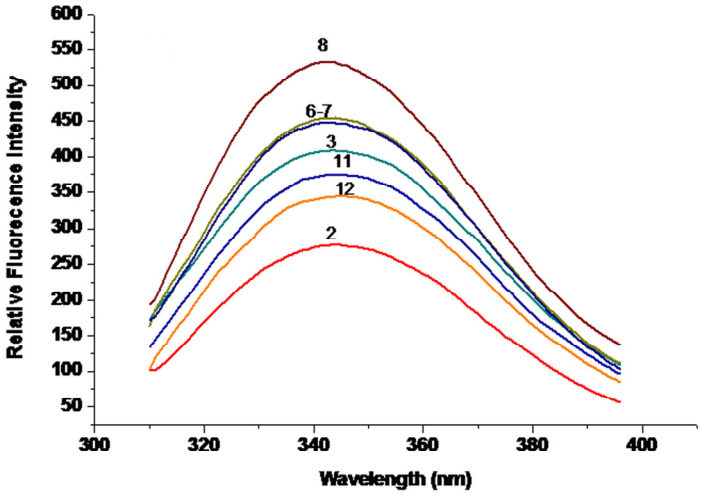


Figure 6.5.1.1: Relative intrinsic fluorescence intensity change of w-t *NDUFS7* at different pH. The numbers on the spectra represent the pH values.

6.5.1.2 *Extrinsic fluorescence:* The extrinsic fluorescence studied by ANS binding in the w-t protein increased significantly at pH 1 and 2 indicating an exposure of the hydrophobic patches at extreme acidic pH (Figure 6.5.1.2).

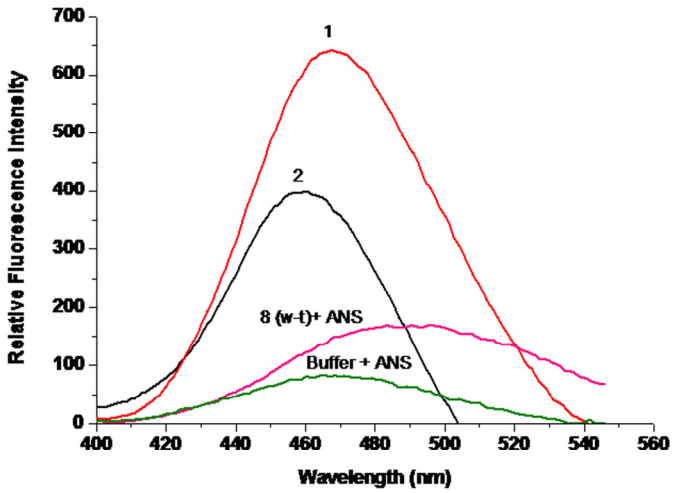


Figure 6.5.1.2: ANS binding spectra of w-t *NDUFS7* at different pH. The numbers on the spectra represent the pH values at which it was recorded.

6.5.1.3 *CD analysis*: The structure of w-t *NDUFS7* was sensitive to changes in pH. Although under the range of pH studied (1-12), the structure was dominated by β-

sheets (Table 6.3) the changes happening to this structure were evident from the changes in the far UV CD spectra (Figure 6.5.1.3 A and B).

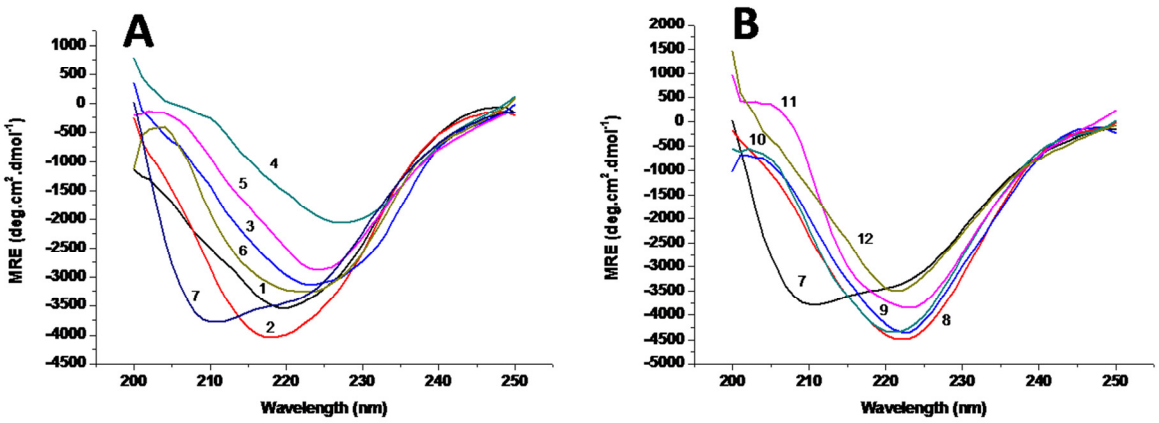


Figure 6.5.1.3: Far UV CD spectra of w-t *NDUFS7* at A: acidic pH and B: Basic pH. The numbers on the spectra represent the pH values.

Table 6.3: Secondary structure composition of *NDUFS7* protein at varying pH. All the calculations were done in CDSSTR of the CDPro suite.

pH	% Helix	% Sheet	% Turn	% Unord	NRMSD
1	6.0	39.5	22.2	32.3	0.021
2	5.8	41.0	21.8	31.4	0.018
3	5.5	41.3	22.3	30.9	0.054
4	4.5	44.9	21.6	29.0	0.072
5	5.2	41.5	22.0	31.3	0.056
6	5.2	40.7	21.5	32.6	0.038
7	8.8	37.7	22.2	31.3	0.029
8	7.0	38.5	22.6	31.9	0.025
9	6.4	41.2	21.9	30.5	0.030
10	5.7	40.8	21.9	31.6	0.025
11	4.2	42.0	22.4	31.4	0.038
12	5.4	41.7	22.4	30.5	0.047

6.5.2 Temperature effects

6.5.2.1 *Intrinsic fluorescence:* A decrease in the intrinsic florescence intensity was observed with increase in temperature. A blue shift in λ_{max} was observed at 60°C the temperature at which aggregation starts (Figure 6.5.2.1).

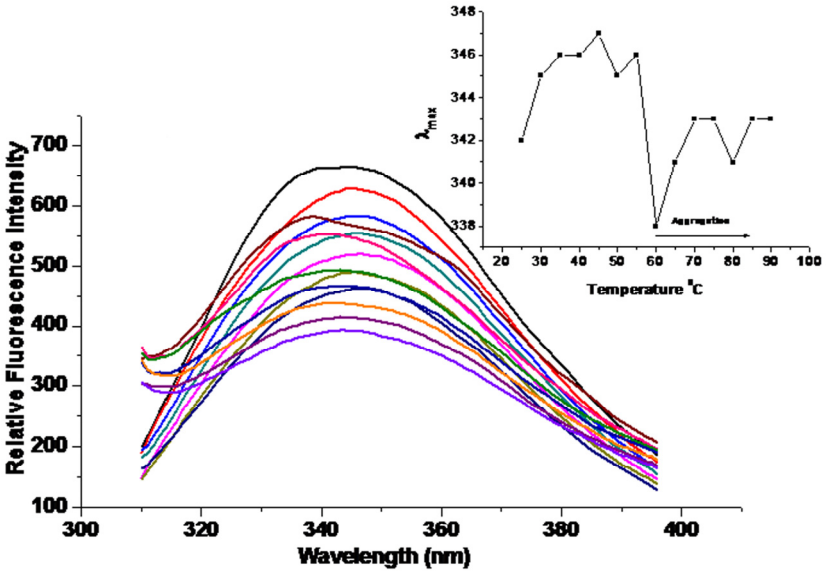


Figure 6.5.2.1: Change in intrinsic fluorescence intensity with increase in temperature. Inset shows a plot of λ_{max} versus temperature.

6.5.2.2 *Extrinsic fluorescence*: No ANS binding was observed at any temperature, even though the aggregation was observed only from 60°C onwards.

6.5.2.3 *CD analysis*: A sudden and drastic change in the CD spectrum was observed above 60°C. Till 70°C the structure remained dominated by β-sheet (Table 6.4) after which there was a drastic decrease in sheets as well as helices with a simultaneous increase in the unordered structure (Figure 6.5.2.2 A and B).

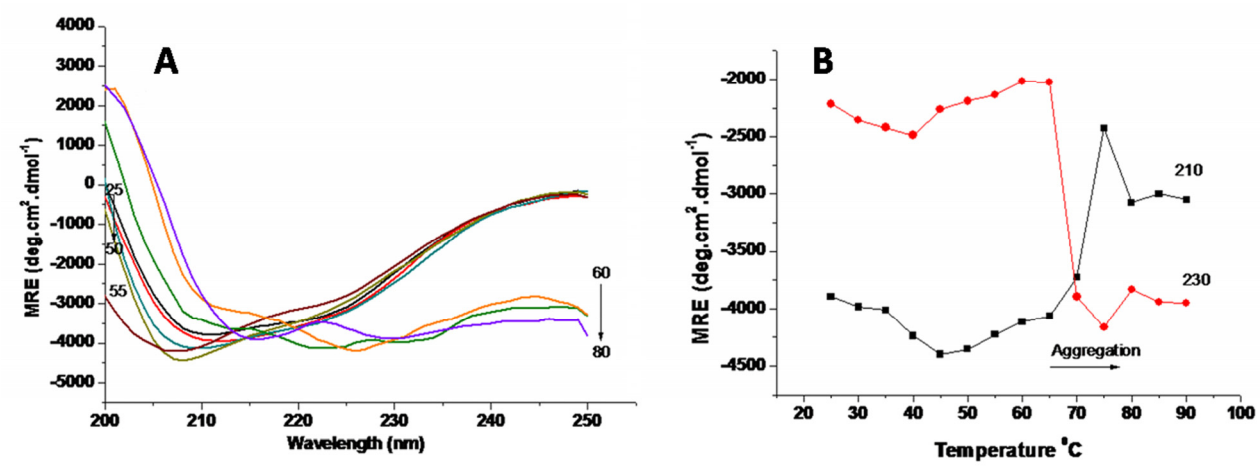


Figure 6.6.2.2: A: Far UV CD spectrum of w-t *NDUFS7* with increasing temperature. B: Plot of MRE versus temperature in ^{0}C; MRE_{210} (black) and MRE_{230} (red).

Table 6.4: Secondary structure composition of *NDUFS7* protein with increasing temperature (^{0}C). All the calculations were done in CDSSTR of the CDPro suite.

Temperature ^{0}C	% Helix	% Sheet	% Turn	% Unord	NRMSD
25	8.8	37.7	22.2	31.3	0.029
30	8.2	39.6	21.2	31.0	0.022
35	9.4	41.2	21.5	27.9	0.027
40	8.5	38.8	21.8	30.9	0.019
45	9.2	37.5	22.3	31.0	0.015
50	8.7	38.9	21.6	30.8	0.014
55	8.7	38.6	22.3	30.4	0.017
60	8.4	36.2	22.5	32.9	0.012
65	11.5	45.1	22.0	32.9	0.012
70	11.5	45.1	22.0	32.9	0.138
75	6.2	16.5	33.4	43.9	0.413
80	4.6	17.3	32.1	46.0	0.447
85	6.0	15.0	33.2	45.8	0.142
90	6.0	15.9	33.2	45.8	0.438

6.5.3 Gdn-HCl induced unfolding

6.5.3.1 *Intrinsic fluorescence*: A decrease in the intrinsic fluorescence intensity and a gradual red shift in the λ_{max} (341 nm) were observed by increasing the concentrations of Gdn-HCl. The class to which tryptophans belonged shifted from class I to II with increasing Gdn-HCl concentrations and finally to class III at 2.5 M Gdn-HCl indicating the continuous increase in exposure of tryptophan residues during unfolding of the structure (Figure 6.5.3.1).

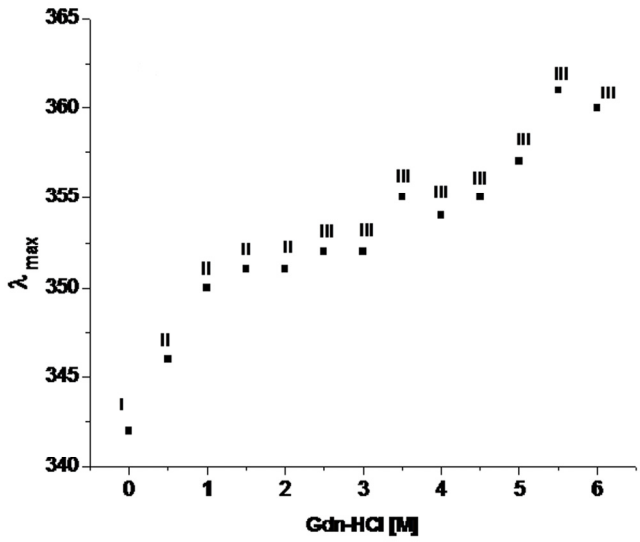


Figure 6.5.3.1: Plot of change in λ_{max} versus Gdn-HCl concentration [M] for the w-t *NDUFS7* protein.

6.5.3.2 *Extrinsic fluorescence*: Interestingly, no ANS binding was observed under any condition.

6.5.3.3 *CD analysis*: The structure gradually opened up, evidenced by the increase in the content of unordered structure, in the presence of increasing concentrations of Gdn-HCl (Table 6.5). A total structural collapse was observed at 2 M Gdn-HCl in the CD spectra though not evident from the CDPro calculations (Table 6.5). The spectra reflected the sensitivity of protein towards the denaturant (Figure 6.5.3.2).

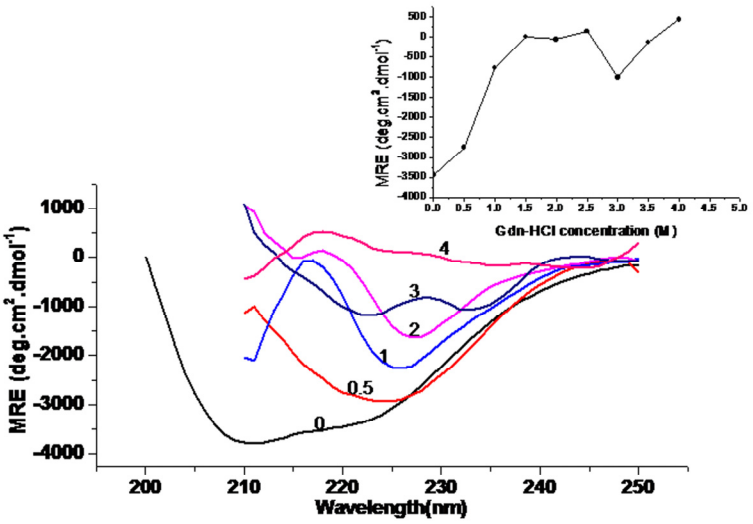


Figure 6.5.3.2: Far UV CD spectrum for the w-t protein incubated at different concentrations of Gdn-HCl. The numbers on the spectra indicate the Gdn-HCl concentration at which it is recorded. Inset is a plot of MRE_{222} versus Gdn-HCl concentration [M].

Table 6.5: Secondary structure composition of *NDUFS7* protein with increasing Gdn-HCl concentration [M]. All the calculations were done in CDSSTR of the CDPro suite.

Gdn-HCl [M]	% Helix	% Sheet	% Turn	% Unord	NRMSD
0	8.8	37.8	22.1	31.3	0.029
0.5	5.1	41.0	22.3	31.6	0.041
1.0	20.9	19.3	24.6	35.2	0.097
1.5	4.5	34.0	22.5	39.0	0.141
2.0	3.0	40.1	18.4	38.5	0.125

The w-t and point mutants, V122M and R145H showed similar λ_{max} and comparable secondary structure with similar quenching and conformational transition parameters.

6.6 Iron sulphur cluster N2

The iron sulphur cluster is an important electron donating group present in the *NDUFS7* subunit. Its role has been implicated in the binding of ubiquinone at the interface of *NDUFS2* and *7*. A UV-visible spectrum from 300-800 nm showed a small peak at 325 nm representing cluster N2 in the w-t protein, which was absent in the case of mutant proteins (Figure 6.6.1). This reflects the destabilizing effect of mutations on the iron sulphur cluster N2. The V122M mutation lies within 4 Å

distance from the cluster in the cysteine motif. However, the R145H mutation is relatively away from the binding site of n-DBQ and cluster N2.

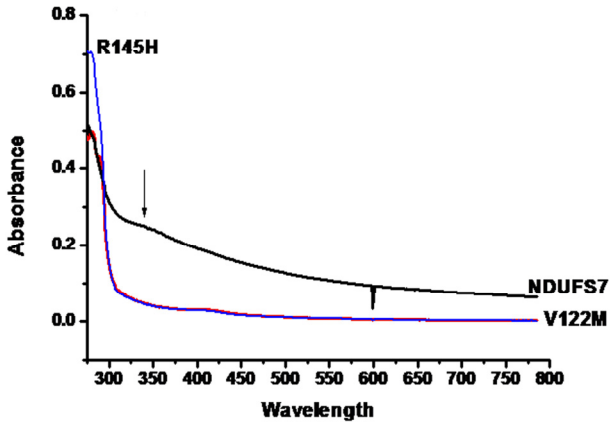
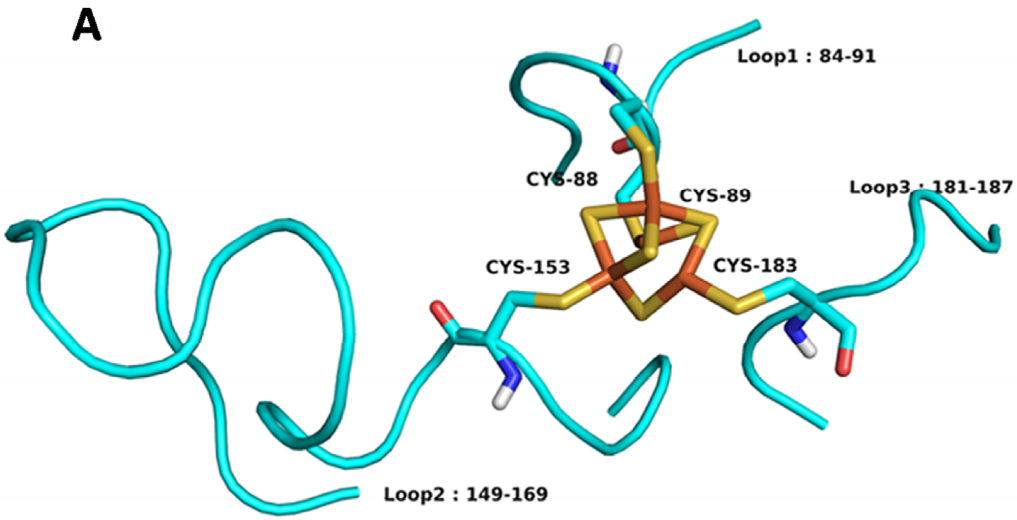


Figure 6.6.1: UV-visible spectrum of the w-t (black), V122M (red)) and R145H (blue) proteins. The arrow indicates the maxima at 325 nm for the w-t *NDUFS7*.

The homology model of the complex *NDUFS2* and 7 built using Prime v3.1 showed that the docked n-DBQ was relatively closer to Fe-S cluster than the sites of mutation (Figure 6.2). The sulphur atoms of 4 cysteines in *NDUFS7* namely C88, 89, 153 and 183 are involved in the formation of the iron sulphur cluster N2. It controls the movement of amino acids present in 3 loops, namely, Loop1: 84-91, Loop2: 149-169 and Loop3: 181-187 (Figure 6.6.2 A and B).



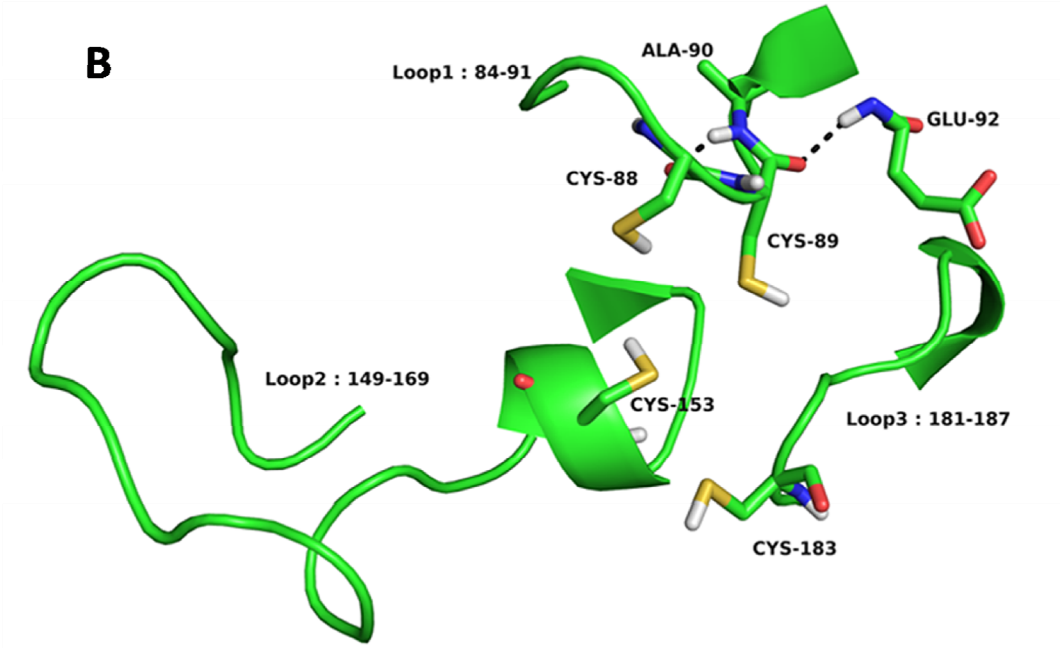
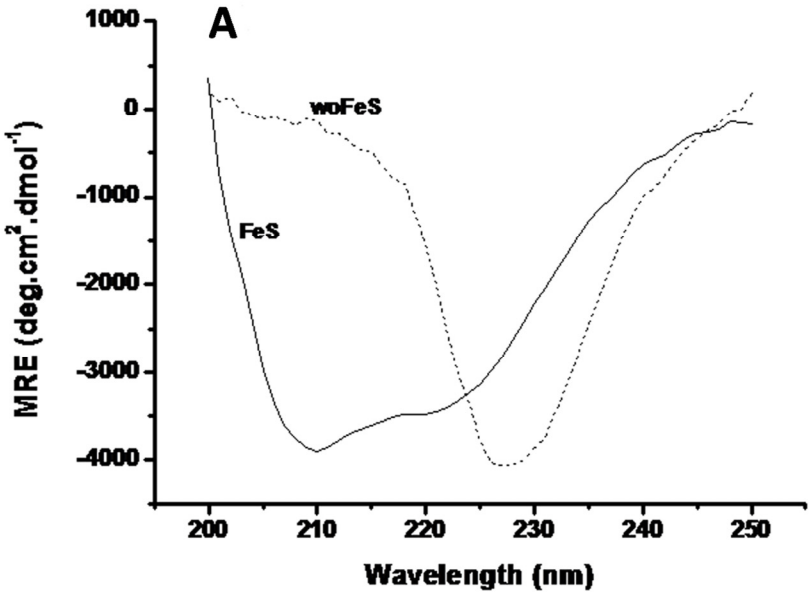


Figure 6.6.2: Three loops A: in presence B: in absence of Fe-S cluster N2.

On removal of the Fe-S cluster by incubating the w-t protein with DTT and EDTA, a drastic change in the secondary and tertiary structure was observed by the change in the CD spectra (Figure 6.6.3 A and B)



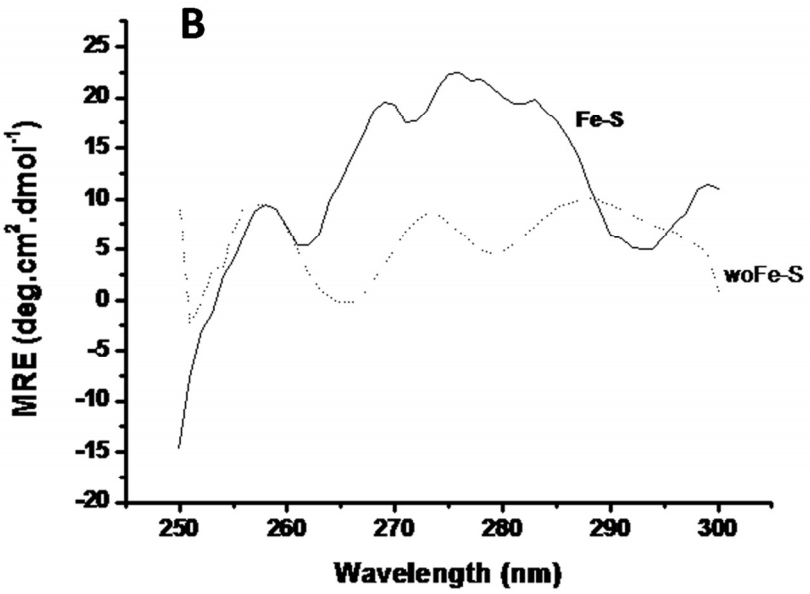


Figure 6.6.3: A: Far UV CD spectrum, B: Near UV CD spectrum of w-t *NDUFS7* in presence (—) and in absence (.....) of Fe-S cluster N2.

Also in the MD simulations, the structure without the Fe-S cluster showed high fluctuation of the above three loops and residues surrounding the binding site (Figure 6.6.4). This instability was reflected in the breakage of the essential hydrogen bond between n-DBQ and Y141 of *NDUFS2* during energy minimization implying the significant role played by N2 in the stable binding of n-DBQ.

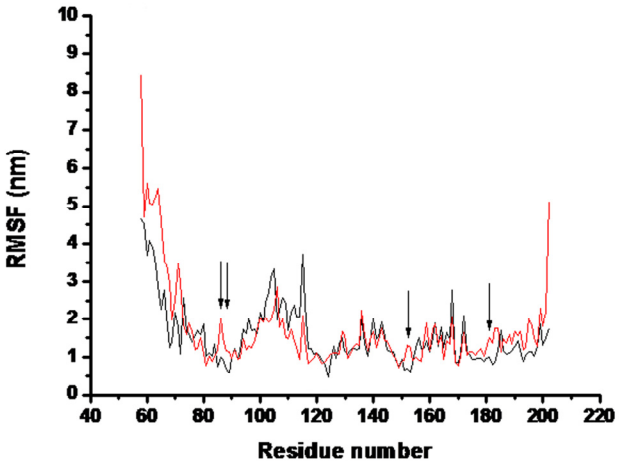


Figure 6.6.4: Plot of RMSF of residues of w-t *NDUFS7* in presence (black) and in absence (red) of the Fe-S cluster N2. The arrows indicate the positions of cysteines forming the cluster.

6.7 Binding of n-DBQ

Recently, Baradaran *et al* [Baradaran *et al* 2013], based on the crystal structure of the Complex-I (3M9S) from *Thermus thermophilus* showed that only a single binding site

at the interface was possible. Since the binding site is a narrow cavity, it causes substantial conformational changes during ubiquinone binding, reduction and simultaneous proton pumping. The ubiquinone binding site is similar to that found in the inhibitor Piericidin A [Baradaran *et al*, 2013; Darrouzet *et al*, 1998]. Although the knowledge about the binding site and functioning of bacterial Complex-I is substantial, not much is known about its eukaryotic counterpart studied from yeast. However, site-directed mutagenesis studies by Angerer *et al* [Angerer *et al*, 2012] in yeast has revealed that in eukaryotes also the ubiquinone binding site is at the interface of *NDUFS2* and *7*. The head carboxyl group of ubiquinone forms a hydrogen bond with the hydroxyl group of a highly conserved tyrosine Y87 in *T. thermophilus*, Y144 in yeast and Y141 in humans (Figure 6.7). It is interesting to note that a methionine rich region is present through which the highly hydrophobic tail of ubiquinone passes [Angerer *et al*, 2012].

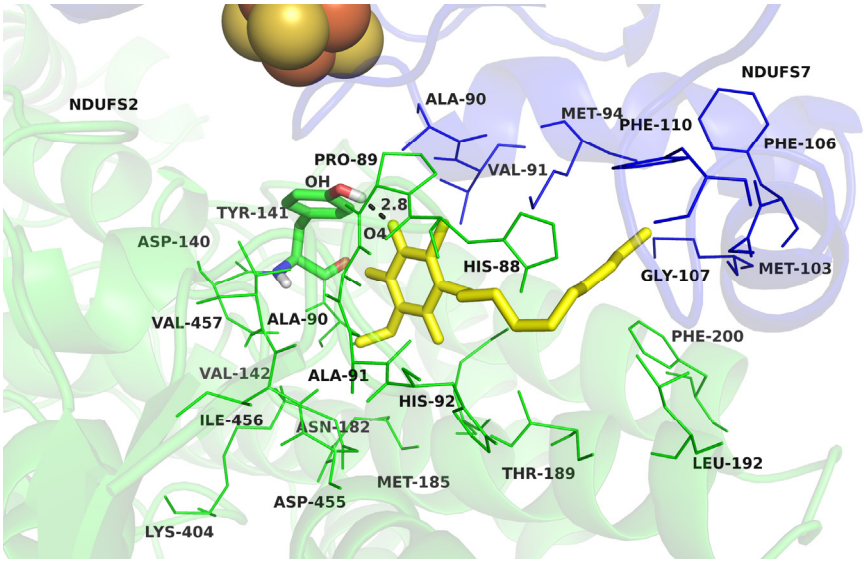


Figure 6.7: Ligand interaction diagram of n-DBQ bound at the interface of *NDUFS2* and *7*. N-DBQ is shown in yellow. The Fe-S cluster is represented as space-fill.

6.7.1 Fluorescence based binding assay

n-DBQ is a substrate showing equal activity as natural ubiquinone [Angerer *et al*, 2012]. n-DBQ showed binding to w-t as well as mutant proteins (Figure 6.8).

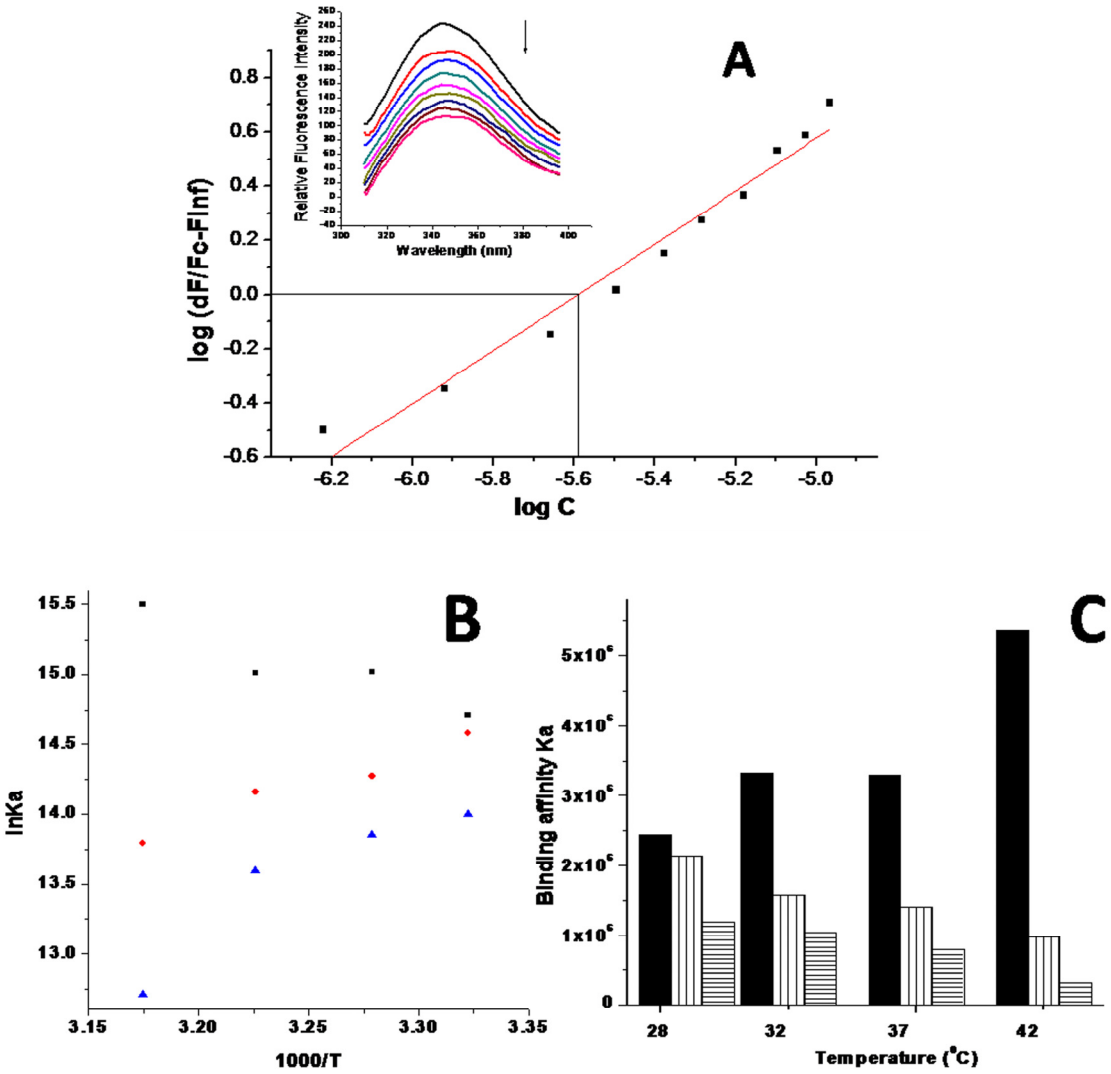


Figure 6.8: Binding of n-DBQ with *NDUFS7* and determination of association constants for interaction of n-DBQ with each of the proteins w-t, V122M and R145H mutants A: Double log plot of log $\{\Delta F/(F_c-F_\infty)\}$ versus log [C] for the w-t *NDUDFS7*, the X intercept of the plot gives pKa value for the interaction, inset shows the quenching of w-t *NDUFS7* fluorescence with n-DBQ. B: Van't Hoff plot for the association w-t (black), V122M (red) and R145H (blue) (r>0.9, N=3), C: Bar graph of binding affinity K_a versus temperature for the w-t (filled), V122M (vertical stripes) and R145H (horizontal stripes) mutants.

The w-t showed an increase in binding affinity with an increase in temperature. However, both the mutants showed a decrease in binding affinity with increase in temperature. Also, the binding affinity of the mutants was less as compared to w-t at all temperatures (Table 6.6, Figure 6.8 B and C). At 37°C (physiological temperature), the binding affinity of the w-t protein ($3.30*10^6$ M^{-1}) was reduced to half for V122M mutant ($1.41*10^6$ M^{-1}) and reduced by four times for the R145H

mutant ($0.80*10^6$ M^{-1}). The binding of n-DBQ with the w-t and mutant proteins was spontaneous as reflected by the negative ΔG values. The binding of n-DBQ with the w-t is endothermic as reflected by the positive ΔH value and that for the mutants is exothermic as reflected by the negative ΔH values. These thermodynamic parameters reflect the different modes of binding of n-DBQ with w-t as against the mutants.

Table 6.6: Binding parameters of n-DBQ binding to *NDUFS7* estimated at different temperatures. Affinity constant Ka is expressed as M^{-1}, ΔH as kJ, ΔG as kJ mol^{-1} and ΔS as kJ $mol^{-1}K^{-1}$.

Temperature	28°C	32°C	37°C	42°C
		w-t		
Ka (M^{-1})	$2.45*10^6$	$3.32*10^6$	$3.30*10^6$	$5.37*10^6$
ΔG	-36.82	-38.08	-38.68	-40.58
ΔS	0.25	0.25	0.25	0.25
ΔH	39.24			
		V122M		
Ka (M^{-1})	$2.14*10^6$	$1.58*10^6$	$1.41*10^6$	$0.98*10^6$
ΔG	-15.84	-15.72	-15.85	-15.69
ΔS	-0.006	-0.007	-0.006	-0.007
ΔH	-17.79			
		R145H		
Ka (M^{-1})	$1.20*10^6$	$1.04*10^6$	$0.80*10^6$	$0.33*10^6$
ΔG	-35.03	-35.13	-35.03	-33.28
ΔS	-0.003	-0.002	-0.003	-0.003
ΔH	-35.92			

6.7.2 MD-simulation of w-t and mutants V122M and R145H with bound n-DBQ

The MD-simulation of the w-t and the mutants performed for a time-scale of 5 ns was useful in understanding the time for which a stable bond is observed between Y141 of *NDUFS2* and n-DBQ. The bond is between the hydroxyl group of Y141 and head carbonyl group of n-DBQ (Figure 6.7). In w-t the bond was observed for 1.8 ns after which n-DBQ moved away from the residue Y141. In the V122M mutant, the critical

hydrogen bond broke within 0.17 ns. In the R145H mutant, though a stable bond was observed till 1.4 ns it was still for less time than with w-t where the bond was maintained up to 1.8 ns (Figure 6.9 A).

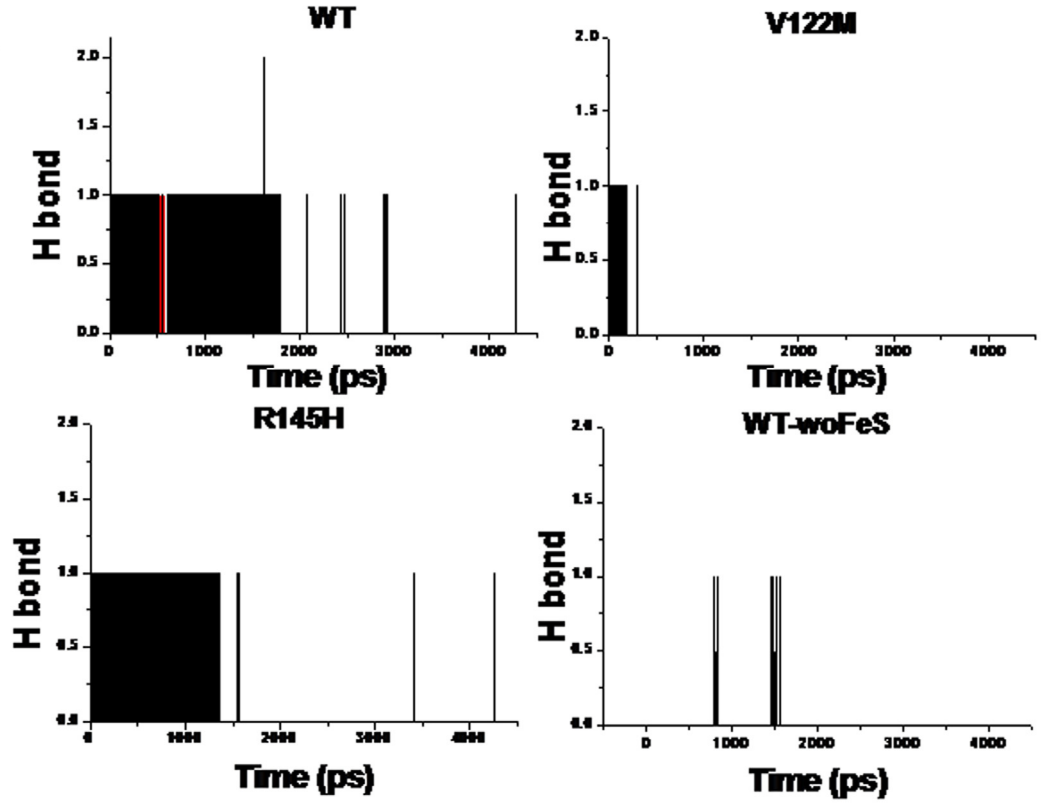


Figure 6.9.1: Plots of hydrogen bond formed between n-DBQ and Y141 of *NDUFS2* versus time (ps) for the w-t, V122M, R145H and w-t without Fe-S cluster.

On monitoring the distance between hydroxyl group of Y141 and the head carbonyl group of n-DBQ in all the 3 trajectories, it was noticed that the time point of bond breakage correlated with the increase in distance between the two interacting groups (Figure 6.9.2).

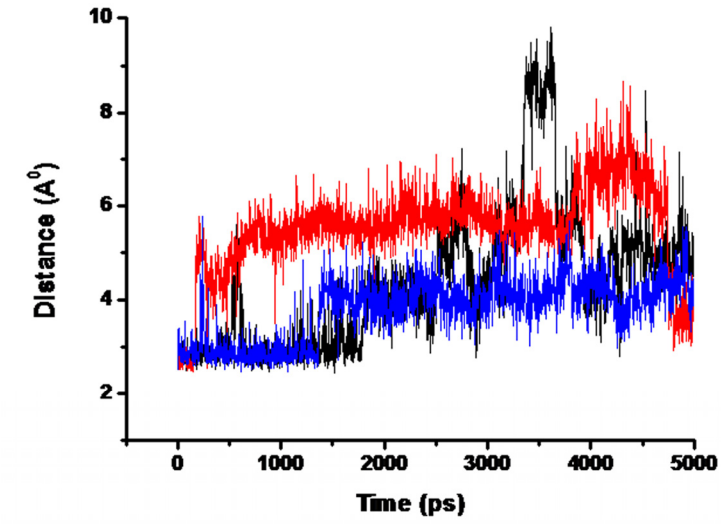


Figure 6.9.2: Plot of distance between hydroxyl group of Y141 and carbonyl head group of n-DBQ versus time (ps). The proteins are represented as w-t *NDUFS7* (black), V122M (red) and R145H (blue).

Large fluctuations of residues near the binding site observed in both mutants would probably affect the binding of n-DBQ (Figure 6.9.3).

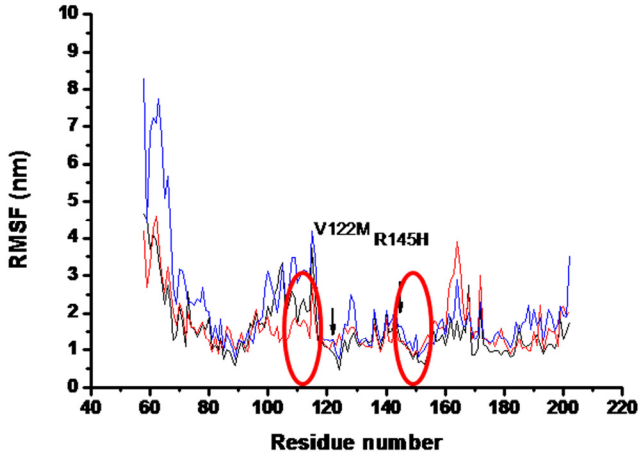


Figure 6.9.3: Root mean square fluctuation (RMSF) (nm) of the residues of *NDUFS7*. The red circles indicate the positions of the residues present at the binding site. The proteins are represented as w-t (black), V122M (red) and R145H (blue).

This observation correlated with the observed reduced affinity of binding in the case of mutants seen in fluorescence assays.

6.8 Aggregation properties

6.8.1 Rayleigh light scattering: Both the mutant proteins, V122M and R145H showed higher scattering intensity as compared to the w-t *NDUFS7* at the same protein concentration, at temperatures 25°C and 37°C (Figure 6.10.1) when measured in the time drive module. Also the scattering intensity of the w-t protein remained constant at 25°C and 37°C which increased significantly for both V122M and R145H mutants (Figure 6.10.1).

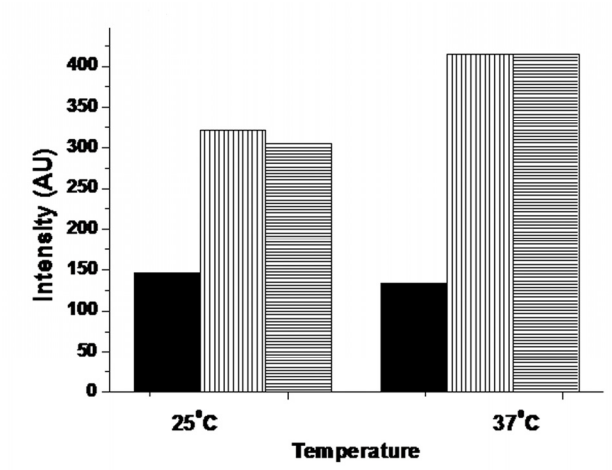


Figure 6.10.1: Rayleigh scattering bar graph, scattering intensity (AU) plotted for w-t (filled) and V122M (vertical stripes) and R145H (horizontal stripes) proteins at 25°C and 37°C.

6.8.2 Th-T dye binding assay: The amyloidophilic Th-T dye shows a characteristic increase in fluorescence upon binding to amyloid fibrils. The w-t *NDUFS7* protein didn't show an increase in fluorescence emission in the concentration range of 50-100 µg ml^{-1} of the protein at 25° and 37°C. However, both the mutant proteins showed a characteristic increase in fluorescence intensity at 25°C as well as 37°C in the same protein concentration range (Figure 6.10.2). R145H mutant showed the highest fluorescence emission at 37°C at any given concentration followed by V122M.

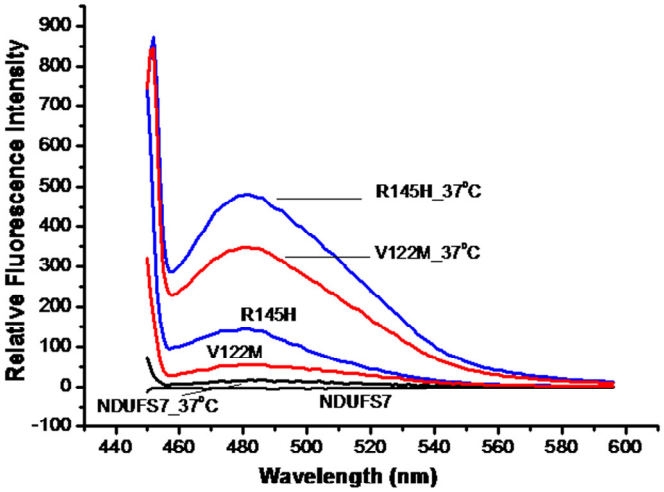


Figure 6.10.2: Thioflavin-T dye binding assay, at 25°C and 37°C, the proteins are represented as w-t (black), V122M (red) and R145H (blue).

6.8.3 Congo red (CR) dye binding assay: CR is similar to Th-T and binds to β-rich structures, inducing an increase in absorption and a red shift in the CR absorption band from 490 to 540 nm. Both the w-t protein and R145H protein showed an increase in absorption and maxima at 500 nm. However, the V122M mutant showed an increase in absorption along with a red shift to 540 nm indicating a propensity for aggregation resulting in β-sheet formed from intermolecular strands (Figure 6.10.3).

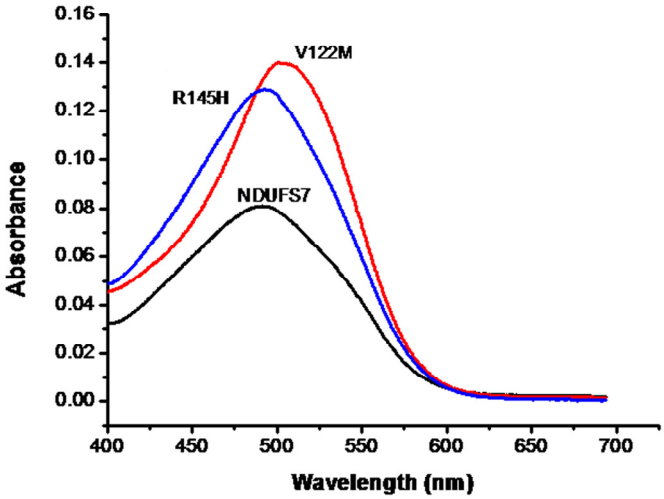


Figure 6.10.3: Congo red dye binding assay, the proteins are represented as w-t (black), V122M (red) and R145H (blue).

6.9 Conclusion

In the present study, the changes in the structure and behaviour of the mutants V122M and R145H as compared to the w-t protein have been studied. The characterization of the proteins has revealed their sensitivity towards pH changes, guanidium-hydrochloride concentration and temperature variation indicating the specificity of environmental conditions required for their stability. The iron sulphur cluster N2 present in the w-t protein is not only essential for the activity of the protein but also gives structural integrity to the protein which is reflected by the drastic changes taking place in the secondary and tertiary structures on removal of iron-sulphur cluster. The high fluctuation of the amino acid residues in the absence of N2 cluster observed during MD-simulation may be the result of instability in the binding of n-DBQ. Thus, the cluster not only controls the movement of the three loops in *NDUFS7* but it also influences the residues near the binding site. The w-t and the two mutant proteins showed differences in the binding of n-DBQ and aggregation properties. The mutant proteins not only aggregated more promptly but also showed reduced binding affinity towards n-DBQ. Thus, even though the mutation sites were away from the n-DBQ binding site, they did influence the activity of the protein; this might be by the destabilizing effect they had on the iron-sulphur cluster N2. The mutated amino acid residue V122M lies within 4Å of the cluster N2 in a cysteine motif, and thus, its influence on the cluster N2 may be consequential. But the amino acid R145 lies comparatively away from the cluster. Several studies indicated the functional relevance of posttranslational modifications of arginine residues in the *NDUFS2, 7*

and *ND1* subunits [Caroll *et al*, 2013]. These modifications have been implicated in the stabilization of subcomplexes that form the initial nucleus of the peripheral arm [Rhein *et al*, 2013]. Although posttranslational modification studies have not focused on the residue R145 of *NDUFS7* as such, it could be a potential candidate for such modifications, thus, playing a role in stabilizing the subcomplexes.

The mutations V122M and R145H may be reducing the binding affinity of n-DBQ by destabilizing the iron-sulphur cluster N2. They form aggregates with spectroscopic characteristics similar to that of amyloid fibrils. All these factors contribute to an unviable Complex-I which will not assemble properly. Enhanced generation of superoxides and energy deficiency leads to clinical variability in symptoms and broad spectrum disorders like Leigh syndrome.

CHAPTER 7

Conclusions

Complex-I is a multi faceted enzyme protein complex with several functions. It is not only involved in transfer of electrons from NADH to ubiquinone but it also generates a proton motive force by pumping protons across the inner mitochondrial membrane. It also has provisions for reduction of reactive oxygen species (ROS). Hence, a large number of iron sulphur clusters are present one after the other at suitable distances to facilitate desired transfer of electrons. For several years the mitochondrial Complex-I was considered to be an L-shaped black box. This was not only because of the lack of structural information but also due to its mutation prone nature. Several mutations have been reported over the years in the 42 mammalian mitochondrial subunits resulting in different neuromuscular disorders [MITOMAP 2014]. Leigh syndrome is one such sub-acute necrotizing encephalomyelopathy with a characteristic neuropathology occurring due to mutations of Complex-I subunits. It is an energy deficit disorder where clinical symptoms depend on the areas of central nervous system that are involved. Due to the dual genetic origin of the Complex-I subunits, n-DNA or mtDNA, the disease could be inherited in an autosomal recessive or maternal fashion. In the present thesis, the human Q module has been extensively studied using computational and biophysical approaches. The aim was to evaluate the structural and functional effects of the Leigh syndrome mutations at molecular level to enhance the understanding of consequences of mutations on Complex-I. Before assembling in the mitochondria, the subunits are individually expressed either within the mitochondrial matrix or transported into it. Thus these proteins exist as individual entities before coming together to form functional assembly. So, any defect in the folding at the initial stage may be carried over to subunit assembly or directly affects assembly formation itself. Thus it is worth studying the subunits individually in isolation and with respect to each other in the Q module environment.

Among the three modules of Complex-I the Q module is the intermediate domain which links the N and P module. In addition to its structural role in Complex-I Q module is also involved in ubiquinone reduction. The four core subunits of the Q module: *NDUFS2, 3, 7* and *8* are highly conserved across species. *In silico* studies of the four subunits and the Q module reflect that point/ double mutations affect the structural integrity of the individual subunits. The mutations are at evolutionarily highly conserved positions in the protein sequence. By identifying the location of the mutation its consequences to some extent can be judged. Breaking of essential

hydrogen bonds and changes in solvent accessibility observed in the mutant subunits have made them behave differently than the w-t. The changes were observed in terms of changes in fluctuation of residues during dynamic simulation, the observed changes in compactness of the structure etc. The energetic instability of subunits by mutation also affected their affinity towards the other subunits, iron-sulphur (Fe-S) cluster and n-DBQ. It would be interesting to note that in *NDUFS2, 3* and *8* the mutations occurred in pairs. Also it was noted that in the case of *NDUFS3* and *8* the single mutations were not causing a lethal phenotype, as the individual mutations were inherited from healthy heterozygous parents [Benit *et al*, 2004; Loeffen *et al*, 1998]. Thus, the disease causing phenotype was observed only when the mutations occur together. Thus, the instability of each mutation calculated by the *in silico* studies would have a cumulative effect on the subunit leading to disastrous consequences.

Although the four subunits of the Q module were cloned and detailed studies were planned, in the case of three of them, *NDUFS2, 8* and ND1, there were difficulties during protein expression stage itself or faced with solubility problems of expressed proteins, hence further studies on them could not be completed. The subunits amenable to cloning and expression, *NDUFS3* and *NDUFS7*, were studied using biochemical and biophysical techniques. These are amongst the reported top 10 proteins most prone to mutations leading to Complex-I deficiency. *In vitro* studies of *NDUFS3* and *7* highlight the common features of their mutants causing Leigh syndrome. Mutants of both subunits shown to be energetically unstable *in silico* had higher aggregation propensities with spectroscopic features of amyloid like fibrils. The T145I+R199W mutant of *NDUFS3* showed a different native structure compared to the w-t although the stabilities under different pH conditions, temperature dependence and Gdn-HCl concentrations were found to be somewhat similar as that of w-t. The higher aggregation propensity, differences in secondary and tertiary structure, tryptophan microenvironment and the higher fluctuation of residues in mutant during molecular dynamics simulation indicate the unstable nature of the mutant. *NDUFS3* is involved early in the assembly of the peripheral arm. Hence, mutation in the subunits disrupting its correct folding would result in hindrance during the early stages of assembly. Hence, even though the expression levels of mutant *NDUFS3* may remain normal in the cell, mutations affect the assembly process of Complex-I thereby resulting in the accumulation of assembly intermediates as observed in patient cell lines [Benit *et al*, 2004].

NDUFS7 plays a functional role in the Q module. It harbours partial binding site for ubiquinone. It is expressed in all tissues similar to housekeeping genes and is the fourth most mutated protein resulting in Complex-I deficiency. Hence, mutations in this gene cause a spectrum of disorders depending on its expression levels in the tissues. The protein possesses the iron-sulphur cluster N2 which plays a vital role in maintaining the structural integrity and functionality of the subunit. Molecular dynamics studies indicate that the absence of cluster N2 has a detrimental effect on the binding of n-DBQ revealing the involvement of cluster N2 in binding. The mutants V122M and R145H showed the absence of signature peak of N2 cluster in UV-visible spectra. The mutants destabilize the cluster N2 which would have an effect on their binding ability. Thus, although the mutation sites are away from the n-DBQ binding site in the structure they can influence the binding of n-DBQ. This was actually proved when reduced binding affinity for n-DBQ was shown in the fluorescence based assay of mutants V122M and R145H of *NDUFS7*. The MD simulation studies also showed the early breakage of one critical hydrogen bond formed with the substrate in the mutants. Thus, even when the mutants showed a fairly similar secondary and tertiary structure with similar stabilities in comparison with the w-t; differences were seen in the aggregation propensities and n-DBQ binding ability as observed in both biophysical and computational studies.

Thus, the mutations affected both the subunits structurally as well as functionally. The subunits *NDUFS3* and *7* have been shown to assemble early during the Complex-I assembly process. *NDUFS3* forms the early peripheral arm while *NDUFS7* acts as an anchor joining the assembled Q module to ND1 in the membrane. Thus, mutations lead to the lack of a fully assembled Complex-I and consequently leads to energy deficit. This may be due to the structurally modified mutant proteins or availability of less number of molecules for assembly pre-empted by aggregation. In the absence of a fully and correctly assembled complex large amount of ROS generation results and that further leads to cellular damage. Muscles and nerves are the most energy dependant tissues and hence are the worst hit by the deleterious mutations. The wide spectrum of symptoms observed in the cases of such mutations are the result of different regions affected by the Complex-I deficiency and resultant ROS damage.

The points to highlight on the observations recorded and conclusions drawn from this study are:

(1) Modelling the subunits structures has helped in identifying the location of mutations in the structure, which in turn has exemplified mutational effects on function based on structure.

(2) Molecular dynamics simulation has demonstrated the effects of mutations on the local and global stability of subunits.

(3) Modelling the Q module has brought out the varied influence of mutations on the stability of assembly, iron-sulphur cluster and ubiquinone binding.

(4) Cloning the subunits and biophysical studies has helped to assess the folding and stability of individual subunits.

(5) Preparation of selected mutants using site-directed mutagenesis and further experimentation has helped to compare the stability of mutants with that of the wild-type.

(6) Results from both computational and experimental approaches have helped us to conclude that the deleterious mutations affect folding and stability of individual subunits as well as assembly, affect stability of iron-sulphur cluster and reduce ubiquinone binding affinity.

(7) Experimental studies have specifically demonstrated the increased propensity of mutant subunits for aggregation in solution compared to their corresponding wild-type protein.

(8) Investigations on subunit NDUFS7 and mutants showed that the signature peak of UV-visible spectrum for iron-sulphur cluster N2 was absent in mutants.

These studies, thus, take us closer towards understanding the molecular effects of pathological mutations in subunits of Complex-I leading to genetic disorders. The effects have been studied via both computational and biophysical approaches. Crystal structure of the human subunits or the entire eukaryotic complex will of course give further insights into understanding the nature of the structural changes resulting in disease phenotype. However, present studies highlight the important roles played by some of the core subunits making up the Q module.

APPENDIX

Cloning of two nuclear encoded and one mitochondrial encoded subunits of human mitochondrial Complex-I

Introduction

NDUFS2 and *8* are two of the core subunits of Complex-I Q module. Mt-ND1 (mitochondrially encoded NADH dehydrogenase 1) gene is encoded by the mt-DNA. It is a 955 bp gene containing no introns. The predicted polypeptide has a molecular weight of 35.6 kDa. This protein is part of the membrane domain of Complex-I and links the membrane domain to the hydrophilic domain.

Mt-ND1 is considered a mutational hotspot in the case of mitochondrial disorders. The initial diagnosis of any suspected mitochondrial disorder always involves the screening of Mt-ND1 gene sequence for mutations. Mutations in ND1 are known to cause disorders like LHON, MELAS, dystonia etc.

Materials and Methods

1. RNA isolation, cDNA preparation, primer design and PCR amplification

Total RNA was isolated from the human colorectal adenocarcinoma cell line HT29 ($1x10^6$ cells) using Trizol® Reagent (Life Technologies, Cat#10296-010) as per the manufacturer's instructions. Purified RNA samples were analyzed by denaturing agarose gel electrophoresis and concentration was spectrophotometrically determined using Nanodrop (Thermo Scientific, USA). One µg of purified RNA was used for the preparation of cDNA using the SuperScript™ III First Strand Synthesis System (Life Technologies, Cat#18080-051).

1.1 *NDUFS2*

Suitable primers (NDUFS2F: 5' AAC ATA TGG CGG CGC TGA GGG 3' & NDUFS3R: 5' ATA AGC TTC ACC GAT CTA CTT CTC CAA ATA CAA TAT C 3') were designed based on the RNA sequences downloaded from NCBI (http://www.ncbi.nlm.nih.gov/) for amplification of the full length *NDUFS3* ORF. The PCR reactions were set in a 50 µl volume containing 1X Pfu buffer (20 mM Tris-HCl pH 8.8 at 25ºC, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-

100, & 0.1 mg ml^{-1} nuclease free BSA), 1 unit of Pfu polymerase, 200 µM each of dNTP and forward and reverse primers and 100 ng of the amplified cDNA. Cycling conditions were 95°C for three minutes, followed by 30 cycles of 95°C for ten seconds, 45°C for forty-five seconds, and 72°C for one minute thirty seconds. The amplicon of desired size (808 bp) was gel extracted using the QIAQuick Gel Purification kit (Qiagen, Cat#28704), which was used for TA cloning in pGEM-T vector (Promega).

1.2 *NDUFS8*

Suitable primers (NDUFS8F: 5'CCG AAT TCA TGC GCT GCC TGA CCA 3' & NDUFS8R: 5' TGC TCG AGT CAC CGA TAC AAG TAG TCA 3') were designed based on the RNA sequences downloaded from NCBI (http://www.ncbi.nlm.nih.gov/) for amplification of the full length *NDUFS7* ORF. The PCR reactions were set in a 50 µl volume containing 1X Pfu buffer (20 mM Tris-HCl pH 8.8 at 25ºC, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100, & 0.1 mg ml^{-1} nuclease free BSA), 1 unit of Pfu polymerase, 200 µM each of dNTP and forward and reverse primers and 100 ng of the amplified cDNA. Cycling conditions were 95°C for 3 min, followed by 30 cycles of 95°C for 10 s, 50°C for 45 s, 72°C for 1 min and final extension of 72ºC for 20 min with hold of 25°C forever. The amplicon of desired size (656 bp) was gel extracted using the QIAQuick Gel Purification kit (Qiagen, Cat#28704), which was used for TA cloning in pGEM-T vector (Promega).

2. TA cloning and sub-cloning in the bacterial expression vector

The purified amplicon was A-tailed using Taq polymerase in 1X Thermopol buffer (NEB) containing 200 µM dATP and 5 units of Taq DNA polymerase (NEB) at 72°C for 20 minutes. The A-tailed amplicons were then cloned into pGEM-T vector (Promega), followed by chemical transformation into *E. coli* DH5α (Invitrogen). The plasmids from the positive colonies, screened through colony PCR, were purified by the standard alkaline lysis method. Full length *NDUFS2* cDNA ORF or *NDUFS8* cDNA ORF from pGEM-T vector was further subcloned into pET 28b(+) vector (Novagen) between Nde I and Xho I sites. Sequences were confirmed by DNA sequencing using BigDye™ Terminator Cycle Sequencing Ready Reaction Kit v3.1

(ABI, Cat#4337457) in an automated 3730 DNA analyzer (ABI). Appropriate plasmids were transformed into *E. coli* BL21(DE3) for protein expression.

The plasmids from the transformed colonies were isolated by the alkaline lysis method. *NDUFS2* was sub-cloned in pET 28b(+) vector (Promega) between Hind III and Xho I. *NDUFS8* was sub-cloned in pGEX-4T1 vector (Promega) between EcoR I and Xho I. The plasmids were transformed chemically into *E. coli* DH5α (Invitrogen). The plasmids were then isolated from transformed colonies by the alkaline lysis method and further transformed in *E. coli* BL21(DE3) expression cells. Colony PCR and Sequencing was done using BigDye™ Terminator Cycle Sequencing Ready Reaction Kit v3.1 (ABI, Cat#4337457) in an automated 3730 DNA analyzer (ABI) to confirm positive clones.

3. Synthesis of the ND1 gene

The sequence of the ND1 gene was downloaded from NCBI. 24 overlapping oligonucleotides of 60 base pairs each were designed using the software DNAWorks which on annealing would form the ND1 construct. The primers were annealed and ND1 was amplified using the protocol described by Hoover DM and Lubkowski J, 2002 [Hoover *et al* 2002]. In short it involves 2 PCR steps. All the 24 oligonucleotides were synthesized at 50 nmol scale. They were mixed and diluted to a final concentration of 1 ng/µl. The PCR reaction 1X Pfu buffer, 200 µM dNTP's and the primer mix was subjected to the following cycling conditions 95°C for five minutes after which the reaction is kept on hold at 80°C for 1 min and Pfu polymerase is added. This is followed by 10 cycles of 95°C for 30 seconds, 62°C for 30 seconds, 72°C for 20 seconds with an increase of 5 seconds per cycle. One µl of this reaction mixture is mixed with 1X Pfu buffer, 200 µl dNTP's, outermost oligonucleotides used as forward and reverse primers and Pfu polymerase and subjected to following reaction conditions 95°C for 5 min, 95°C for 20 seconds, 65°C for 30 seconds, 72°C for 1 minute 30 seconds for 30 cycles. The amplified product was checked by Agarose gel electrophoresis and the amplicon of desired size were purified using the QiaQuick Gel Purification kit (Qiagen Cat.#28704).

Table 1: 24 oligonucleotides designed for in-vitro synthesis of ND1 gene.

Primer name	Sequence
ND1-1	ATGCCAATGGCGAATCTGCTCCTGCTGATCGT
ND1-2	GTTCGGTCAGCATCAGAAACGCCATCGCAATCAGGATCGGAACGATCAGCAGGAGCAGAT
ND1-3	GTTTCTGATGCTGACCGAACGTAAAATTCTGGGTTACATGCAGCTCCGTAAAGGTCCAAA
ND1-4	GTCGGCGAACGGCTGGAGGAGACCATACGGACCAACGACGTTTGGACCTTTACGGAGCTG
ND1-5	CAGCCGTTCGCCGACGCGATGAAACTGTTCACCAAAGAACCGCTGAAACCGGCGACTAGC
ND1-6	CGTCAGCGCCAGGGTAGGCGCGGTGATGTACAGGGTGATGGTGCTAGTCGCCGGTTTCAG
ND1-7	TACCCTGGCGCTGACGATCGCACTCCTCCTCTGGACTCCGCTGCCGATGCCTAACCCGCT
ND1-8	AGAGGTCGCCAGGATAAAGAGCAGGCCCAGGTTCAGGTTTACCAGCGGGTTAGGCATCGG
ND1-9	TCTTTATCCTGGCGACCTCTTCTCTCGCAGTTTATAGCATTCTGTGGTCTGGTTGGGCGT
ND1-10	GCAACCGCACGCAGCGCACCGATGAGCGCGTAGTTAGAGTTAGACGCCCAACCAGACCAC
ND1-11	GCTGCGTGCGGTTGCGCAGACGATCTCTTACGAAGTTACGCTCGCTATCATTCTCCTGTC
ND1-12	AGGGTGGAGAGGTTGAAAGAGCCGGACATCAGGAGCGTAGACAGGAGAATGATAGCGAGC
ND1-13	TCTTTCAACCTCTCCACCCTCATCACCACTCAAGAACACCTGTGGCTGCTGCTCCCATCT
ND1-14	TTTCCGCGAGAGTAGAGATGAACCACATCATCGCCAGCGGCCAAGATGGGAGCAGCAGCC
ND1-15	TCATCTCTACTCTCGCGGAAACCAACCGTACGCCGTTTGACCTGGCAGAAGGTGAATCCG
ND1-16	AGGACCTGCGGCGTATTCGATGTTGAAGCCAGAAACCAGTTCGGATTCACCTTCTGCCAG
ND1-17	AATACGCCGCAGGTCCTTTCGCGCTCTTCTTCATGGCCGAATACACCAACATCATTATGA
ND1-18	TAGTACCGAGGAAGATGGTCGTCGTCAGGGTGTTCATCATAATGATGTTGGTGTATTCGG
ND1-19	CGACCATCTTCCTCGGTACTACCTACGACGCGCTGTCTCCGGAACTGTACACCACTTACT
ND1-20	AGAGAAACAGGCTAGTCAGCAGCAGAGTCTTGGTGACGAAGTAAGTGGTGTACAGTTCCG
ND1-21	GCTGACTAGCCTGTTTCTCTGGATCCGTACCGCATACCCGCGCTTCCGTTATGATCAGCT
ND1-22	CAGAGTCAGAGGCAGGAAGTTTTTCCACAGGAGGTGCATGAGCTGATCATAACGGAAGCG
ND1-23	ACTTCCTGCCTCTGACTCTGGCCCTGCTCATGTGGTACGTTTCTATGCCGATTACCATTA
ND1-24	GGTTTGTGGTGGAATAGAGCTAATGGTAATCGGCATAGAAACG
MTND1F	AACATATGCCGATGGCTAACCTG
MTND1R	TACTCGAGTCAGGTCTGAGGTGGAATG

4. TA cloning and sub-cloning of the synthesized ND1 gene

The synthesized and purified ND1 gene was A tailed and cloned into pGEM-T vector (Promega) as already described. This was then sub-cloned into pGEX-4T1 (Promega) expression vector.

All the constructs were confirmed by colony PCR and sequencing using BigDye™ Terminator Cycle Sequencing Ready Reaction Kit v3.1 (ABI, Cat#4337457) in an automated 3730 DNA analyzer (ABI).

5. Expression studies

5.1 IPTG induction

Single colony of *E coli* BL21(DE3) clone was inoculated in 5 ml LB containing a suitable antibiotic (Ampicillin 100 μg ml^{-1} or Kanamycin 60 μg ml^{-1}) and incubated overnight at 37°C (200 rpm). The overnight grown culture was spun at 3500 rpm for 10 min at 25 °C. The cell pellet was resuspended in 5 ml LB containing suitable antibiotic and transferred to 100 ml LB with suitable antibiotic. The cells were grown till OD read 0.3; at 37 °C and shaking at 150 rpm. The cells were then aseptically induced with 0.1-0.5 mM IPTG and were further grown at 18 °C, 28 °C and 37 °C with a shaking at 200 rpm till $O.D_{600}$ = 0.6-0.7. 1ml of the above sample was used to check the expression profile. Cells were harvested by spinning at 4000 rpm for 15 minutes at 4°C. The pellet was redissolved in 20 ml Lysis Buffer (100 mM Tris, 200 mM Nacl, 0.01 % IGEPAL® CA630 and 1 mg ml^{-1} lysozyme) and was kept on ice for 30 minutes with intermittent shaking. Cells were lyzed by sonication at 60% Power in Esquire Biotech Ultrasonic homogenizer with a pulse of 6 s followed by pause of 8 s for 10 minutes. The sonicated solution was centrifuged at 10,000 rpm for 45 minutes at 4°C. The pellet and supernatant were separated. The expression profile was checked on 12% SDS-PAGE gel followed by Coomassie Brilliant Blue R-250 staining.

5.2 Autoinduction

The autoinduction protocol was similar to the protocol described in Materials and methods section 11. In brief, the primary culture grown overnight in LB containing suitable antibiotic was transferred to fresh LB containing suitable antibiotic. The autoinduction medium was prepared as described in section 11. The primary culture was inoculated in the autoinduction medium (1 % inoculum). The flasks were incubated at 37 °C for 2, 4 and 6 hours (standardized for soluble expression) followed by a prolonged incubation at either 16 °C, 18 °C and 20 °C. 1ml of the above sample was used to check the expression profile. Cells were harvested by spinning at 4000 rpm for 15 minutes at 4°C. The pellet was redissolved in 20 ml Lysis Buffer (100 mM Tris, 200 mM Nacl, 0.01 % IGEPAL® CA-630 and 1 mg ml^{-1} lysozyme) and was kept on ice for 30 minutes with intermittent shaking. Cells were lyzed by sonication at

60% Power in Esquire Biotech Ultrasonic homogenizer with a pulse of 6 s followed by pause of 8 s for 10 minutes. The sonicated solution was centrifuged at 10,000 rpm for 45 minutes at 4°C. The pellet and supernatant were separated. The expression profile was checked on 12% SDS-PAGE gel followed by Coomassie Brilliant Blue R-250 staining.

Results

1. RNA was isolated from both healthy human blood and HT29 cell line. The concentration of RNA isolated from HT 29 cell line was estimated to be 415 ng/µl. This RNA was further used for cDNA preparation.

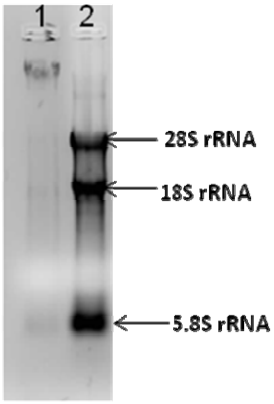


Fig 1: Formaldehyde Agarose gel for identifying RNA extracted, Lane 1: from blood and Lane 2: from HT29 cell line.

cDNA was prepared and used for further amplification of the genes with suitable primers. NDUFS2 and 8 amplified and were identified depending upon their appropriate sizes. NDUFS2 was ~1.4 kb and NDUFS8 was 646 bp.

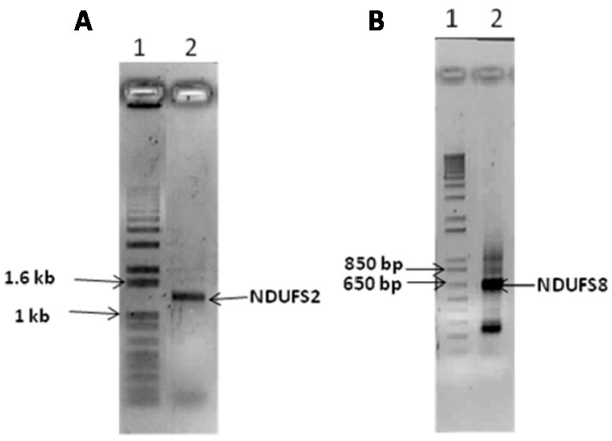


Fig 2: Agarose gel showing amplification of the gene constructs (A) Lane 1: 1 kb plus DNA ladder, Lane 2: PCR product amplified with NDUFS2 primers and (B) Lane 1: 1 kb plus DNA ladder, Lane 2: PCR product amplified with NDUFS8 primers from cDNA prepared from RNA isolated from HT29 cell line.

2. Cloning in pGEMT vector and colony PCR to confirm positive clones.

Several colonies were obtained after cloning the amplified PCR product into pGEMT vector. Of these a few colonies were subjected to colony PCR and appropriate clone was chosen for further plasmid isolation.

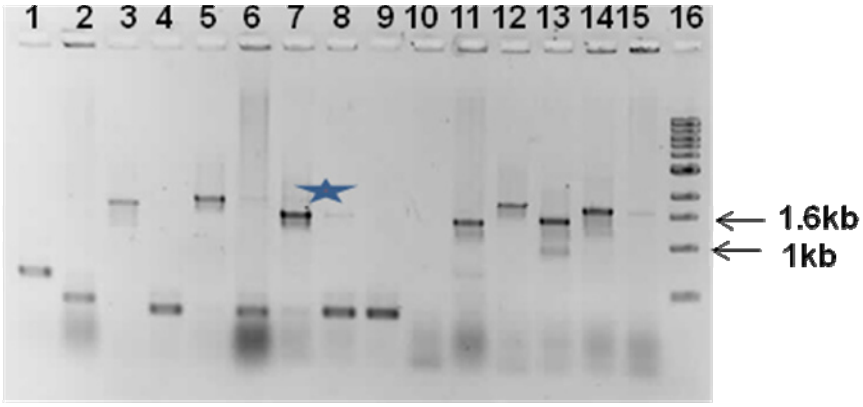


Figure 3.1: Agarose gel electrophoresis Lane 1-15: Colony PCR products for colony number 1-15 for NDUFS2 gene cloned in pGEMT vector and transformed in *E coli* DH5α cells. Lane 16: 1 kb NEB ladder. Star indicates: Amplified NDUFS2 colony PCR product.

Colony number 7, 11 and 13 show amplified constructs of ~1.4 kb which is the size of the NDUFS2 gene. (Figure 3.1)

Colony number 2, 3, 4 and 7 showed amplified bands at ~850 bp. The size of the NDUFS8 gene is 655 bp and including the extra bases incorporated by the primers this is the appropriate size (Figure 3.2).

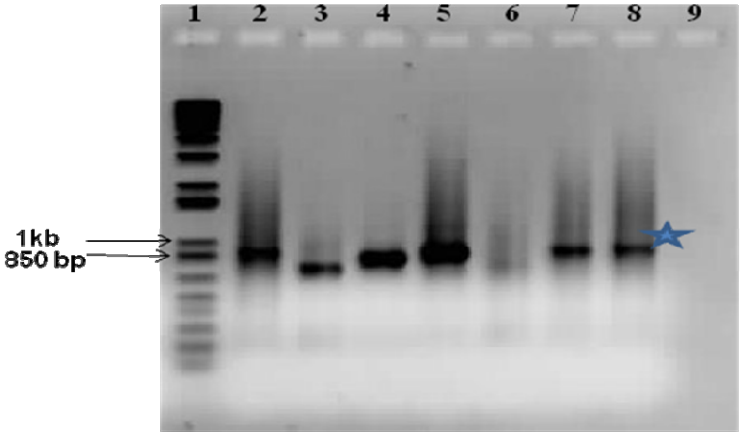


Figure 3.2: Agarose gel electrophoresis Lane 1: 1 kb plus DNA ladder, Lane 2-9: Colony PCR products for colony number 1-8 for NDUFS8 gene cloned in pGEMT vector and transformed in *E. coli* DH5α cells. Star indicates the presence of NDUFS8 Colony PCR product.

The potential positive clones were used for plasmid isolation and the plasmids were used for sub-cloning in expression vector and were sequenced to confirm the presence of the gene.

3. Sub-cloning in expression vector.

Colony PCR was performed to confirm positive clones followed by sequencing.

NDUFS2 was sub-cloned in pET 28b(+) between Nde I and Hind III.

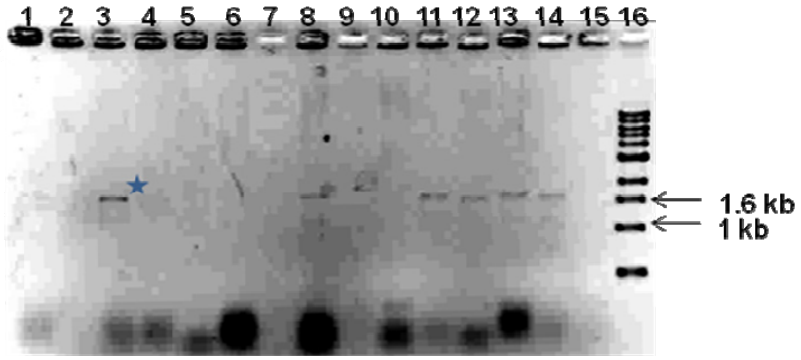


Figure 4.1: Agarose gel electrophoresis Lane 1-15: Colony PCR products for colony number 1-15 for NDUFS2 gene cloned in pET 28b(+) vector and transformed in *E coli* DH5α cells, Lane 16: 1 kb NEB Ladder. Star indicates the presence of NDUFS2 Colony PCR product.

Colony number 3, 8, 11, 13 and 14 show bands of appropriate sizes and are positive clones (Figure 4.1).

NDUFS8 was sub-cloned in pGEX 4T1 between Eco RI and Xho I.

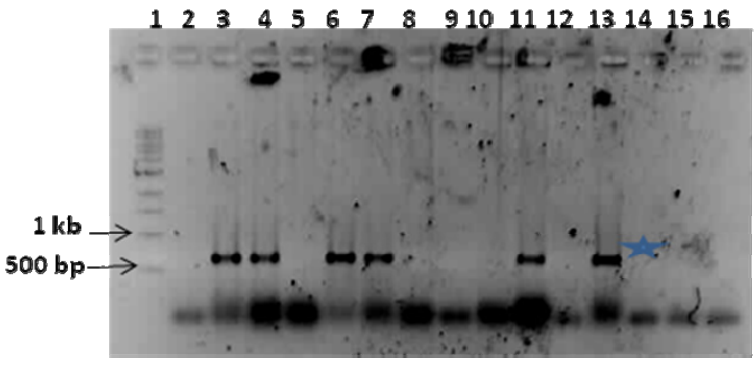


Figure 4.2: Agarose gel electrophoresis Lane 1: 1 kb NEB Ladder, Lane 2-16 Colony PCR products for colony number 1-15 for NDUFS8 gene cloned in pET 28b(+) vector and transformed in *E coli* DH5α cells. Star indicates the presence of NDUFS8 Colony PCR product.

Colony number 2, 3, 5, 6, 10 and 12 showed bands of appropriate size and are positive clones (Figure 4.2)

The plasmids from the positive clones were sequenced to confirm the presence of desired gene and transformed into *E coli* BL21(DE3) expression cells. The clones were then checked for expression.

4. Soluble protein expression

Recombinant *NDUFS2* and *8* proteins were expressed only as inclusion bodies in all the conditions of expression.

5. *In vitro* synthesis of ND1 gene.

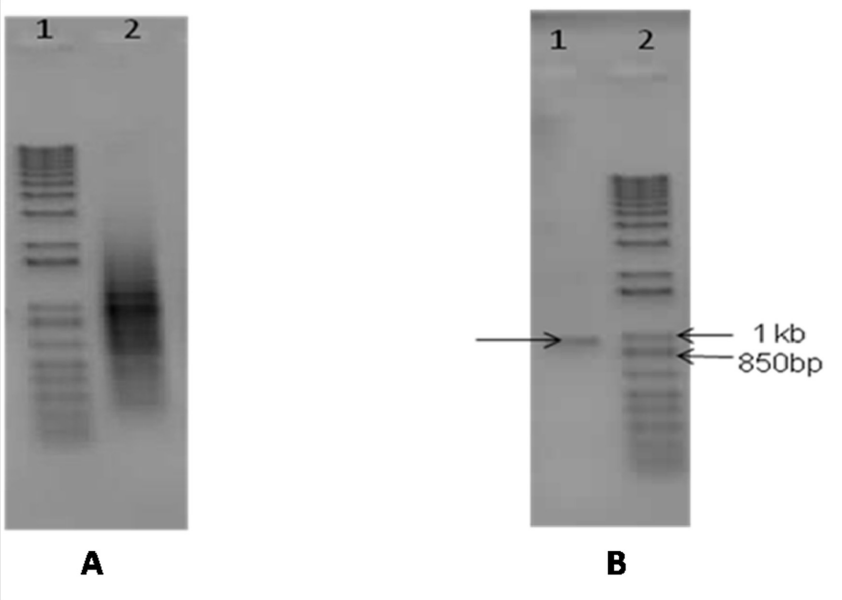


Figure 5: A: PCR product after 1st PCR reaction. B: PCR product after 2nd PCR reaction. Arrow indicates the presence of amplified ND1 gene.

ND1 is 956 bp. In the first PCR reaction a general assembly of the oligonucleotides takes place. ND1 got synthesized in the 2nd PCR reaction where the outermost oligonucleotides were used for amplification (Figure 5). The amplified PCR product was cloned into pGEM-T vector. 96 colonies were screened and sequenced to get an appropriate clone. The construct was further sub-cloned in pET 26b(+) (Nde I – Xho I), 28b(+) (Nde I-Xho I), pET 30b(+) (EcoRI – Xho I), pET 44b(+) (EcoR I – Hind III) and pGEX4T1. Expression was tried by inducing the cells at different IPTG

concentrations (0.1-1 mM) at different temperatures (16°C, 18°C, 21°C, 28°C, 37°C) in different expression hosts (Rosetta, pLysS, pLysE); however the clones didn't show overexpression of the ND1 gene. Autoinduction was also tried, however, it also didn't improvise the expression profile for carrying out further research on it.

BIBLIOGRAPHY

1. Achilli A, Iommarini L, Olivieri A, Pala M, Hooshiar Kashani B, Reynier P, La Morgia C, Valentino ML, Liguori R, Pizza F, Barboni P, Sadun F, De Negri AM, Zeviani M, Dollfus H, Moulignier A, Ducos G, Orssaud C, Bonneau D, Procaccio V, Leo-Kottler B, Fauser S, Wissinger B, Amati-Bonneau P, Torroni A, Carelli V (2012) Rare primary mitochondrial DNA mutations and probable synergistic variants in Leber's hereditary optic neuropathy. Plos One. **7**, e42242.
2. Alexeyev MF, Venediktova N, Pastukh V, Shokolenko I, Bonilla G, Wilson GL (2008) Selective elimination of mutant mitochondrial genomes as therapeutic strategy for the treatment of NARP and MILS syndromes. Gene Ther. **15**, 516-523
3. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic Local Alignment Search Tool. J. Mol. Biol. **215**, 403-410.
4. Angerer H, Nasiri HR, Niedergesab V, Kerscher S, Schwalbe H, Brandt U (2012) Tracing the tail of ubiquinone in mitochondrial complex-I. Biochim Biophys Acta. **1817**, 1776-1784.
5. Antonicka H, Leary SC, Guercin GH, Agar JN, Horvath R, Kennaway NG, Harding CO, Jaksch M, Shoubridge EA (2003) Mutations in COX10 result in a defect in mitochondrial heme A biosynthesis and account for multiple, early onset clinical phenotypes associated with isolated COX deficiency. Hum. Mol. Genet. **12**, 2693-2702.
6. Antonicka H, Ogilvie I, Taivassalo T, Anitori R, Haller R, Vissing J, Kennaway N, Shoubridge E (2003) Identification and characterization of a common set of complex I assembly intermediates in mitochondria from patients with complex I deficiency. J. Biol. Chem. **278**, 43081-43088.
7. Antonicka H, Ostergaard E, Sasarman F, Weraarpachai W, Wibrand F, Pedersen AM, Rodenburg RJ, van der Knaap MS, Smeitink JA, Chrzanowska-Lightowlers ZM, Shoubridge EA (2010) Mutations in C12orf65 in patients with encephalomyopathy and a mitochondrial translation defect. Am. J. Hum. Genet. **87**, 115-122.
8. Arii J, Tanabe Y (2000) Leigh syndrome: serial MR imaging and clinical follow-up. AJNR Am. J. Neuroradiol. **21**, 1502-1509.

9. Balton E, Wang Y, Thiessen PA, Bryant SH (2008) PubChem: Integrated Platform of Small Molecules and Biological Activities. Chapter12 Annual Reports in Computational Chemistry.4, American Chemical Society. Washington DC.
10. Baradaran R, Berrisford J, Minhas G, Saznov L (2013) Crystal structure of the entire respiratory complex-I. Nature **494**, 443-448.
11. Barghuti F, Elian K, Gomori JM, Shaag A, Edvardson S, Saada A, Elpeleg O (2008) The unique neuroradiology of complex I deficiency due to NDUFA12L defect. Mol.Genet. Metab. **94,** 78–82.
12. Bellance N, Lestienne P, Rossignol R (2009) Mitochondria: from bioenergetics to the metabolic regulation of carcinogenesis. Front. Biosci. (Landmark Ed.) **14**, 4015-4034.
13. Benit P, Beugnot R, Chretien D, Giurgea I, De Lonlay-Debeney P, Issartel JP, Corral-Debrinski M, Kerscher S, Rustin P, Rötig A, Munnich A (2003) Mutant NDUFV2 subunit of mitochondrial complex I causes early onset hypertrophic cardiomyopathy and encephalopathy. Hum. Mutat. **21,** 582–586.
14. Benit P, Chretien D, Kadhom N, de Lonlay-Debeney P, Cormier –Daire V, Cabral A, Peudenier S, Rustin P, Munnich A, Rotig A (2001) Large scale deletions and point mutations of the nuclear NDUFV1 and NDUFS1 genes in the mitochondrial complex I deficiency. Am. J. Hum. Genet. **68**, 1344-1352.
15. Benit P, Goncalves S, Dassa EP, Briiere JJ, Rustin P (2008) The variability of the harlequin mouse phenotype resembles that of human mitochondrial complex I deficiency. Plos One. **3**, e3208.
16. Benit P, Slama A, Cartault F, Giurgea I, Chretien D, Lebon S, Marsac C, Munnich A, Rotig A, Rustin P (2004) Mutant NDUFS3 subunit of mitochondrial complex I causes Leigh syndrome. J. Med. Genet. **41**, 14-17.
17. Berendsen HJC, Postma JPM, van Gunsteren WF, Nola AD, Haak JR (1984) Molecular dynamics with coupling to an external bath. J. Chem. Phys. **81**, 3684-3690.
18. Berger I, Hershkovitz E, Shaag A, Edvardson S, Saada A, Elpeleg O (2008) Mitochondrial complex I deficiency caused by a deleterious NDUFA11 mutation. Ann. Neurol. **63**, 405–408.
19. Berrisford JM, Sazanov LA (2009) Structural basis for the mechanism of respiratory complex I. J. Biol. Chem. **284**, 29773-29783.

20. Bourgeron T, Rustin P, Chretien D, Birch-Machin M, Bourgeois M, Viegas-Pequignot E, Munnich A, Rotig A (1995) Mutation of a nuclear succinate dehydrogenase gene results in mitochondrial respiratory chain deficiency. Nat. Genet. **11**, 144-149.
21. Bradford M (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. **72**, 248-254.
22. Brautbar A, Wang J, Abdenur JE, Chang RC, Thomas JA, Grebe TA, Lim C, Weng SW, Graham BH, Wong LJ (2008) The mitochondrial 13513G>A mutation is associated with Leigh disease phenotypes independent of complex I deficiency in muscle. Mol. Genet. Metab. **94**, 485-490.
23. Brody S, Mikolajczyk S (1988) Neurospora mitochondria contain an acyl carrier protein. Eur. J. Biochem. **173**, 353-359.
24. Budde SM, van den Heuvel LP, Janssen AJ, Smeets RJ, Buskens CA, DeMeirleir L, Van Coster R, Baethmann M, Voit T, Trijbels JM, Smeitink JA (2000) Combined enzymatic complex I and III deficiency associated with mutations in the nuclear encoded NDUFS4 gene. Biochem. Biophys. Res. Commun. **275**, 63–68.
25. Bugiani M, Invernizzi F, Alberio S, Briem E, Lamantea E, Carrara F, Moroni I, Farina L, Spada M, Donati MA, Uziel G, Zeviani M (2004) Clinical and molecular findings in children with complex I deficiency. Biochim. Biophys. Acta **1659**, 136–147.
26. Burstein EA, Abornev SM, Reshetnyak YK (2001) Decomposition of protein tryptophan fluorescence spectra into log normal components I. Decomposition algorithms. Biophys. J. **81**, 1699-1709.
27. Bych K, Kerscher S, Netz DJ, Pierik AJ, Zwicker K, Huynen MA, Lill R, Brandt U, Balk J (2008) The iron–sulphur protein Ind1 is required for effective complex I assembly. EMBO J. **27**, 1736–1746.
28. Calvo SE, Tucker EJ, Compton AG, Kirby DM, Crawford G, Burtt NP, Rivas M, Guiducci C, Bruno DL, Goldberger OA, Redman MC, Wiltshire E, Wilson CJ, Altshuler D, Gabriel SB, Daly MJ, Thorburn DR, Mootha VK (2010) High-throughput, pooled sequencing identifies mutations in NUBPL and FOXRED1 in human complex I deficiency. Nat. Genet. **42**, 851–858.

29. Cardol P, Boutaffala L, Memmi S, Devreese B, Matagne RF, Remacle C (2008) In Chlamydomonas, the loss of ND5 subunit prevents the assembly of whole mitochondrial complex I and leads to the formation of a low abundant 700 kDa subcomplex. Biochim. Biophys. Acta **1777**, 388–396.
30. Caroll J, Ding S, Fearnley IM, Walker JE (2013) Post-translational modifications near the quinone binding site of mammalian complex I. J. Biol. Chem. **288**, 24799-24808.
31. Carroll J, Fearnley IM, Shannon RJ, Hirst J, Walker JE (2003) Analysis of subunit composition of complex I from bovine heart mitochondria. Mol. Cell Proteomics **2**, 117-126.
32. Carroll J, Fearnley IM, Skehel M, Shannon RJ, Hirst J, Walker JE (2006) Bovine Complex I is a Complex of 45 Different Subunits. J. Biol. Chem. **281**, 32724-32727.
33. Castro-Gago M, Blanco-Barca MO, Campos-Gonzalez Y, Arenas-Barbero J, Pintos-Martinez E, Eiris-Punal J (2006) Epidemiology of pediatric mitochondrial respiratory chain disorders in northwest Spain. Pediatr. Neurol. **34**, 204-211.
34. Chang E, Liao T-Y, Lim T-S, Fann W, Chen R-P (2009) A new Amyloid-Like β-aggregate with Amyloid Characteristics, Except Fibril Morphology. J. Mol. Biol. **385**, 1257-1265.
35. Ching CK, Mak CM, Au KM, Chan KY, Yuen YP, Yau EK, Ma LC, Chow HL, Chan AY (2013) A patient with congenital hyperlactatemia and Leigh syndrome: an uncommon mitochondrial variant. Hong Kong Med. J. **19**, 357-361.
36. Chipman D, Grisaro V, Sharon N (1967) The binding of oligosaccharides containing N-acetylglucosamine and N-acetylmuramic acid to lysozyme. The specificity of binding subsites. J. Biol. Chem. **242**, 4388-4394.
37. Claros MG, Vincens P (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur. J. Biochem. **241**, 779-786.
38. Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. **2**, 1511-1519.
39. Darden T, York D, Pedersen LG (1993) Particle Mesh Ewald: An Nlog (N) method for Ewald sums in large systems. J. Chem. Phys. **98**, 10089-10092.

40. Darin N, Oldfors A, Moslemi AR, Holme E, Tulinius M (2001) The incidence of mitochondrial encephalomyelopathies in childhood: clinical features and morphological, biochemical and DNA abnormalities. Ann. Neurol. **49**, 377-383.
41. Darrouzet E, Issartel J, Lunardi J, Dupuis A (1998) The 49-kDa subunit of NADH-ubiquinone oxidoreductase (Complex I) is involved in the binding of piericidin and rotenone, two quinine-related inhibitors. FEBS Lett. **431**, 34-38.
42. DeHaan C, Habibi-Nazhad B, Yan E, Salloum N, Parliament M, Allalunis-Turner J (2004) Mutation in mitochondrial complex I ND6 subunit is associated with defective response to hypoxia in human glioma cells. Mol. Cancer **3**, 19-34.
43. Desmond Molecular Dynamic System version 3.1, D.E. Shaw Research, New York, NY, 2012. Maestro-Desmond Interoperability Tools, version 3.1, Schrȍdinger, New York, NY, 2012.
44. Dieteren CE, Willems PH, Vogel RO, Swats HG, Fransen J, Roepman R, Crienen G, Smeitnik JA, Nijtmans LG, Koopman WJ (2008) Subunits of mitochondrial complex I exist as a part of matrix- and membrane associated subcomplexes in living cells. J. Biol. Chem. **283**, 34753-34761.
45. DiMauro S, Schon EA (2003) Mitochondrial respiratory-chain diseases. N. Engl. J. Med. **348**, 2656-2668.
46. Duarte M, Populo H, Videira A, Friedrich T, Schulte U (2002) Disruption of iron-sulphur cluster N2 from NADH: ubiquinone oxidoreductase by site-directed mutagenesis. Biochem. J. **364**, 833-839.
47. Dunning CJ,McKenzie M, Sugiana C, Lazarou M, Silke J, Connelly A, Fletcher JM, Kirby DM, Thorburn DR, Ryan MT (2007) Human CIA30 is involved in the early assembly of mitochondrial complex I and mutations in its gene cause disease. EMBO J. **26**, 3227–3237.
48. Efremov RG, Sazanov LA (2011) Structure of the membrane domain of respiratory Complex I. Nature **476**, 414-420.
49. Elliott HR., Samuels DC, Eden JA, Relton CL, Chinnery PF (2008) Pathogenic Mitochondrial DNA Mutations Are Common in the General Population. Am. J. Hum. Genet. **83**, 254–260.

50. Emahazion T, Beskow A, Gyllensten U, Brookes AJ (1998) Intron based radiation hybrid mapping of 15 complex I genes of the human electron transport chain. Cytogenet. Cell Genet. **82**, 115-119.
51. Eswar N, Eramian D, Webb B, Shen MY, Sali A (2008) Protein structure modeling with MODELLER. Methods Mol. Biol. **426**, 145-59.
52. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A (2007) Comparative Protein Structure Modeling With MODELLER. Curr Protoc Protein Sci, Chapter 2: Unit 2.9. doi: 10.1002/0471140864.ps0209s50, 2007.
53. Fassone E, Rahman S (2012) Complex I deficiency: clinical features, biochemistry and molecular genetics. J. Med. Genet. **49**, 578-590.
54. Fernandez-Moreira D, Ugalde C, Smeets R, Rodenburg RJ, Lopez-Laso E, Ruiz-Falco ML, Briones P, Martin MA, Smeitnik JA, Arenas J (2007) X-linked NDUFA1 gene mutations associated with mitochondrial encephalomyopathy. Ann. Neurol. **61**, 73-83.
55. Fernández-Vizarra E, Tiranti V, Zeviani M (2009) Assembly of the oxidative phosphorylation system in humans: What we have learned by studying its defects. Biochim. Biophys. Acta **1793**, 200–211.
56. Friedrich T, Steinmüller K, Weiss H (1995) The proton-pumping respiratory complex I of bacteria and mitochondria and its homologue in chloroplasts. FEBS Lett. **367**, 107-111.
57. Friedrich T, Weiss H (1997) Modular Evolution of the Respiratory NADH:Ubiquinone Oxidoreductase and the origin of its Modules. J. Theor. Biol. **187**, 529-540.
58. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelly M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) Glide: A New Approach for Rapid Accurate Docking and Scoring.1. Method and Assessment of Docking Accuracy. J. Med. Chem. **47**, 1739-1749.
59. Gerards M, Sluiter W, van den Bosch BJ, de Wit LE, Calis CM, Frentzen M, Akbari H, Schoonderwoerd K, Scholte HR, Jongbloed RJ, Hendrickx AT, de Coo IF, Smeets HJ (2010) Defective complex I assembly due to C20orf7 mutations as a new cause of Leigh syndrome. J. Med. Genet. **47**, 507-512.
60. Gerards M, van den Bosch BJ, Danhauser K, Serre V, vanWeeghel M, Wanders RJ, Nicolaes GA, Sluiter W, Schoonderwoerd K, Scholte HR,

Prokisch H, Rotig A, de Coo IF, Smeets HJ (2011) Riboflavin-responsive oxidative phosphorylation complex I deficiency caused by defective ACAD9: new function for an old gene, Brain **134**, 210–219.

61. Ghezzi D, Seviroukova I, Invernizzi F, Lamperti C, Mora M, D Adamo P, Novara F, Zuffardi O, Uziel G, Zeviani M (2010) Severe X-linked mitochondrial encephalopathy associated with a mutation in apoptosis-inducing factor. Am. J. Hum. Genet. **86**, 639-649.
62. Glide, version 5.8, Schrodinger, LLC, New York, NY, 2012.
63. Grigorieff N (1998) Three-dimensional structure of bovine NADH:ubiquinone oxidoreductase (complex I) at 22 A in ice. J. Mol. Biol. **277**, 1033-1046.
64. Grinvald A, Steinberg IZ (1974) On analysis of fluorescence decay kinetics by the method of least squares. Anal. Biochem. **59**, 583-598.
65. Groenning M (2010) Binding mode of Thioflavin T and other molecular probes in the context of amyloid fibrils-current status. J. Chem. Biol. **3**, 1-18.
66. Guo Z, Mohanty U, Noehre J, Sawyer TK, Sherman W, Krilov G (2010) Probing the alpha-helical structural stability of stapled p53 peptides: molecular dynamics simulations and analysis. Chem. Biol. Drug Des. **75**, 348-359.
67. Haack TB, Danhauser K, Haberberger B, Hoser J, Strecker V, Boehm D, Uziel G, Lamantea E, Invernizzi F, Poulton J, Rolinski B, Iuso A, Biskup S, Schmidt T, Mewes HW, Wittig I, Meitinger T, Zeviani M, Prokisch H (2010) Exome sequencing identifies ACAD9 mutations as a cause of complex I deficiency. Nat. Genet. **42,** 1131–1134.
68. Haack TB, Haberberger B, Frisch E-M, Wieland T, Iuso A, Gorza M, Strecker V, Garf E, Mayr JA, Herberg U, Hennermann JB, Klopstock T, Kuhn KA, Ahting U, Sperl W, Wilichowski E, Hoffmann GF, Tesarova M, Hansikova H, Zeman J, Plecko B, Zeviani M, Wittig I, Strom TM, Schuelke M, Freisinger P, Meitinger T, Prokissch H (2012) Molecular diagnosis in mitochondrial complex I deficiency using exome sequencing. J. Med. Genet. **49**, 277-283.
69. Hall BG (2013) Building phylogenetic trees from molecular data with MEGA. Mol. Biol. Evol. **30**, 1229-1235.
70. He H-W, Zhang J, Zhou H-M, Yan Y-B (2005) Conformational Change in the C terminal Domain is responsible for the initiation of creatine kinase thermal aggregation. Biophys. J. **89**, 2650-2658.

71. Heinrich H, Werner S (1992) Identification of the ubiquinone-binding site of NADH:ubiquinone oxidoreductase (complex I) from Neurospora crassa. Biochemistry. **31**, 11413-11419.
72. Hess B, Bekker H, Berendsen HJC, Fraaije GEM (1997) LINCS: A linear constraint solver for molecular simulations. J. Comput. Chem. **18**, 1463-1472.
73. Hess B, Kutzner C, Van der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J. Chem. Theory Comput. **4**, 435-447.
74. Hinchliffe P, Sazanov L (2005) Organization of Iron-Sulfur Clusters in Respiratory Complex I. Science **309**,771-774.
75. Hintala R, Smeets R, Moilanen JS, Ugalde C, Uusimaa J, Smeitnik JAM, Majamaa K (2006) Analysis of mitochondrial DNA sequences in patients with isolated or combined oxidative phosphorylation system deficiency. J. Med. Genet. **43**, 881-886.
76. Hixon J, Reshetnyak YK (2009) Algorithm for the analysis of tryptophan fluorescence spectra and their correlation with protein structural parameters. Algorithms. **2**, 1155-1176.
77. Hoefs SJ, Dieteren CE, Distelmaier F, Janssen RJ, Epplen A, Swarts HG, Forkink M, Rodenburg RJ, Nijtmans LG, Willems PH, Smeitink JA, van den Heuvel LP (2008) NDUFAF2 complex I mutation leads to Leigh disease. Am. J. Hum.Genet. **82**, 1306-1315.
78. Hoefs SJ, Dieteren CE, Rodenburg RJ, Naess K, Bruhm H, Wibom R, Wagena E, Willems PH, Smeitnik JA, Nijtmans LG, Van den Heuvel LP (2009) Baculovirus complementation restores a novel NDUFSAF2 mutation causing complex I deficiency. Hum. Mutat. **30**, E728-736.
79. Hoefs SJ, Skjeldal OH, Rodenburg RJ, Nedregaard B, van Kaauwen EP, Spiekerkötter U, von Kleist-Retzow JC, Smeitink JA, Nijtmans LG, van den Heuvel LP (2010) Novel mutations in the NDUFS1 gene cause low residual activities in human complex I deficiencies. Mol. Genet. Metab. **100,** 251–256.
80. Hoefs SJ, van Spronsen FJ, Lenssen EW, Nijtmans LG, Rodenburg RJ, Smeitnik JA, van den Heuvel LP (2011) NDUFA10 mutations cause complex I deficiency in a patient with Leigh disease. Eur. J. Hum. Genet. **19**, 270-274.

81. Hoover DM, Lubkowski J (2002) DNAWorks:an automated method for designing oligonucleotides for PCR-based gene synthesis. Nucleic Acids Res. **30**, e43.
82. Howell N, Bindoff LA, McCullough DA, Kubacka I, Poulton J, Mackey D, Taylor L, Turnbull DM (1991) Leber hereditary optic neuropathy : identification of the same mitochondrial ND1 mutations in six pedigrees. Am. J. Hum. Genet. **49**, 939-950.
83. Hunte C, Zickermann V, Brandt U (2010) Functional Modules and Structural Basis of Conformational Coupling in Mitochondrial Complex I. Science **329**, 448-451.
84. Hyslop SJ, Ducan AMV, Pitkanen S, Robinson BH (1996) Assignment of the PSST subunit gene of human mitochondrial complex I to chromosome 19p13. Genomics **37**, 375-380.
85. Inokuti M, Hirayama F (1965) Influence of energy transfer by the exchange mechanism on donor luminescence. J. Chem. Phys. **43**, 1978-1989.
86. Ise W, Haiker H, Weiss H (1985) Mitochondrial translation of subunits of the rotenone-sensitive NADH:ubiquinone reductase in Neurospora crassa. EMBO J. **4**, 2075-2080.
87. Jacobson MP, Pincus DL, Rapp CS, Day TJF, Honig B, Shaw DE, Friesner RA (2004) A Hierarchical Approach to All-Atom Protein Loop Prediction. Proteins. **55**, 351-367.
88. Jaokar TM, Sharma R, Suresh CG (2013) Structural Effects of Leigh Syndrome Mutations on the Function of Human Mitochondrial Complex-I Q module. Biochem. Physiol. S2 doi: 10.4172/2168-9652.S2-004.
89. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J. Am. Chem. Soc. **118**, 11225-11236.
90. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL (2001) Evaluation and reparametrization of the OPLA-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. J. Phy. Chem. B. **105**, 6474-6487.
91. Ke BX, Pepe S, Grubb DR, Komen JC, Laskowski A, Rodda FA, Hardman BM, Pitt JJ, Ryan MT, Lazarou M, Koleff J, Cheung MM, Smolich JJ, Thorburn DR (2012) Tissue-specific splicing of an Ndufs6 gene-trap insertion

generates a mitochondrial complex I deficiency-specific cardiomyopathy. Proc. Natl. Acad. Sci. USA **109**, 6165-6170.

92. Khurana R, Coleman C, Ionescu-Zanetti C, Carter S, Krishna V, Grover R, Roy R, Singh S (2005) Mechanism of Thioflavin T binding to amyloid fibrils. J. Struct. Biol. **151**, 229-238.
93. Klunk WE, Jacob RF, Mason RP (1999) Quantifying amyloid beta peptide (Abeta) aggregation using the Congo red-Abeta (CR-abeta) spectrophotometric assay. Anal Biochem. **266**, 66-76.
94. Kogelnik AM, Lott MT, Brown MD, Navathe SB, Wallace DC (1995) MITOMAP: A Human Mitochondrial Genome Database. Nucleic Acids Res. **24**, 177-179.
95. Kruse SE, Watt WC, Marcinek DJ, Kapur RP, Schenkman KA, Palmiter RD (2008) Mice with mitochondrial complex I deficiency develop fatal encephalomyopathy. Cell Metab. **7**, 312-320.
96. Kumar M, Kaur P, Kumar M, Saxena R, Sharma P, Dada R (2012) Clinical characterization and mitochondrial DNA sequence variations in Leber hereditary optic neuropathy. Mol. Vis. **18**, 2687-2699.
97. Kumar M, Tanwar M, Saxena R, Sharma P, Dada R (2010) Identification of novel mitochondrial mutations in Leber's hereditary optic neuropathy. Mol. Vis. **16**, 782-792.
98. Lacowicz EM, Weber G (1973) Quenching of protein fluorescence by oxygen. Detection of structural fluctuations in proteins on the nanosecond time scale. Biochemistry. **12**, 4171-4179.
99. Lakowicz JR (1983) Principles of fluorescence spectroscopy. Plenum, new York, p45.
100. Laskowski RA (2009) PDBSum new things. Nucleic Acid Res. **37**, D355-359.
101. Lazarou M, McKenzie M, Ohtake A, Thorburn DR, Ryan MT (2007) Analysis of assembly profiles for mitochondrial- and nuclear-DNA encoded subunits into complex I. Mol. Cell. Biol. **27**, 4228-4237.
102. Lebon S, Minai L, Chretein D, Corcos J, Serre V, Kadhom N, Steffann J, Pauchard JY, Munnich A, Bonnefont JP, Rotig A (2007a) A novel mutation of the NDUFS7 gene leads to activation of a cryptic exon and impaired assembly

of mitochondrial complex I in a patient with Leigh syndrome. Mol. Genet. Metab. **92**, 104-108.

103. Lebon S, Rodriguez D, Bridoux D, Zerrad A, Rotig A, Munnich A, Legrand A, Salma A (2007b) A novel mutation in the human complex I NDUFS7 subunit associated with Leigh syndrome. Mol. Genet. Metab. **90**, 379-382.
104. Lehrer SS (1971) Solute perturbation of protein fluorescence. The quenching of tryptophanyl fluorescence of model compounds and of lysozyme by iodide ion. Biochemistry. **10**, 3254-3263.
105. Lehrer SS, Leavis PC (1978) Solute quenching of protein fluorescence. Methods Enzymol. **49**, 222-236.
106. Leif H, Sled VD, Ohnishi T, Weiss H, Friedrich T (1995) Isolation and characterization of the proton translocating NADH: ubiquinone oxidoreductase from Escherichia coli. Eur. J. Biochem. **230**, 538-548.
107. Leigh D (1951) Sub acute necrotizing encephalomyelopathy in an infant. J. Neurol Neurosurg Psychiatry, **14**, 216-221.
108. LigPrep, version 2.5, Schrödinger, LLC, New York, NY, 2011.
109. Loeffen J, Elpeleg O, Smeitink J, Smeets R, Stöckler-Ipsiroglu S, Mandel H, Sengers R, Trijbels F, van den Heuvel L (2001) Mutations in the complex I NDUFS2 gene of patients with cardiomyopathy and encephalomyopathy, Ann. Neurol. **49**, 195–201.
110. Loeffen J, Smeitink J, Triepels R, Smeets R, Schuelke M, Sengers R, Trijbels F, Hamel B, Mullaart R, van den Heuvel L (1998a) The First Nuclear-Encoded Complex I mutation in a patient with Leigh syndrome. Am. J. Hum. Genet. **63**, 1598-1608.
111. Loeffen J, van den Heuvell L, Smeets R, Triepels R, Sengers R, Trijbels F, Smeitnik J (1998b) cDNA sequence and chromosomal localization of the remaining three human nuclear encoded iron sulphur protein (IP) subunits of complex I: the human IP fraction is completed. Biochem. Biophys. Res. Commun. **247**, 751-758.
112. Loeffen JL, Smeitinik JA, Trijbels JM, Janssen AJ, Triepels RH, Sengers RC, Van den Heuvel LP (2000) Isolated Zcomplex I deficiency in children: clinical, biochemical and genetic aspects. Hum. Mutat. **15**, 123-134.

113. Lopez LC, Schuelke M, Quinzii CM, Kanki T, Rodenburg RJ, Naini A, Dimauro S, Hirano M (2006) Leih syndrome with neuropathy and CoQ10 deficiency due to decaprenyl diphosphate synthase subunit 2 (PDSS2) mutations. Am. J. Hum. Genet. **79**, 1125-1129.
114. Maestro, version 9.3, Schrodinger, LLC, New York, NY, 2012.
115. Marin SE, Mesterman R, Robinson B, Rodenburg RJ, Smeitnik J, Tarnopolsky MA (2013) Leigh syndrome associated with mitochondrial complex I deficiency due to novel mutations in NDUFV1 and NDUFS2. Gene **516**, 162-167.
116. Marthino JMG, Santos AM, Fedorov A, Baptista RP, Taipa MA, Cabral JMS (2003) Fluorescence of a single tryptophan of cutinase: temperature and pH effect on protein conformation and dynamics. Photochem. Photobiol. **78**,15-22.
117. Martín MA, Blázquez A, Gutierrez-Solana LG, Fernández-Moreira D, Briones P, Andreu AL, Garesse R, Campos Y, Arenas J (2005) Leigh syndrome associated with mitochondrial complex I deficiency due to a novel mutation in the NDUFS1 gene. Arch. Neurol. **62**, 659–661.
118. Mimaki M, Wang X, McKenzie M, Thornburn DR, Ryan MT (2012) Understanding mitochondrial complex I assembly in health and disease. Biochim. Bhiophys. Acta **1817**,851-862.
119. MITOMAP: A Human Mitochondrial Genome Database. http://www.mitomap.org, 2013.
120. Mootha VK, Lepage P, Miller K, Bunkenborg J, Reich M, Hjerrild M, Delmonte T, Villeneuve A, Sladek R, Xu F, Mitchell GA, Morin C, Mann M, Hudson TJ, Robinson B, Rioux JD, Lander ES (2003) Identification of a gene causing human cytochrome c oxidase deficiency by integrative genomics. Proc. Natl. Acad. Sci. USA. **100**, 605-610.
121. Munnich A, Rustin P (2001) Clinical Spectrum and Diagnosis of Mitochondrial Disorders. Am. J. Med. Genet. **106**,4-17.
122. Nesbitt V, Morrison PJ, Crushell E, Donnelly DE,, Alston CL, He L, McFarland R, Taylor RW (2012) The clinical spectrum of the m.10191T>C mutation in complex I-deficient Leigh syndrome. Dev. Med. Child Neurol. **54**, 500-506.

123. Ngu LH, Nijtmans LG, Distelmaier F, Venselaar H. Van Emst-de Vries S, van den Brand M, Stoltenborg B, Willems PH, van den Heuvel LP, Smeitink JA, Rodenburg R (2012) A catalytic defect in the mitochondrial respiratory chain complex I due to a mutation in *NDUFS2* in a patient with Leigh syndrome. Biochim. Biophys. Acta **1822**, 168-175.

124. Nouws J, Nijtmans L, Houten SM, van den Brand M, Huynen M, Venselaar H, Hoefs S, Gloerich J, Kronick J, Hutchin T, Willems P, Rodenburg R, Wanders R, van den Heuvel L, Smeitink J, Vogel RO (2010) Acyl-CoA dehydrogenase 9 is required for the biogenesis of oxidative phosphorylation complex I. Cell Metab. **12,** 283–294.

125. Ogilvie I, Kennaway NG, Shoubridge EA (2005) A molecular chaperone for mitochondrial complex I assembly is mutated in a progressive encephalopathy. J. Clin. Invest.**115**, 2784–2792.

126. Oquendo CE, Antonicka H, Shoubridge EA, Reardon W, Brown GK (2004) Functional and genetic studies demonstrate that mutation in the COX15 gene can cause Leigh syndrome. J. Med. Genet. **41**, 540-544.

127. Ostergaard E, Rodenburg RJ, van den Brand M, Thomsen LL, Duno M, Batbayli M, Wibrand F, Nijtmans L (2011) Respiratory chain complex I deficiency due to NDUFA12 mutations as a new cause of Leigh syndrome. J. Med. Genet. **48**, 737–740.

128. Pagliarini DJ, calvo CE, Chang B. Sheth SA, Vafai SB, Ong SE, Walford GA, Sugiana C, Boneh A, Chen WK, Hill DE, Vidal M, Evans JG, Thorburn DR, Carr SA, Mootha VK (2008) A mitochondrial protein compendiumelucidates complex I disease biology. Cell. **134**, 112-123.

129. Pagnamenta AT, Hargreaves IP, Duncan AJ, Taanman JW, Heales SJ, Land JM, Bitner-Glindzicz M, Leonard JV, Rahman S (2006) Phenotypic variability of mitochondrial disease caused by nuclear mutation in complex II. Mol. Genet. Metab. **89**, 214-221.

130. Pagniez-Mammeri H, Lombes A, Brivet M, Ogier-de Baulny H, Landrieu P, Legrand A, Slama A (2009) Rapid screening for nuclear genes mutations in isolated respiratory chain complex I defects. Mol. Genet. Metab. **96**, 196–200.

131. Pagniez-Mammeri H, Loublier S, Legrand A, Benit P, Rustin P, Slama A (2012) Mitochondrial Complex I deficiency of nuclear origin I. Structural genes. Mol. Genet. Metab. **105**, 163-172.

132. Pequignot MO, Dey R, Zeviani M, Tiranti V, Godinot C, Poyau A, Sue C, Di Mauro S, Abitbol M, Marsac C (2001) Mutations in the SURF1 gene associated with the Leigh syndrome and cytochrome C oxidase deficiency. Hum. Mutat. **17**, 374-381.
133. Petruzzella V, Vergari R, Puzziferri I, Boffoli D, Lamantea E, Zeviani M, Papa S (2001)A nonsense mutation in the NDUFS4 gene encoding the 18 kDa (AQDQ) subunit of complex I abolishes assembly and activity of the complex in a patient with Leighlike syndrome. Hum. Mol. Genet. **10**, 529–535.
134. Potluri P, Davila A, Ruiz-Pesini E, Mishmar D, O'Hearn S, Hancock S, Simon M, Scheffler IE, Wallace DC, Procaccio V(2009) A novel NDUFA1 mutation leads to a progressive mitochondrial complex I-specific neurodegenerative disease. Mol. Genet. Metab. **96**, 189–195.
135. Prieur I, Lunardi J, Dupuis A (2001) Evidence for a quinine binding site close to the interface between NUOD and NUOB subunits of Complex I. Biochim. Biophys. Acta **1504**, 173-178.
136. Prime, version 3.1, Schrodinger, LLC, New York, NY 2012.
137. Procaccio V, Lescuyer P, Bourges I, Beugnot R, Duborjal H, Depetris D, Mousson B, Montfort MF, Smeets H, De Coo R, Issartel JP (2000) Human NDUFS3 gene coding for the 30-kDa subunit of mitochondrial complex I: genomic organization and expression. Mamm. Genome **11**, 808-810.
138. Procaccio V, Wallace DC (2004) Late-onset Leigh syndrome in a patient with mitochondrial Complex I NDUFS8 mutations. Neurology **62**, 1899-1901.
139. Prokisch H, Andreoli C, Ahting U, Heiss K, Ruepp A, Scharle C, Meitinger T (2006) MitoP2: the mitochondrial proteome database- now including mouse data. Nucleic Acids Res. **34**, D705-711.
140. Quintana A, Kruse SE, Kapur RP, Sanz E, Palmiter RD (2010) Complex I deficiency due to loss of Ndufs4 in the brain results in progressive encephalopathy resembling Leigh syndrome. Proc. Natl. Acad. Sci. USA **107**, 10996-101001.
141. Radermacher M, Ruiz T, Clason T, Benjamin S, Brandt U, Zickermann V (2006) The three-dimensional structure of complex I from *Yarrowia lipolytica*: a highly dynamic enzyme. J. Struct. Biol. **154**,269-279.

142. Rahman S, Blok RB, Dahl HH, Danks DM, Kirby DM, Chow CW, Christodoulou J, Thorburn DR (1996) Leigh syndrome: clinical features and biochemical and DNA abnormalities. Ann. Neurol. **39**, 343-351.

143. Rambaut A (2009) FigTree v1.3.1 2006-2009.

144. Remacle C, Barbieri MR, Cardol P, Hamel PP (2008) Eukaryotic complex I: functional diversity and experimental systems to unravel the assembly process. Mol. Genet. Genomics **280**, 93–110.

145. Reshentnyak YK, Koshevnik Y, Burstein EA (2001b) Decomposition of protein tryptophan fluorescence spectra into log normal components III. Correlation between fluorescence and microenvironment parameters of individual tryptophan residues. Biophys. J. **81**, 1735-1758.

146. Reshetnyak YK, Burstein EA (2001a), Decomposition of protein tryptophan fluorescence spectra into log normal components II. The statistical proof of discreteness of tryptophan classes in proteins. Biophys. J. **81**, 1710-1734.

147. Rhein VF, Caroll J, Ding S, Fearnley IM, Walker JE (2013) NDUFAF7 Methylates Arginine 85 in the NDUFS2 Subunit of the Human Complex I. J. Biol. Chem. **288**, 33016-33026.

148. Ripple MO, Kim N, Springett R (2013) Mammalian complex I pumps 4 protons per 2 electrons at high and physiological proton motive force in living cells. J. Biol. Chem. **22**, 5374-5380.

149. Rossi A, Biancheri R, Bruno C, Di Rocco M, Calvi A, Pessagno A, Tortori-Donati P (2003) Leigh syndrome with COX deficiency and SURF1 gene mutations: MR imaging findings. ANJR Am. J. Neuroradiol. **24**, 1188-1191.

150. Rossignaol R, Faustin B, Rocher C, Malgat M, Mazat JP and Letellier T (2003) Mitochondrial Threshold effects. Biochem J. **370**,751-762.

151. Saada A, Vogel RO, Hoefs SJ, van den Brand MA, Wessels HJ, Willems PH, Venselaar H, Shaag A, Barghuti F, Reish O, Shohat M, Huynen MA, Smeitink JA, van den Heuvel LP, Nijtmans LG (2009) Mutations in NDUFAF3 (C3ORF60), encoding an NDUFAF4 (C6ORF66)-interacting complex I assembly protein, cause fatal neonatal mitochondrial disease. Am. J. Hum. Genet. **84**, 718–727.

152. Saada A, Edvardson S, Rapoport M, Shaag A, Amry K, Miller C, Lorberboum-Galski H, Elpeleg O (2008) C6ORF66 is an assembly factor of mitochondrial complex I. Am. J. Hum. Genet. **82**, 32–38.

153. Saada A, Edvardson S, Shaag A, Chung WK, Segel R, Miller C, Jalas C, Elpeleg O (2012) Combined OXPHOS complex I and IV defect, due to mutated complex I assembly factor C20ORF7. J. Inherit. Metab. Dis. **35**, 125-131.
154. Sazanov LA, Hinchliffe P (2006) Structure of the Hydrophilic domain of the Respiratory Complex I from *Thermus thermophilus*. Science. **311**, 1430-1436.
155. Schmiedel J, Jackson S, Schäfer J, Reichmann H (2003) Mitochondrial Cytopathies. J. Neurol. **250**, 267–277.
156. Schuelke M, Smeitink J, Mariman E, Loeffen J, Plecko B, Trijbels F, Stöckler-Ipsiroglu S, van den Heuvel L (1999) Mutant NDUFV1 subunit of mitochondrial complex I causes leukodystrophy and myoclonic epilepsy. Nat. Genet. **21**, 260–261.
157. Schulte U, Fecke W, Krüll C, Nehls U, Schmiede A, Schneider R, Ohnishi T, Weiss H (1994) In vivo dissection of the mitochondrial respiratory NADH: ubiquinone oxidoreductase (complex I). Biochim. Biophys. Acta **1187**, 121-124.
158. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L (2005) The FoldX web server: an online force field. Nucleic Acids Res. **33**, W382-8.
159. Semisotnov GV, Radionova NA, Razgulyaev OI, Uversky VN, Gripas AF, Gilmanshin RI (1991) Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. Biopolymers. **31**, 119-128.
160. Sgarbi G, casalena GA, Baracca A, Lenaz G, DiMauro S, Solaini G (2009) Human NARP mitochondrial mutation metabolism corrected with alpha-ketoglutarate/aspartate : a potential new therapy. Arch. Neurol. **66**, 951-957.
161. Shen C, Menon R, Das D, Bansal N, Nahar N, Guduru N, Jaegle S, Peckham J, Reshetnyak YK (2008) The protein fluorescence and structural toolkit: Database and programs for the analysis of protein fluorescence and structural data. Proteins **71**, 1744-1754.
162. Shrikhande DY, Kalakoti P, Syed AMM, Ahya K, Singh G (2010) A rare mitochondrial disorder: Leigh syndrome- a case report. Italian Journal of Pediatrics **36**, 62-66.
163. Smeitink J, van den Heuvel L (1999) Human mitochondrial complx I in health and disease. Am. J. Hum. Genet. **64**,1505-1510.

164. Smeitink JAM, Loeffen JLCM, Triepels RH, Smeets RJP, Trijbels JMF and Van den Heuvel LP (1998) Nuclear genes of human complex I of the mitochondrial electron transport chain: state of the art. Hum. Mol. Genet. **7**,1573-1579.
165. Spiegel R, Shaag A, Mandel H, Reich D, Penyakov M, Hujeirat Y, Saada A, O. Elpeleg O, Shalev SA (2009) Mutated NDUFS6 is the cause of fatal neonatal lactic academia in Caucasus Jews. Eur. J. Hum. Genet. **17,** 1200–1203.
166. Sreerama N, Woody RW (2004) Computation and analysis of protein circular dichroism spectra. Methods Enzymol. **383**, 318-351.
167. Studier FW (2005) Protein production by auto-induction in high density shaking cultures. Protein Expr. Purif. **41**, 207-234.
168. Su JT, Kim SH, Yan YB (2007) Dissecting the pretransitional conformational changes in aminoacylase I thermal denaturation. Biophys. J. **92**, 578-587.
169. Sugiana C, Pagliarini DJ, McKenzie M, Kirby DM, Salemi R, Abu-Amero KK, Dahl HH, Hutchison WM, Vascotto KA, Smith SM, Newbold RF, Christodoulou J, Calvo S, Mootha VK, Ryan MT, Thorburn DR (2008) Mutation of C20orf7 disrupts complex I assembly and causes lethal neonatal mitochondrial disease. Am. J. Hum. Genet. **83**,468–478.
170. Suhane S, Berel D, Ramanujan VK (2011) Biomarker signatures of mitochondrial NDUFS3 in invasive breast carcinoma. Biochem. Biophys. Res. Commun. **412**, 590-595.
171. Suhane S, Kanazaki H, Arumugaswami V, Murali R, Ramanujan VK (2013) Mitochondrial NDUFS3 regulates the ROS-mediated onset of metabolic switch in transformed cells. Biol. Open **2**, 295-305.
172. Suite 2012: BioLuminate 1.0 , LLC, New York, NY, 2012.
173. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolution genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. Mol. Biol. Evol. **28**, 2731-2739.
174. Tanaka M, Borgeld HJ, Zhang J, Muramatsu S, Gong JS, Yoneda M, Maruyama W, Naobi M, Ibi T, Sahashi K, Shamoto M, Fuku N, Kurata M, Yamada Y, Nishizawa K, Akao Y, Ohishi N, Miyabayashi S, Umemoto H,

Muramatsu T, Furukawa K, Kikuchi A, Nakano I, Ozawa K, Yagi K (2002) Gene therapy for mitochondrial disease by delivering restriction endonuclease SmaI in mitochondria. J. Biomed. Sci. **9**, 534-541.

175. Taylor RW, Schaefer AM, Barron MJ, McFarland R, Turnbull DM (2004) The diagnosis of mitochondrial muscle disease. Neuromuscul. Disord. **14**, 237–245.
176. The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.
177. The Pymol molecular Graphics system, Version 1.5.04 Schrödinger, LLC.
178. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**, 4673-4680.
179. Thorburn DR (2004) Mitochondrial disorders: prevalence, myths and advances. J. Inherit. Metab. Dis. **27**, 349-362.
180. Thorburn DR, Rahman S (2011) Mitochondrial DNA-Associated Leigh syndrome and NARP. GeneReviews™ .
181. Thorburn DR., Sugiana C, Salemi R, Kirby DM., Worgan L, Ohtake A and Ryan MT (2004) Biochemical and molecular diagnosis of mitochondrial respiratory chain disorders. Biochim. Biophys. Acta **1659**, 121– 128.
182. Tocilescu M, Zickermann V, Zwicker K, Brandt U (2010) Quinone binding and reduction by respiratory complex I. Biochim. Biophys. Acta **1797**,1883-1890.
183. Towbin H, Staehelin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA. **76**, 4350-4354.
184. Triepels RH, van den Heuvel LP, Loeffen JL, Buskens CA, Smeets RJ, Rubino Gozalbo ME, Budde SM, Mariman EC, Wijburg FA, Barth PG, Trijbels JM, Smeitink JA (1999) Leigh syndrome associated with a mutation in the NDUFS7 (PSST) nuclear encoded subunit of complex I. Ann. Neurol. **45**,787-790.
185. Tuppen HA, Hogan VE, He L, Blackely EL, Worgan L, Al-Dosary M, Saretzki G, Alston CL, Morris AA, Clarke M, Jones S, Devlin AM, MMansour S, Chrzanowska-Lightowlers ZM, Thorburn DR, McFarland R,

taylor RW (2010) The p.M292T NDUFS2 mutation causes complex-I deficient Leigh syndrome in multiple families. Brain **133**, 2952-2963.

186. Ueki I, Koga Y, Povalko N, Akita Y, Nishioka J, Yatsuga S, Fukiyama R and Matsuishi T (2006) Mitochondrial tRNA gene mutations in patients having mitochondrial disease with lactic acidosis. Mitochondrion **6**, 29–36.
187. Ugalde C, Hinttala R, Timal S, Smeets R, Rodenburg RJ, Uusimaa J, van Heuvel LP, Nijtmans LG, Majamaa K, Smeitink JA (2007) Mutated Nd2 impairs mitochondrial complex I assembly and leads to Leigh syndrome. Mol. Genet. Metab. **90**, 10-14.
188. Ugalde C, Janssen RJ, Heuvel LP, Smeitink JA, Nijtmans LG (2004a) Differences in assembly or stability of complex I and other mitochondrial OXPHOS complexes in inherited complex I deficiency. Hum. Mol. Genet. **13**, 659-667.
189. Ugalde C, Vogel R, Huijbens R, Heuvel BV, Smeitink J, Nijtmans L (2004b) Human mitochondrial complex I assembles through the combination of evolutionary conserved modules: a framework to interpret complex I deficiencies. Hum. Mol. Genet. **13**, 2461-2472.
190. Van den Ecker D, van den Brand MA, Ariaans G, Hoffmann M, Bossinger O, Mayatepek E, Nijtmans LG, Distelmaier F (2012) Identification and functional analysis of mitochondrial complex I assembly factor homologues in C elegans. Mitochondrion **12,** 399-405.
191. Van Maldergem L, Trijbels F, DiMauro D, Sindelar PJ, Musumeci O, Janssen A, Delberghe A, Martin JJ, Gillerot Y (2002) Coenzyme Q responsive Leigh's encephalopathy in two sisters. Ann. Neurol. **52**, 750-754.
192. Videira A (1998) Complex I from fungus *Neurospora crassa*. Biochim Biophy. Acta. **1364**, 89-100.
193. Vilain C, Rens C, Aeby A, Baleriaux D, Van Bogaert P, Remiche G, Smet J, Van Coster R, Abramowicz M, Pirson I (2012) A novel NDUFV1 gene mutation in complex I deficiency in consanguineous siblings with brainstem lesions and Leigh syndrome. Clin. Genet. **82**, 264-270.
194. Visch HJ, Rutter GA, Koopman WJH, Koenderink JB, Verkaart S, de Groot T, Varadi A, Mitchell KJ, van den Heuvel LP, Smeitink JAM, Willems PHGM (2004) Inhibition of mitochondrial Na+-Ca2+ exchange restores agonist-

induced ATP production and Ca2+ handing in human complex I deficiency. J. Biol. Chem. **279**, 40328-40336.

195. Vogel RO, Dieteren CE, Van den Heuvel LP, Willems PH, Smeitink JA, Koopman WJ, Nijtmans LG (2007) Identification of mitochondrial complex I assembly intermediates by tracing tagged NDUFS3 demonstrates the entry point of mitochondrial subunits. J. Biol. Chem. **282**, 7582-7590.
196. Vogel RO, Janssen RJ, Ugalde C, Grovenstein M, Huijbens RJ, Visch HJ, van denHeuvel LP, Willems PH, Zeviani M, Smeitink JA, Nijtmans LG (2005) Human mitochondrial complex I assembly is mediated by NDUFAF1. FEBS J. **272**, 5317–5326.
197. Wallace DC (2013) A mitochondrial bioenergetic etiology of disease. J. Clin. Invest. **123**, 1405-1412.
198. Wang DC, Meinhardt SW, Sackmann U, Weiss H, Ohnishi T (1991) The iron-sulphur clusters in the two related forms of mitochondrila NADH: ubiquinone oxidoreductase made by *Neurospora crassa*. Eur. J. Biochem. **197**, 257-264.
199. Wang K, Takahashi Y, Gao ZL, Wang GX, Chen XW, Goto J, Lou JN, Tsuji S (2009) Mitochondrial ND3 as the novel causative gene for Leber hereditary optic neuropathy and dystonia. Neurogenetics. **10**, 337-345.
200. Weidner U, Geier S, Ptoock A, Friedrich T, Leif H, Weiss H (1993) The Gene locus of the proton-translocating NADH:ubiquinone oxidoreductase in *Escherichia coli*. Organization of the 14 genes and relationship between the derived proteins and subunits of the mitochondrial complex I. J. Mol. Biol. **233**, 109-122.
201. Weiss H, Friedrich T, Hofhaus G, Preis D (1991) The respiratory chain NADH dehydrogenase(complex I) of mitochondria. Eur. J. Biochem. **197**, 563-576.
202. Wiederstein M, Sippl MJ (2007) Pro-SA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acid Res. **35**, W407-410.
203. Zickermann V, Kerscher S, Zwicker K, Tocilescu MA, Radermacher M, Brandt U (2009) Architecture of complex I and iits implications for electron transfer and proton pumping. Biochim. Biophys. Acta **1787**, 574-583.

List of Publications

1. **Tulika M. Jaokar,** Ranu Sharma and C.G. Suresh. Structural effects of Leigh syndrome mutations on the function of human mitochondrial Complex-I Q module. (2013) Biochem. Physiol. S2 doi: 10.4172/2168-9652.S2-004.

2. **Tulika M. Jaokar,** Deepak P. Patil, Yogesh S. Shouche, Sushama M Gaikwad, C.G. Suresh. Human Mitochondrial *NDUFS3* protein bearing Leigh syndrome mutation is more prone to aggregation than its wild-type. (2013) Biochimie. 95:2393-2403.

3. Wakankar MS, Krishnasastry MV, **Jaokar TM**, Patel KA, Gaikwad SM. Solution and in silico studies on the recombinant lectin from *Cicer arietenum* seeds. (2013) Int J Bio Macromol. 56:149-155.

4. **Tulika M Jaokar**, Deepak P Patil, Ranu Sharma, Yogesh S. Shouche, Sushama M Gaikwad, C.G. Suresh. Characterization and binding studies of *NDUFS7* subunit and its mutant causing Leigh syndrome. (manuscript under preparation)