The Molecular and Biochemical Characterization of the MLRQ Subunit of NADH:Ubiquinone Oxidoreductase in the Human Mitochondrial Respiratory Chain

by

Dhushy Kanagarajah

A Thesis submitted in conformity with the requirements for the degree of Master of Science
Graduate Department of Biochemistry
University of Toronto

© Copyright by Dhushy Kanagarajah. 2001
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-58735-5
The Molecular and Biochemical Characterization of the MLRQ Subunit of NADH:Ubiquinone Oxidoreductase in the Human Mitochondrial Respiratory Chain

Master of Science, 2001
Dhushy Kanagarajah
Department of Biochemistry
University of Toronto

Abstract
Isolated deficiency of NADH:ubiquinone oxidoreductase (Complex I), the first enzyme of the mitochondrial respiratory chain is the most common cause of human mitochondriopathies. In order to characterize the nuclear genes contributing to this disease, the cDNA and genomic sequences encoding the MLRQ subunit of complex I were determined. The NDUF.44 gene encoding MLRQ was localized to chromosome 7 p21-22 and a pseudogene was found on chromosome 1 p21. Tissue specific expression of MLRQ at both mRNA and protein levels was examined. Overexpression of this subunit in a patient exhibiting complex I deficiency is also discussed. Extraction, immunoprecipitation and cross-linking studies revealed that while the N-terminus of MLRQ has a great affinity for phospholipids of the inner mitochondrial membrane, it likely also associates with the MWFE subunit through other intermediary subunit(s), forming part of the bulky stalk region that bridges the two arms of complex I.
Acknowledgements

First and foremost, I wish to express my heartfelt thanks and gratitude to Dr. Brian Robinson for his guidance and support during the course of this degree. I truly feel privileged to have had the opportunity of working with such a great supervisor. I am also grateful to my co-supervisors, Dr. B. Sarkar and Dr. R. Baker for their time and assistance with this project.

Behind every graduate student’s thesis lies a support system so precious that failing to acknowledge it would be unpardonable. I count myself lucky to have had the privilege of working with all my colleagues (past and present) at the Robinson lab, but I must especially thank Agnieszka, Jessie, Maryanna, Nevi, So-Young and Tomoko. Their kindness and friendship will never be forgotten. I am forever indebted to Dr. Sandy Raha for all his expertise, encouragement and humour without which I would never have survived. A very special thanks goes to Maureen Waite for her friendship and for always making the time to lend me a hand.

This thesis would not have been completed without the extraordinary love and support of each and every person whom I call family. My most ardent supporters, they have been with me throughout the trials and tribulations of graduate life. For this, I would like to thank my uncle Milroy, my brother Dhilip and especially my husband Ramana for his patience and understanding during these past few years. Finally but most importantly, I would like to express my gratitude to my parents for impressing upon me the principles of hard work, perseverance and a firm belief in the merits of education. Their love and prayers have been instrumental in the completion of this study.

I dedicate this thesis to the memory of my loved ones whose blessings and encouragement still spur me on to greater accomplishments.
CONTRIBUTIONS TO THESIS

Screening of PAC/YAC libraries and FISH Mapping: The Toronto Centre For Applied Genomics

Tissue mitochondria: Dr. Sandeep Raha
Table of Contents

Abstract i
Acknowledgements ii
Contributions to thesis iii
Table of contents iv
List of figures viii
List of tables x
Abbreviations xi
Glossary of Medical Terms xiii

Chapter 1
Introduction and Objectives
The mitochondrion: structure, function, mode of inheritance and associated diseases
Overview 1
Part I. Mitochondrial structure, function and inheritance 2

The mitochondrion 2
Mitochondrial ultrastructure 3
Mitochondrial DNA organization 4
Mitochondrial replication, transcription and translation 6
Mitochondrial protein import 8
Energy metabolism 12

Part II. The mitochondrial respiratory chain complexes 14

Organization of the OXPHOS system 14
The OXPHOS system: Role in electron transport and proton translocation 15
Complex I: The NADH:ubiquinone oxidoreductase complex 17

Evolution of complex I 17
Subunit composition of the NADH:ubiquinone oxidoreductase complex 20
(i) The flavoprotein fraction (FP) 21
(ii) The iron-sulphur protein fraction (IP) 24
(iii) The hydrophobic protein fraction (HP) 25
a) The mitochondrially encoded subunits 25
b) The nuclear encoded subunits 27

Structural model of complex I 29
(i) Assembly of Complex I 31
(ii) Spatial Organization and Subunit Interaction in Human Complex I 32

Energy conversion in complex I 35
(i) Iron-sulphur clusters, flavin and semiquinones 35
(ii) Electron transfer in complex I 37
(iii) Models for coupling electron flow with proton translocation 42
Complex I inhibitors 47
Complex II: The Succinate-ubiquinone oxidoreductase complex 48
Complex III: The Ubiquinol-ferricytochrome c oxidoreductase complex 49
Complex IV: The cytochrome c oxidase complex 51
Complex V: The ATP synthase complex 52

Part III. Mitochondrial disorders 54

Typical symptoms of defects in energy metabolism 54
Mitochondrial respiratory chain diseases 56
(i) MtDNA associated diseases 56
(ii) Nuclear DNA associated diseases 57
(iii) Mitochondrial respiratory chain disorders associated with neurological diseases 58

Human Complex I deficiencies 59
(i) MtDNA encoded defects in complex I 59
(ii) Nuclear DNA encoded defects in complex I 60
(iii) Free radical generation and complex I deficiency 60

Objectives and Rationale 62

Chapter 2
Cloning, molecular characterization and chromosomal localization of the MLRQ subunit of human NADH:ubiquinone oxidoreductase

Abstract 64
Introduction 64
Materials and methods 69
Part I. Molecular Characterization of MLRQ cDNA in various tissues, cells and patient cell lines 69
Tissue culture of cardiomyocytes and fibroblasts 69
RNA isolation from tissues and cells 69
cDNA synthesis and PCR 69
cDNA cloning and sequencing of MLRQ 70
Mutational screening of complex I deficient patients 71

Part II. Genomic characterization and localization of the NDUFA4 (MLRQ) gene and pseudogene 72
Screening the PAC library 72
Southern blot analysis of PAC clones 72
Northern analysis of MLRQ expression 73
Chromosomal localization 73
Amplification, cloning and sequencing of MLRQ from genomic and PAC DNA 73
YAC library screening 74

Results and discussion 74
Part I. Isolation and characterization of MLRQ cDNA 74
Chapter 3
Biochemical characterization, protein expression and immunoprecipitation studies pertaining to MLRQ and related complex I subunits

Abstract 94

Introduction 95

Materials and methods 96

Part I. MLRQ expression in human tissues and cells 96
  Antibody generation 96
  Western blot analysis of MLRQ expression 97

Part II. Bacterial expression of MLRQ protein 98
  Design of MLRQ-fusion protein construct 98
  Induction and purification of MLRQ-GST fusion protein 99
  Factor Xa cleavage of fusion protein 99

Part III. Anti-sense expression of MLRQ in mammalian cells 100
  Design of sense and anti-sense oriented pREP9 constructs 100
  Optimization of transfection conditions 101
  Transfection and selection with pREP9 101
  Transfection and selection with the linearized vector pCDNA 3.1+ 102
  Western blot analysis of sense anti-sense expression in transfected cells 102
  Amplification of MLRQ from transfected cells 103

Part IV. Association of MLRQ with other complex I subunits 103
  Solubilization of beef heart mitochondria 103
  Immunoprecipitation of MLRQ, MWFE and 49 kD subunits 104
  Cross-linking with DST and EGS 105
  Immunoprecipitation of cross-linked bovine heart mitochondria 105
  SDS-PAGE and western blot analysis of MLRQ during extraction, immunoprecipitation and cross-linking 106

Results and discussion 106

Part I. Determining MLRQ form, function and expression 106
  Tissue expression of MLRQ 106
  Bacterial expression of MLRQ: Attempts at defining subunit structure 110
Antisense expression of MLRQ: Attempts to determine subunit function 113
Part II. Subunit interactions of MLRQ within complex I 117
Detergent solubilization of complex I subunits 117
Proximity and association of MLRQ with other complex I subunits 122
Cross-linking and immunoprecipitation studies 128

Chapter 4
Postulating the role of a supernumerary subunit such as MLRQ in complex I function and dysfunction
Conclusions and Future directions

Part I. Molecular structure of MLRQ 135
Part II. Biochemical characterization of MLRQ 137
Part III. Localization of MLRQ within complex I 138
Part IV. Possible roles for MLRQ in complex I function 140
The final word 143

References 145
List of figures

Chapter 1.
1.1 Electron micrograph of a mammalian mitochondrion 3
1.2 Mitochondrial structure and membrane organization 5
1.3 Organization of the mitochondrial DNA genome 7
1.4 The protein import machinery of mitochondria 11
1.5 The glucose oxidation pathway 13
1.6 Mitochondrial respiratory chain and ATP-synthase 15
1.7 Three dimensional models of complex I from E. coli, N. crassa and B. taurus as determined by electron cryo-microscopy 30
1.8 Model of the overlap in subunit composition between the different subcomplexes of complex I 33
1.9 Structural model of complex I 41
1.10 A hypothetical model for direct energy conversion in Complex I 44

Chapter 2.
2.1 The nucleotide sequence of the human MLRQ subunit cDNA and its deduced amino acid sequence 75
2.2 Alignment of the predicted human MLRQ subunit protein sequence with that of bovine and mouse 76
2.3 Schematic of NDUFA4 structure and chromosomal allocation of the gene and its pseudogenes 79
2.4 Sequence of the MLRQ pseudogene on PAC 2F23 81
2.5 Northern blot analysis of NDUFA4 transcripts in normal human tissues 82
2.6 PCR amplification of NDUFA4 from genomic DNA 84
2.7 Southern analysis of PAC clones 2F23 and 96E24 85
2.8 Regulatory motifs and putative transcription factor binding sites in the 5' lower part of the NDUFA4 gene 88
2.9 Comparison of the MLRQ cDNA sequence with the pseudogene sequences from PACs 2F23 and 69E11 in the region showing greatest homology 92

Chapter 3.
3.1 Tissue specific expression of the MLRQ subunit of complex I 107
3.2 Western blot analysis on mitochondria isolated from cultured fibroblasts of a patient (5624-HT) and control (4212) using various complex I antibodies 109
3.3 Bacterial expression of the MLRQ subunit as a GST-fusion protein 111
3.4 Purification and cleavage of GST-MLRQ fusion protein 112
3.5 Western blot analysis of MLRQ expression in SV40 immortalized fibroblasts transfected with sense and anti-sense pCNDA 3.1 constructs 115
3.6 PCR amplification of sense and anti-sense MLRQ sequences to confirm transfection of fibroblasts 116
3.7 Extraction of the MLRQ subunit from bovine heart mitochondria ........................................ 118
3.8 Solubilization of the MLRQ subunit compared to other complex I subunits ....................... 120
3.9 Hydropathy profile and membrane orientation of the MLRQ polypeptide ............................ 121
3.10 Immunoprecipitation studies using protein A agarose ......................................................... 123
3.11 Immunoblotting of the immunoprecipitated MLRQ, MWFE and 49 kDa subunits of complex I to determine subunit association ................................................................. 125
3.12 Immunoblots of 1mM DST and 0.2mM EGS cross-linked bovine heart mitochondria with antibodies to the MLRQ and 49 kDa subunits ................................................................. 127
3.13 Immunoblotting of EGS cross-linked bovine heart mitochondria with complex I antibodies ................................................................................................................................. 129
3.14 Immunoblotting of EGS cross-linked bovine heart mitochondria immunoprecipitated with MLRQ and MWFE antibodies ............................................................................... 133

Chapter 4.
4.1 Schematic representation of the proximity of MLRQ to other subunits of complex I .......... 139


## List of tables

### Chapter 1.

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Nomenclature and properties of homologous complex I subunit genes of <em>E. coli</em>, <em>B. taurus</em> and <em>H. sapiens</em></td>
<td>19</td>
</tr>
<tr>
<td>1.2</td>
<td>Current molecular genetic knowledge of human nuclear-encoded subunits of complex I of the mitochondrion</td>
<td>22</td>
</tr>
<tr>
<td>1.3</td>
<td>Current hypotheses on the subunit location of FMN and iron-sulphur (Fe-S) clusters in the minimal nuclear-encoded functional unit of bovine heart Complex I</td>
<td>37</td>
</tr>
<tr>
<td>1.4</td>
<td>Composition and genetic origin of mitochondrial respiratory chain (OXPHOS) subunits</td>
<td>54</td>
</tr>
<tr>
<td>1.5</td>
<td>The clinical presentation and incidence of isolated complex I deficiency attributed to nuclear encoded defects</td>
<td>61</td>
</tr>
</tbody>
</table>

### Chapter 2.

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Nuclear gene mutations in patients with isolated complex I deficiency</td>
<td>66</td>
</tr>
<tr>
<td>2.2</td>
<td>Summary of clinical information on patients screened for mutations in the cDNA of <em>NDUFA4</em></td>
<td>71</td>
</tr>
<tr>
<td>2.3</td>
<td>Exon-intron splice junctions of the human <em>NDUFA4</em> gene</td>
<td>87</td>
</tr>
</tbody>
</table>

### Chapter 3.

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Subunit proximity in complex I as determined by immunoprecipitation Studies</td>
<td>126</td>
</tr>
<tr>
<td>3.2</td>
<td>Cross-linked products detected by immunodetection with various complex I antibodies</td>
<td>134</td>
</tr>
</tbody>
</table>
Abbreviations

A  adenine
aa  amino acid
ATP  adenosine triphosphate
ATPase  adenosine triphosphate synthase
bp  base pairs
BCIP  5-bromo-4-chloro-3-indolylphosphate p-toluidine salt
C  cytosine
cAMP  cyclic adenosine monophosphate
CC  cardiomyopathy and cataracts
CD  circular dichroism
CD  cataracts and developmental delay
cDNA  DNA complementary to RNA
CHAPS  3-[(3-cholamidopropyl)dimethylammonio]-1-propan-sulfonat
COX  cytochrome c oxidase
CPEO  chronic progressive external ophthalmoplegia
Da  daltons
DCCD  N,N'-dicyclohexylcarbodiimide
DDM  n-dodecyl-β-D-maltoside
DNA  deoxyribonucleic acid
DST  disuccinimidyl tartrate
DTT  dithiothreitol
EDTA  ethylenediaminetetra-acetate
EGS  ethylene glycolbis(succinimidylsuccinate)
EPR  electron paramagnetic resonance
EST  expressed sequence tag
FAD  flavin adenine dinucleotide
FeS  iron sulfur center
FILA  fatal infantile lactic acidosis
FISH  fluorescence in situ hybridization
FMN  flavin mononucleotide
FP  flavoprotein
G  guanine
GTP  guanosine triphosphate
HQNO  2-n-heptyl-4-hydroxyquinoline N-oxide
HT  hepatopathy and tubulopathy
HTGS  high throughput genome sequence
HP  hydrophobic protein
IgG  immunoglobulin G
IP  iron-sulfur protein
IPTG  isopropylthiogalactoside
kb  kilo base pair
kDa  kilodaltons
KSS  Kearns-Sayre syndrome
LD  Leigh's disease
LDAO  lauryldimethylamine oxide
LDH  lactate dehydrogenase
LHON  Leber's hereditary optic neuropathy
L/P  lactate to pyruvate ratio
MELAS  mitochondrial encephalomyelopathy with lactic acidosis and stroke-like episodes
MERRF  myoclonus epilepsy with ragged red fibres
MM  mitochondrial myopathy
MMC  myopathy and cardiomyopathy
MNGIE  mitochondrial neurogastrointestinal encephalomyopathy
mRNA  messenger ribonucleic acid
MS  mild symptoms
mtDNA  mitochondrial deoxyribonucleic acid
NaCl  sodium chloride
NAD  nicotinamide adenine dinucleotide oxidized form
NADH  nicotinamide adenine dinucleotide reduced form
NADPH  nicotinamide adenine dinucleotide phosphate reduced form
NARP  neurogenic muscular weakness, ataxia and retinitis pigmentosa
NBT  para-nitro-blue tetrazolium chloride
nDNA  nuclear deoxyribonucleic acid
NMR  nuclear magnetic resonance
Oligo  oligodeoxyribonucleotide
ORF  open reading frame
OXPHOS  oxidative phosphorylation system
PAC  P1-artificial chromosome
PAGE  polyacrylamide gel electrophoresis
PCR  polymerase chain reaction
PD  Parkinson's disease
Q  ubiquinone
QFR  menaquinol-fumarate oxidoreductase
QH₂  ubiquinol
RNA  ribonucleic acid
rpm  rounds per minute
rRNA  ribosomal ribonucleic acid
S  svedberg unit
SDS  sodium dodecyl sulfate
SMP  submitochondrial particles
SQ  semiquinone
SQR  succinate-ubiquinone oxidoreductase
T  thymine
tRNA  transfer ribonucleic acid
TFFA  theonyltrifluoro-acetone
UTR  untranslated region(s)
YAC  yeast artificial chromosome
Glossary of medical terms

**Alzheimer’s disease** – A progressive, neurodegenerative disease characterized by loss of function and death of nerve cells in several areas of the brain leading to loss of cognitive function such as memory and language.

**Ataxia** – Defective muscular coordination affecting balance, gait, limb or eye movements.

**Basal ganglia disease** – Disease of the three large subcortical nuclei of the vertebrate brain that participate in the control of movement.

**Cardiomyopathy** – A general diagnostic term designating primary myocardial disease. Often of obscure or unknown aetiology.

**Cataracts** – An ocular opacity, partial or complete, of one or both eyes, on or in the lens or capsule, especially an opacity impairing vision or causing blindness.

**Chronic progressive external ophthalmoplegia** – Disorder where there is a progressive weakness of the extraocular muscles, eventually leading to a complete ophthalmoplegia.

**Dystonia** – Disordered tonicity of muscle.

**Encephalomyelopathy** – Any disease involving the brain and spinal cord.

**Encephalopathy** – Any degenerative disease of the brain.

**Epilepsy** – Recurring disorder characterized by sudden seizure activity or temporary alterations of one or more brain functions arising from abnormal electrical brain activity.

**Familial megalencephaly** – An inherited disorder where the patient exhibits an enlargement of the head caused by blockage of outflow of cerebrospinal fluid.

**Fatal infantile lactic acidosis** – Acidosis caused by accumulation of lactic acid more rapidly than it can be metabolized causing fatality in infants.

**Friedreich’s ataxia** – An autosomal recessive inherited disorder that leads to the progressive dysfunction of the cerebellum, spinal cord and peripheral nerves. Symptoms consist of an unsteady gait (ataxia), slurred speech and jerky eye movements.

**Hepatopathy and renal tubulopathy** – Enlargement of liver accompanied by disorders of the reabsorptive functions of the kidney with regard to specific nephron segments responsible for specific transport functions.

**Hyperventricular cardiomyopathy** – Myocardial disease where there is an enlargement of the ventricles in the heart.
**Kearns-Sayre syndrome** – Phenotype associated with single deletions of mtDNA. Core clinical features are a progressive weakness of the muscle which moves the eyes (CPEO) and pigmentary retinopathy.

**Lactic acidemia** – Condition of high blood lactate resulting from an inborn error of metabolism.

**Leigh’s disease** - A disease of pyruvate metabolism manifesting in infancy with psychomotor retardation, dysphagia, hypotonia, ataxia, weakness, external ophthalmoplegia, vision loss, hearing loss, and convulsions.

**Leukodystrophy** - An inherited metabolic disorder of the nervous system, particularly the white matter.

**Myoclonus** – Twitching or spasm of a muscle or group of muscles.

**Neurogenic** - Arising from or caused by the nervous system.

**Neuropathy** – Disease involving inflammation or damage to the peripheral nerves.

**Ophthalmoplegia** – Paralysis of the ocular muscles

**Parkinson’s disease** – A progressive neurological disease where symptoms include shuffling gait, stooped posture, resting tremor, speech impediments, movement difficulties and an eventual slowing of mental processes and dementia.

**Retinitis pigmentosa** – Disease caused by overactivity of the pigmented retinal epithelial cells, leading to damage and occlusion of photoreceptors and blindness.

**Wilson’s disease** - An inherited (autosomal recessive) disorder where there is excessive quantities of copper in the tissues, particularly the liver and central nervous system.
Chapter 1

Introduction and objectives
The mitochondrion: structure, function, mode of inheritance and associated diseases

Overview

In recent years, mitochondrial defects have been implicated as playing a role in a wide range of degenerative diseases, aging and cancer. Although studies on various human disorders resulting from mitochondrial dysfunction have given some insight into the complexities of mitochondrial genetics, the pathophysiology of mitochondrial diseases remains a perplexing problem due to the interplay between the mitochondrial and nuclear genomes. The essential role of mitochondrial oxidative phosphorylation in cellular energy production, the generation of reactive oxygen species, and the initiation of apoptosis has suggested a number of novel mechanisms for mitochondrial pathology. The importance and interrelationship of these pathways are now being studied. In order to illustrate these interrelationships, this section begins with the examination of mitochondrial structure, DNA organization and protein import into the organelle. It then proceeds to look at the function of the different complexes of the respiratory chain with particular emphasis on the first enzyme of the chain namely, NADH:ubiquinone oxidoreductase or complex I. In order to establish a solid foundation for the concepts and ideas that are presented in succeeding chapters, a thorough discussion of the subunit composition, structural model and energy conversion pathways of complex I is presented. This chapter concludes by introducing the clinical spectrum of mitochondrial pathology arising from defects in both
the mitochondrial and nuclear genomes, with emphasis being placed on human complex I deficiency.

Part I. Mitochondrial structure, function and inheritance

The mitochondrion

All reactions in cells involving growth and metabolism require energy. Mitochondria were identified 51 years ago as the organelles responsible for most cellular energy-production (reviewed by Gray et al, 1999). It is generally believed that mitochondria represent the descendents of primitive bacterial cells (cyanobacteria) that became symbiotically associated with primitive ancestors of the present eukaryotic organisms, thereby increasing the host's energy-generating capacity (Gray et al, 1999). Each mitochondrion performs a multitude of functions which include the reactions of the Krebs cycle, oxidative phosphorylation and β-oxidation (Darnell et al, 1986). Mitochondria display an amazing plasticity of form and distribution. The size, shape and quantity of mitochondria vary between tissues and even among different locations within the same tissue (Munn, 1974). Although, their internal structural organization is highly conserved, the external shape of mitochondria is variable. In addition to the classic kidney-bean shaped organelles observed in electron micrographs (Fig. 1.1), mitochondria are also frequently found as extended reticular networks (Chen, 1988). These networks are extremely dynamic in growing cells, with tubular sections dividing in half, branching and fusing to create a fluid tubular web (Bereiter-Hahn and Voth, 1994). In differentiated cells, such as those found in cardiac muscle or kidney tubules, mitochondria are often
localized to specific cytoplasmic regions rather than randomly distributed (Yaffe, 1999).

Typical mitochondria within a rat liver cell will exhibit an elliptical shape with an approximate length of 1-3μm and a width of 0.1-1μm (Lehninger, 1964). The average number of mitochondria within a rat liver cell is approximately 1000 (Munn, 1974), but a range of 500-2500 has been reported. Therefore, mitochondria can occupy nearly 20% of the total cellular volume (Lehninger, 1964).

Figure 1.1. Electron micrograph of a mammalian mitochondrion. The outer membrane as well as the highly folded, finger-like cristae which make up the inner membrane are clearly visible. Adapted from Fawcett, A., 1994.

**Mitochondrial ultrastructure**

The presence of a double membrane is a common feature of all mitochondria (Fig. 1.2). The outer and inner mitochondrial membranes define the two submitochondrial spaces:
the intermembrane space between the two membranes, and the central matrix compartment (Darnell et al, 1986). The outer mitochondrial membrane is freely permeable to most of the small molecules (<10 kDa) because it is covered with hydrophilic pores or channels composed of the protein porin (Manella, 1982). The inner mitochondrial membrane, because of its high protein and cardiolipin content is only freely permeable to O₂, CO₂, H₂O and small metabolites (Darnell et al, 1986). It is highly folded, with finger-like projections termed cristae which greatly increase its internal surface area (Darnell et al, 1986). While the respiratory chain is situated in the inner mitochondrial membrane, most of the reactions involving the oxidation of pyruvate and fatty acids take place in the mitochondrial matrix (Darnell et al, 1986).

**Mitochondrial DNA organization**

The majority of mitochondrial proteins are nuclear encoded and imported into the mitochondria from the cytoplasm. In addition, mitochondria possess their own unique genome and this mitochondrial deoxyribonucleic acid (mtDNA) is inherited maternally in humans, because sperm mitochondria do not survive after fertilization (Giles et al, 1980). The human mitochondrion typically contains 2 to 10 copies of the mitochondrial genome (Harding and Holt, 1993). MtDNA is a closed circular, double stranded (ds) molecule consisting of 16,569 bp and has been completely sequenced (Anderson et al, 1981; reviewed by Wallace, 1993) (Fig. 1.3). It is a compact piece of genetic information with little intervening, non-coding sequence with the exception of its short regulatory region termed the D- or displacement loop (Anderson et al, 1981; Tzagoloff and Myers, 1986).
Figure 1.2. Mitochondrial structure and membrane organization. A common feature of all mitochondria is the presence of a double membrane. The outer membrane completely envelopes the inner membrane. The highly invaginated inner membrane houses the OXPHOS system. ● - Complex I; ○ - Complex II; ● - Complex III; ○ - Complex IV; ❀ - Complex V
MtDNA encodes genes for the 12S and 16S ribosomal ribonucleic acids (rRNAs), 22 transfer ribonucleic acids (tRNAs) and 13 messenger ribonucleic acids (mRNAs) for polypeptides which are components of the mitochondrial oxidative phosphorylation system (Wallace, 1993).

**Mitochondrial replication, transcription and translation**

Due to their different buoyant densities in alkaline cesium chloride gradients, the two strands of mitochondrial DNA are referred to as the heavy (H) or guanine rich strand, and the complementary light (L) strand which is cytosine rich (Larsson and Clayton, 1995). DNA replication initiates within the D-loop region at the $O_H$ origin of replication on the H-strand (Larsson and Clayton, 1995). When the leading strand has elongated to two-thirds of its total length, the $O_L$ origin of replication on the L-strand (which is nested in a cluster of five tRNA genes) is then exposed and initiates lagging-strand replication (Larsson and Clayton, 1995). Components that are crucial for the replication process such as the mitochondrial specific $\gamma$-DNA polymerase (Bolden et al, 1977), mtDNA helicase (Hehman and Hauswirth, 1992) and primase enzymes (Wong and Clayton, 1985) are nuclear encoded factors which are imported into the mitochondria. The H-strand encodes the 12S and 16S rRNAs, 14 tRNAs and 12 polypeptides of the respiratory chain, while the L-strand encodes the ND6 subunit and eight tRNAs (Shoffner and Wallace, 1994). Mitochondrial transcripts begin at two promoter regions, $P_H$ and $P_L$ for the H- and L-strand transcripts, respectively (Shoffner and Wallace, 1994)). With the help of mitochondrial transcription factor (h-mtTFA) and possibly some other factor, a
Figure 1.3. Organization of the mitochondrial DNA genome. The circular, double stranded 16.5 kb human mitochondrial DNA is illustrated, indicating the locations of the encoded genes. MtDNA encodes two rRNAs (12S, 16S), 22 tRNAs (\[\ldots\]\] for its own protein synthesis, and 13 mRNAs for protein subunits of the respiratory chain complexes: ND1-6 for complex I, cyt b for complex III, CO I-III for complex IV and ATP 6 and 8 for complex V. (Adapted from Pitkanen, S., Academic dissertation, 1997).
mitochondrial specific RNA polymerase produces long polycistronic transcripts which are later processed (Larsson and Clayton, 1995). An additional transcript is also produced from the PH promoter at a rate approximately 10-30 fold greater than the full length PH transcript (Shoffner and Wallace, 1994). Following their release from the primary polycistronic transcript by a mitochondrial endonuclease (P), both tRNA molecules and coding transcripts are further modified to form mature functional tRNAs and polyadenylated mRNAs, respectively (Tzagoloff and Myers, 1986). The 13 mRNAs are then translated in the mitochondria by mitochondrial ribosomes which utilize the two mitochondrial rRNAs and 22 mitochondrial tRNAs (Shoffner and Wallace, 1994). The mitochondrial ribosomal complex is composed of small 28S and large 39S subunits (Tzagoloff and Myers, 1986). Mitochondrial ribosomes have been proposed to bind in a non-specific manner to mitochondrial transcripts, whereby the small 28S ribosome subunit interacts with the mRNA molecule and scans for the initiation codon (Liao and Spremulli, 1990). Because of this unique interaction, the translation of mitochondrial transcripts is believed to involve monoribosomes rather than polyribosomes as is seen in cytoplasmic protein synthesis (Liao and Spremulli, 1990). This process may be aided by mitochondrial translational initiation factors and subunit S5, a GTP binding component of the small ribosomal complex (O'Brien et al, 1990).

**Mitochondrial protein import**

Proteins destined for mitochondrial import have either an amino terminal targeting sequence (most common) or an internal targeting sequence (Rassow et al, 1999) (Fig. 1.4).
The matrix targeting signals (MTSs) at the amino terminal are usually 20 to 60 amino acid residues in length, rich in positively charged amino acids (arginine and lysine), rich in hydroxylated amino acids (serine and threonine), lack acidic residues (aspartic acid and glutamic acid), and have an amino acid sequence necessary for forming an amphipathic helical structure (von Heijne, 1986). The internal targeting sequences however are poorly characterized and are mainly found in hydrophobic preproteins (Herrmann and Neupert, 2000).

The proteins which participate in the translocation of proteins across the mitochondrial outer membrane are systematically named TOM proteins; the corresponding proteins of the inner membrane are named TIM proteins (Hurt et al, 1984). The TOM complex is comprised of an array of import receptors and a protein conducting channel (Komiya et al, 1988). Targeting sequences are recognized by the receptors Tom20 and Tom70 and are transferred to a general insertion pore made up of Tom40, Tom22, Tom7 and Tom6 (Komiya et al, 1998). After passage across the outer membrane, some preproteins first bind to proteins in the intermembrane space, others immediately insert into import sites of the inner membrane (Komiya et al, 1998). Preproteins that translocate into the intermembrane space, interact with the TIM23 complex made up of proteins Tim17, Tim23 (which form the protein-conducting channel), Tim44 (a hydrophilic matrix protein) and presumably an as yet unidentified 14 kDa subunit (Herrman and Neupert, 2000). Translocation across the mitochondrial inner membrane is strictly dependant on the mitochondrial membrane potential (Berthold et al, 1995). It is believed that the membrane potential exerts an electrophoretic effect on the
positively charged parts of the preproteins, irrespective of whether the charges are localized at the amino terminus or in mature parts of the preprotein (Berthold et al, 1995). In the matrix, the incoming preprotein is bound by the chaperone mtHSP70 which is associated with Tim44 (Berthold et al, 1995). This process is ATP dependent and is assisted by the co-chaperone Mge1, a mitochondrial homolog of the prokaryotic GrpE protein which facilitates the release of Tim44 from mt-Hsp70 and tight binding to the preprotein (Westermann et al, 1995). As soon as an amino-terminal presequence enters the matrix compartment, it is usually cleaved off by the specific processing enzyme MPP (mitochondrial processing peptidase) (Hurt et al, 1984).

Three different pathways are responsible for translocation into the inner membranes depending on the targeting sequence of the preprotein (reviewed by Herrmann and Neupert, 2000). Proteins can be imported as described above, but become arrested at the level of the TIM23 complex and laterally inserted into the lipid bilayer (Kaput et al, 1982). Secondly, proteins can be completely transported into the matrix from where they reinsert into the inner membrane (Hartl et al, 1986). To date, the only identified component of this insertion complex is Oxal (Hell et al, 1998). A third pathway into the inner membrane is used by the hydrophobic preproteins that do not contain MTSs but instead have internal targeting signals (Palmisano et al, 1998). They bind to soluble proteins Tim9 and Tim10 and are then imported via Tim12 and a larger complex of membrane proteins containing Tim22 and Tim54 (Palmisano et al, 1998). Recently, a second soluble intermembrane space complex formed by Tim8 and Tim13 was described
which also seems to be involved in mitochondrial protein import (Leuenberger et al., 1999).

Figure 1.4. The protein import machinery of mitochondria. Transport of preproteins across the outer membrane is mediated by Tom proteins while transport across the inner membrane is mediated by Tim proteins. Depending on the targeting sequence, protein import into and across the inner membrane can follow four different pathways and requires the membrane potential $\Delta \psi$. OM, outer membrane; IM, inner membrane; MPP, mitochondrial processing peptidase; Mge1, mitochondrial GrpE homologue; mtHSP70, mitochondrial heat shock protein of 70kDa; Tom, translocase of the outer membrane; Tim, translocase of the inner membrane.
**Energy metabolism**

Cellular energy requirements are satisfied mainly by the hydrolysis of phosphate bonds in ATP (reviewed by Zubay, 1993). Therefore, in order to sustain all cellular processes, ATP pools must be maintained at high enough levels. Mitochondria are the organelles responsible for transforming energy from the oxidative breakdown of carbohydrate, fatty acids and amino acids into ATP (Zubay, 1993). Glucose metabolism begins in the cytosol with glycolysis, the conversion of one glucose molecule into two molecules of pyruvate (Zubay, 1993). This results in the reduction of two NAD$^+$ (nicotinamide adenine dinucleotide oxidized form) to two NADH (nicotinamide adenine dinucleotide reduced form), as well as a net energy yield of two ATP (Zubay, 1993). In mammalian systems, the mitochondrial malate-aspartate shuttle transfers the two molecules of NADH formed in glycolysis into mitochondria (reviewed by Meijer and van Dam, 1974). A less efficient glycerophosphate shuttle serves the same function in tissue such as insect muscle (Voet and Voet, 1990). Oxidative decarboxylation of pyruvate takes place in the mitochondrial matrix, followed by oxidation by the Krebs cycle enzymes (Fig. 1.5) (Voet and Voet, 1990). The energy released by the oxidation of one molecule of glucose is conserved in ten NADH, two FADH$_2$ (flavin adenine dinucleotide reduced form) and two ATP molecules (Voet and Voet, 1990). NADH and FADH$_2$ are further oxidized by the mitochondrial respiratory chain (Voet and Voet, 1990). In the absence of oxygen, the lactate dehydrogenase enzyme catalyzes a reaction to produce lactate, the end-product of anaerobic glycolysis, and NAD$^+$, which can then be recycled for the oxidative reaction (Zubay, 1993). Therefore, reasons for lactic acid
accumulation in the cell include lack of sufficient oxygen or dysfunction of mitochondrial enzymes required for oxidative phosphorylation (reviewed by Robinson, 1989).

**Figure 1.5. The glucose oxidation pathway in mammals.** Complete oxidation of one molecule of glucose gives a net of 38 ATP molecules. Dashed lines indicate necessary reactions for continuation of glycolysis. LDH, lactate dehydrogenase; PDH-complex, pyruvate dehydrogenase complex. — Complex I; - Complex II; - Complex III; - Complex IV; - Complex V; - Quinone; - Cytochrome c.
Part II. The Mitochondrial respiratory chain complexes

**Organization of the OXPHOS system**

The oxidative phosphorylation system (OXPHOS) located in the inner mitochondrial membrane consists of five distinct multimeric protein assemblies (Fig. 1.6):

1) NADH-ubiquinone oxidoreductase (Complex I)
   also called NADH-dehydrogenase or NADH-coenzyme Q reductase

2) Succinate-ubiquinone oxidoreductase (Complex II)

3) Ubiquinol-ferricytochrome c oxidoreductase (Complex III)

4) Cytochrome c oxidase (Complex IV)

5) ATP synthase or $F_1F_0$-type ATPase (Complex V)

A new study by Schagger and Pfeiffer (2000) shows that these complexes are not randomly distributed, but instead assemble into supramolecular structures. The work involved solubilization of the membrane protein complexes through mild one-step protocols using dodecylmaltoside (DDM), digitonin and Triton X-100, followed by blue-native PAGE to isolate the supramolecular structures. In mammalian mitochondria, almost all complex I is seen to assemble into supercomplexes comprising complexes I and III and up to four copies of complex IV (Schagger and Pfeiffer, 2000). In *Saccharomyces cerevisiae*, complex IV is predominantly found associated with complex III and exists in three forms, the free dimer, and two supercomplexes with another one or two complex IV monomers (Schagger and Pfeiffer, 2000). The amount of supercomplexes formed depends on the cell's demand for energy. A respirasome model is proposed with two copies of a $I_1III_2IV_4$ building block and one copy of a $III_2IV_4$ building block (Schagger and Pfeiffer,
2000) to fit the overall 1:3:6 stoichiometries of complexes I:III:IV determined by Hatefi (1985). Association of complex II with any of the other respiratory chain complexes was not identified (Schagger and Pfeiffer, 2000).

**Figure 1.6. Mitochondrial respiratory chain and ATP-synthase.** Arrows indicate the direction of electron and proton ($H^+$ = hydrogen ion) flow. C I — V, complexes I — V; Q, coenzyme Q; cyt, cytochrome; FAD, flavin adenine dinucleotide; Fe-S, iron-sulphur clusters.

**The OXPHOS system: Role in electron transport and proton translocation**

The first four complexes, with ubiquinone and cytochrome c, make up the mitochondrial respiratory chain, also called the electron transport chain (Saraste, 1999).
The operation of the respiratory chain is characterized by two distinct processes that are linked: electron transport and proton pumping (Saraste, 1999). NADH is oxidized by complex I and succinate by complex II, followed by the electron transfer to ubiquinone and then further to complex III, to cytochrome c, to complex IV and to the final electron acceptor, oxygen (Voet and Voet, 1990). The electron transport via complexes is coupled to proton pumping from the matrix to the intermembrane space, producing an electrochemical gradient (Voet and Voet, 1990). Because of this gradient, the protons flow back from the intermembrane space to the mitochondrial matrix through complex V, ATP synthase, and the released energy is captured in the form of ATP (Voet and Voet, 1990).

The mitochondrial electron transport chain houses at least 4 types of electron carriers: flavins, iron sulfur clusters, quinone and cytochromes (Ohnishi, 1998). Complexes I and II contain as prosthetic groups, flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD), respectively (Ohnishi, 1998). Protein subunits in complexes I, II and III bind iron sulfur clusters with two to four iron and sulphur atoms: [2Fe-2S], [3Fe-3S] or [4Fe-4S] (Saraste, 1999). The ubiquinone present in the mitochondrial inner membrane is responsible for passing electrons on to the respiratory chain cytochrome system, consisting of three types of cytochromes: $a$ ($a$ and $a_3$), $b$, and $c$ ($c$ and $c_1$) (Saraste, 1999). The iron atoms in the iron sulphur clusters and in hemes of the cytochromes undergo oxidation and reduction during respiration, cycling between the ferrous ($Fe^{2+}$) and ferric ($Fe^{3+}$) oxidation states (Saraste, 1999).
**Complex I: The NADH-ubiquinone oxidoreductase complex**

NADH-ubiquinone oxidoreductase or complex I is the largest of the membrane-bound enzymes of the mitochondrial respiratory chain with a total molecular mass of ~10^3 kilodaltons (kDa) for monomeric complex I (Fig. 1.7) (Smeitink et al, 1998a). The mammalian complex I enzyme is composed of at least 43 subunits of which 7 are mitochondrially encoded while the remaining subunits are encoded by the nuclear genome (Smeitink et al, 1998a). In comparison, the fungus *Neurospora crassa* which is widely used as a simple eukaryotic model organism in the study of complex I, has 35 subunits (Schulte et al, 1994), *Paracoccus denitrificans* (designated NDH-1) has 14 subunits (Takano et al, 1996) and *Escherichia coli* complex I has 13 subunits (Blattner et al, 1997). However, *Saccharomyces cerevisiae* and other fermentative yeasts do not contain this multi-enzyme complex of the respiratory chain but instead have a simple dimeric diaphorase (discussed below) (Brody et al, 1997).

**Evolution of Complex I**

Complex I is thought to have originated by fusion of pre-existing protein assemblies constituting modules for coupled electron transfer and proton transport. These functional modules are defined by the homology of parts of complex I to other bacterial enzymes. Complex I is found in purple bacteria and in the mitochondria of most eukaryotes. While the eukaryotic NADH:ubiquinone oxidoreductase is referred to as complex I, its bacterial counterpart has traditionally been called NADH dehydrogenase type I (Friedrich and Weiss, 1997). The known examples of bacterial complex I from purple bacteria consist of
14 different subunits while the mitochondrial complex contains at least 28 accessory proteins which do not directly participate in electron and proton transport (Friedrich and Weiss, 1997). There is no evidence that bacteria other than purple bacteria contain this respiratory enzyme. In fact, a non proton pumping NADH:ubiquinone oxidoreductase with a single FAD redox group called NADH dehydrogenase type II appears to be more widespread than complex I in bacteria (Matsushita et al, 1987; Yagi et al, 1992). Fungi and plant mitochondria contain two of these non proton-pumping NADH:ubiquinone oxidoreductases in addition to complex I (Friedrich and Weiss, 1997). These complexes have a lower affinity for NADH as compared to complex I and most likely operate as overflow outlets for an excess of reducing equivalents (Friedrich and Weiss, 1997). Fermentative yeasts which lack complex I, use these complexes exclusively to oxidize mitochondrial NADH (Friedrich and Weiss, 1997). A minimal bacterial complex which is homologous to the respiratory complex I, is found in cyanobacteria and chloroplasts carrying only 11 subunits (Berger et al, 1993). However, this system is thought to work as a NADPH:plastoquinone oxidoreductase in a cyclic photosynthetic electron transfer (Friedrich et al, 1995).

Most information about bacterial complex I comes from E. coli where complex I genes are organized in the so-called nuo-operon (Friedrich et al, 1995). Seven genes code for peripheral proteins, including all proteins with binding motifs for NADH, FMN and all Fe-S clusters (Friedrich et al, 1995). The seven remaining genes code for the hydrophobic, intrinsic membrane proteins. These 7 intrinsic membrane subunits, the homologues of the E. coli NuoA, H and J-N are mitochondrially encoded in animals and
fungi (Attardi and Schatz, 1988), while all other subunits are nuclear-encoded in most eukaryotes (Friedrich et al, 1995) (Table 1.1). The evolution model proposed by Finel (1998) suggests that the hydrophobic subunits of complex I evolved together with the nuclear-encoded subunits until they were transferred from the mitochondrial chromosome to the nucleus.

**Table 1.1. Nomenclature and properties of homologous complex I subunit genes of *E. coli, B. taurus* and *H. sapiens***

<table>
<thead>
<tr>
<th>E. coli</th>
<th>B. taurus</th>
<th>H. sapiens</th>
<th>Predicted Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>NuoA</td>
<td>ND3</td>
<td>ND3</td>
<td></td>
</tr>
<tr>
<td>NuoB</td>
<td>PSST</td>
<td>NDUF57</td>
<td>Q binding/e⁻ transport</td>
</tr>
<tr>
<td>NuoC</td>
<td>30 (IP)</td>
<td>NDUF3</td>
<td>NADH binding</td>
</tr>
<tr>
<td>NuoD</td>
<td>49 (IP)</td>
<td>NDUF2</td>
<td>Q binding</td>
</tr>
<tr>
<td>NuoE</td>
<td>24 (FP)</td>
<td>NDUF2</td>
<td>e⁻ transport</td>
</tr>
<tr>
<td>NuoF</td>
<td>51 (FP)</td>
<td>NDUFV1</td>
<td>NADH binding/e⁻ transport</td>
</tr>
<tr>
<td>NuoG</td>
<td>75 (IP)</td>
<td>NDUF1</td>
<td>e⁻ transport</td>
</tr>
<tr>
<td>NuoH</td>
<td>ND1</td>
<td>ND1</td>
<td>H⁺ pumping</td>
</tr>
<tr>
<td>NuoI</td>
<td>TYKY</td>
<td>NDUF8</td>
<td>e⁻ transport</td>
</tr>
<tr>
<td>NuoJ</td>
<td>ND6</td>
<td>ND6</td>
<td>H⁺ pumping</td>
</tr>
<tr>
<td>NuoK</td>
<td>ND4L</td>
<td>ND4L</td>
<td></td>
</tr>
<tr>
<td>NuoL</td>
<td>ND5</td>
<td>ND5</td>
<td></td>
</tr>
<tr>
<td>NuoM</td>
<td>ND4</td>
<td>ND4</td>
<td>H⁺ pumping</td>
</tr>
<tr>
<td>NuoN</td>
<td>ND2</td>
<td>ND2</td>
<td>H⁺ pumping</td>
</tr>
</tbody>
</table>

The electron transfer moiety of complex I can be traced back to two different origins. The first is to the diaphorase part of soluble NAD⁺-reducing hydrogenase found in purple bacteria such as *Alcaligenes eutrophus, Desulfovibrio fructosovorans* and cyanobacteria *Anabaena variabilis, Anacystis nidulans*, while the second is to the formate hydrogenlyase complex of *E.coli* (Friedrich and Weiss, 1997). Based on homology, the origin of the proton transporting moiety of complex I can be related to bacterial Na⁺/H⁺ and K⁺/H⁺...
antiporters (Friedrich and Weiss, 1997). It has also been proposed that NuoL and NuoH and the sugar permeases of the bacterial phosphoenolpyruvate-dependent phosphotransferase system belong to a superfamily of pore-forming proteins (Reizer et al, 1991). As the electron transferring and proton transporting modules were joined together, they gave rise to the ancestor of complex I and the formate hydrogenlyase (Friedrich and Weiss, 1997).

To promote electron transport in a specific direction and to prevent the capture of electrons by other redox acceptors in the aqueous and membrane phases, there must be a layer of insulating protein surrounding the electron pathway that should be about 17 to 20 in thickness (Moser and Dutton, 1996). Many of the additional subunits of the mitochondrial enzyme complex are thought to form this scaffold, keeping the redox groups in the right position to prevent the electrons from escaping and forming reactive oxygen species (Friedrich and Weiss, 1997). This results in safer energy conversion in eukaryotes compared to bacteria. Because several of these additional subunits show no sequence similarity between animal and fungi, they are thought to have emerged late in evolution when these species had already diverged or they diverged so fast that a possible homology cannot be seen (Friedrich and Weiss, 1997).

**Subunit composition of the Human NADH:ubiquinone oxidoreductase complex**

As previously elaborated, human mitochondrial complex I appears to be made up of 43 subunits of which 7 are mitochondrially encoded. The nuclear encoded subunits are synthesized in the cytosol and transported into the mitochondria (Chomyn et al, 1988)
Treatment with chaotropic salts like sodium bromide resolves complex I into three fractions. They are, namely 1) the flavin protein (FP) fraction containing polypeptides with a high flavin, iron and sulphide content, 2) the iron protein fraction (IP) which contains a high iron and low flavin content and 3) the hydrophobic (HP) fraction containing polypeptides with the lowest non-heme iron protein ratio (Galante et al., 1979). The extrinsic membrane domain has at least 20 subunits, most of which are hydrophilic, including all the subunits of the FP and IP fractions (Galante et al., 1979). It also contains the FMN and most of the Fe-S clusters. The intrinsic membrane domain which contains the HP fraction is an assembly of ~24 nuclear-encoded subunits and the 7 mitochondrial gene products (Belogrudov et al., 1994). However, this fraction also contains globular water-soluble subunits and being a part of the HP fraction does not necessarily indicate that a subunit is hydrophobic or that it belongs to the membrane domain (Walker et al., 1992). Fractionation using detergents (see later section) gives a clearer picture of complex I structure.

(i) The Flavoprotein Fraction (FP)

The genes encoding the 3 subunits which make up the Flavoprotein (FP) fraction, namely, *NDUFV1* (51 kDa), *NDUFV2* (24 kDa) and *NDUFV3* (10 kDa), have all been characterized at both complementary DNA (cDNA) and nuclear DNA (nDNA) levels (Table 1.2).

The 51 kDa subunit (encoded by *NDUFV1*) is known to hold the binding site for NADH (Patel et al., 1991) and also contains the FMN (Krishnamoorthy and Hinkle,
Table 1.2. Current molecular genetic knowledge of human nuclear-encoded subunits of complex I of the mitochondrial electron transport chain

<table>
<thead>
<tr>
<th>Gene</th>
<th>Subunit</th>
<th>Mr</th>
<th>Chromosome</th>
<th>cDNA length (ORF)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flavoprotein fraction (FP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDUFV1</td>
<td>51 kDa</td>
<td>51</td>
<td>11q13</td>
<td>1331 bp</td>
<td>Schuelke et al, 1998:</td>
</tr>
<tr>
<td>NDUFV2</td>
<td>24 kDa</td>
<td>24</td>
<td>18p11.21-11.22</td>
<td>650 bp</td>
<td>de Coo et al, 1995</td>
</tr>
<tr>
<td>NDUFV3</td>
<td>10 kDa</td>
<td>10</td>
<td>21q22.3</td>
<td>221 bp</td>
<td>de Coo et al, 1997</td>
</tr>
<tr>
<td>Iron-Sulphur protein fraction (IP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDUF51</td>
<td>75 kDa</td>
<td>75</td>
<td>2q33-34</td>
<td>2111 bp</td>
<td>Chow et al, 1991</td>
</tr>
<tr>
<td>NDUF52</td>
<td>49 kDa</td>
<td>49</td>
<td>1q23</td>
<td>1388 bp</td>
<td>Loeffen et al, 1998a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30 kDa</td>
<td>30</td>
<td>1p11.11</td>
<td>791 bp</td>
<td>Loeffen et al, 1998a</td>
</tr>
<tr>
<td></td>
<td>AODQ</td>
<td>18</td>
<td>5q11.1</td>
<td>398 bp</td>
<td>van den Heuvel et al, 1998</td>
</tr>
<tr>
<td></td>
<td>NDUS55</td>
<td>15 kDa</td>
<td>1p33-34.2</td>
<td>317 bp</td>
<td>Loeffen et al, 1999</td>
</tr>
<tr>
<td></td>
<td>DDDG</td>
<td>13</td>
<td>5p15.33-pter</td>
<td>371 bp</td>
<td>Loeffen et al, 1998a</td>
</tr>
<tr>
<td></td>
<td>NDUFA5</td>
<td>B13</td>
<td>7q31.33</td>
<td>347 bp</td>
<td>Pata et al, 1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>39 kDa</td>
<td>39</td>
<td>12p13.3</td>
<td>1130 bp</td>
<td>Baens et al, 1994</td>
</tr>
<tr>
<td></td>
<td>NDUFA10</td>
<td>42 kDa</td>
<td>12p</td>
<td>959 bp</td>
<td>Loeffen et al, 1998b</td>
</tr>
<tr>
<td></td>
<td>SDAP</td>
<td>10</td>
<td>16p12.1-12.3</td>
<td>263 bp</td>
<td>TriepeIs et al, 1999a</td>
</tr>
<tr>
<td></td>
<td>NDUFB1</td>
<td>MNLL</td>
<td>5.5</td>
<td>173 bp</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NDUFB2</td>
<td>AGGG</td>
<td>7q34-35</td>
<td>215 bp</td>
<td>Loeffen et al, 1998b</td>
</tr>
<tr>
<td></td>
<td>NDUFB3</td>
<td>B12</td>
<td>7q34-35</td>
<td>215 bp</td>
<td>Loeffen et al, 1998b</td>
</tr>
<tr>
<td></td>
<td>NDUFB4</td>
<td>B15</td>
<td>7q34-35</td>
<td>215 bp</td>
<td>Loeffen et al, 1998b</td>
</tr>
<tr>
<td></td>
<td>NDUFB5</td>
<td>SGDH</td>
<td>18</td>
<td>428 bp</td>
<td>Loeffen et al, 1998b</td>
</tr>
<tr>
<td></td>
<td>NDUFB6</td>
<td>B17</td>
<td>9p13.2</td>
<td>383 bp</td>
<td>Smeitink et al, 1998b</td>
</tr>
<tr>
<td></td>
<td>NDUFB7</td>
<td>B18</td>
<td>9p13.2</td>
<td>383 bp</td>
<td>Smeitink et al, 1998b</td>
</tr>
<tr>
<td></td>
<td>NDUFB8</td>
<td>ASH1</td>
<td>10q23.2-23.3</td>
<td>473 bp</td>
<td>Loeffen et al, 1998b</td>
</tr>
<tr>
<td></td>
<td>NDUFB9</td>
<td>B23</td>
<td>8p24.21</td>
<td>466 bp</td>
<td>Gu et al, 1996</td>
</tr>
<tr>
<td></td>
<td>NDUFB10</td>
<td>PDSW</td>
<td>16p13.3</td>
<td>411 bp</td>
<td>Loeffen et al, 1998b</td>
</tr>
<tr>
<td></td>
<td>NDUC1</td>
<td>KFY1</td>
<td>4q28.2-28.3</td>
<td>146 bp</td>
<td>Ton et al, 1997</td>
</tr>
<tr>
<td></td>
<td>NDUC2</td>
<td>MMG1</td>
<td>14.5</td>
<td>356 bp</td>
<td>Loeffen et al, 1998b</td>
</tr>
<tr>
<td></td>
<td>NDUS7</td>
<td>PSST</td>
<td>10p13.2</td>
<td>524 bp</td>
<td>Hyslop et al, 1996</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>17.2 kDa</td>
<td>435 bp</td>
<td>TriepeIs et al, 2000</td>
</tr>
</tbody>
</table>


* Kim et al, 1997 were another group that published the cDNA sequence of the human MLRQ homolog. Chromosomal localization of the NDUF44 gene was not carried out by these authors.
1988), a tetranuclear iron-sulfur cluster (Ohnishi et al, 1985) and a consensus motif for the binding of the nuclear respiratory factor II (NRF-2) in its genomic structure, which is thought to be involved in the transcriptional regulation of nuclear genes which code for mitochondrial proteins (Schuelke et al, 1998). 100% anti-sense homology was found between the 3' UTR of NDUFVI-mRNA and the 5' UTR of the mRNA for the γ-interferon inducible protein (IP-30) precursor (Schuelke et al, 1998). It is hypothesized that the NDUFVI-mRNA may act as an anti-sense suppressor, restraining translation of IP-30 in tissues with high energy demand. This could therefore be the molecular link between complex I deficiency and inflammatory myopathy which have been repeatedly described to occur together (Schuelke et al, 1998). The 24 kDa subunit contains four strictly conserved cysteine residues for the binding of a binuclear iron-sulfur cluster (Ohnishi et al, 1985). This subunit has also been shown to bind GTP and possibly exhibit GTPase activity when bound to the native complex I (Hegde, 1998). Mutational studies on the 24 kDa subunit in Neurospora crassa has showed that this subunit is absolutely essential for complex I activity and this may explain cases where the 24 kDa subunit is reduced or absent in human mitochondrial diseases (Almeida et al, 1999). A good example of this is shown by Schapira et al (1988) where the 24 kDa subunit appears to be absent in a patient with mitochondrial myopathy. Hattori et al (1998) have reported a Ala29Val substitution in the mitochondrial targeting sequence of the 24 kDa subunit in patients with Parkinson's disease (PD). This frequency of homozygotes for the mutation was significantly higher in PD patients than in control subjects and the
mutation may well be a cause of complex I deficiency in Parkinson's disease (Hattori et al., 1998).

The 10 kDa subunit has no redox centers (de Coo et al., 1997) and is situated at close proximity to the 24- and 51-kDa subunits (Yamaguchi and Hatefi, 1993). The localization of the \textit{NDUFV3} gene on chromosome 21q22.3 borders the location of the gene for the mitochondrial ATP50 protein, which is thought to contribute to the Down Syndrome phenotype (Chen and Antonarakis, 1995). Because Down syndrome has been postulated to be a contiguous gene syndrome, de Coo et al. (1997) believe that the 10 kDa subunit might also be associated with this disease.

(ii) The Iron-Sulphur Protein Fraction (IP)

The Iron-sulfur (IP) fraction is made up of at least 7 subunits, encoded by \textit{NDUFS1} (75 kDa), \textit{NDUFS2} (49 kDa), \textit{NDUFS3} (30 kDa), \textit{NDUFS4} (AQDQ/18 kDa), \textit{NDUFS5} (15 kDa), \textit{NDUFS6} (DDGD/13 kDa) and \textit{NDUFA5} (B13). All of these subunits have been characterized at the cDNA level (Table 1.2). The genomic DNA sequence has only been determined for the B13 subunit encoded by the \textit{NDUFA5} gene (Tensing et al., 1999) and the 30 kDa subunit encoded by the \textit{NDUFS3} gene (Procaccio et al., 2000).

The 75 kDa subunit contains conserved cysteine motifs allowing for the existence of one tetra-nuclear, one binuclear and possibly another tetranuclear iron-sulfur cluster (Ohnishi, 1998). The 49 kDa subunit seems to be essential for complex I function as was demonstrated by knockout mutants of the 49 kDa gene in \textit{N.crassa}, which lacked NADH dehydrogenase activity completely because the matrix arm of the complex failed to
assemble (Schulte and Weiss, 1995). Sequencing of the gas-1 gene in *Caenorhabditis elegans* has revealed that it is a homologue of the 49 kDa subunit in the worm's respiratory chain and that it plays a role in the determination of anesthetic sensitivity in *C. elegans* (Kayser et al, 1999). Darrouzet et al (1998) using bacterial genetics have shown that the 49 kDa subunit is involved in the binding of piericidin and rotenone (both quinone-related inhibitors) and thereby implicate this subunit in quinone binding. Both the 49- and 30 kDa subunits contain highly conserved phosphorylation sites which are thought to be involved in regulatory functions (Loeffen et al, 1998a). The 30 kDa subunit shows similarity to a protein encoded by a gene on the chloroplast genome (Weiss et al, 1991). An active cAMP-dependent protein kinase consensus phosphorylation site has been determined in the AQDQ subunit (Papa et al, 1996) and phosphorylation of this subunit in response to cholera-toxin treatment has been shown to be accompanied by a 2-3 fold enhancement of rotenone-sensitive NADH oxidase respiration and NADH:ubiquinone oxidoreductase activity of complex I (Scacco et al, 2000).

(iii) The Hydrophobic Protein Fraction (HP)

In humans, the largest fraction of complex I, the Hydrophobic fraction (HP), contains the 7 mitochondrially encoded subunits as well as the 24 nuclear encoded subunits which have been characterized on the cDNA level (Table 1.2).

a) The mitochondrially encoded subunits

The 7 subunits encoded by the mitochondrial genome are ND1 (318 aa), ND2 (347 aa), ND3 (115 aa), ND4 (459 aa), ND5 (603 aa), ND6 (174 aa) and ND4L (98 aa) (Chomyn et
Apolar stretches long enough to traverse the membrane are characteristic of these subunits in a variety of organisms, although homologous subunits vary considerably in size (Weiss et al., 1991). The great degree of sequence identity between the ND2, ND4 and ND5 encoded subunits suggests that these genes evolved from a single ancestral gene (Kikuno and Miyata, 1985). It has been proposed that these mitochondrially encoded subunits are involved in proton translocation (Guenebaut et al., 1998). The carboxyl group modifying reagent $N,N'$-dicyclohexyl-carbodiimide (DCCD) was found to act on complex I at the ND1 subunit site, thereby blocking proton translocation and to the same extent electron transfer in complex I (Yagi and Hatefi, 1988). The ND1 subunit has historically been associated with the ubiquinone reduction site from photolabeling studies done with rotenone analogues in isolated complex I (Earley et al., 1987). However, recent photoaffinity labeling studies by Schuler et al. (1999) on mitochondrial electron transport particles has suggested that ND1 is not directly involved in the action of quinone binding. In human cells, the absence of the ND4 subunit has led to a failure to assemble other mtDNA-encoded subunits and a complete loss of NADH:Q$_1$ oxidoreductase activity (Hofhaus and Attardi, 1993). The ND5 subunit has four cysteine residues that are conserved and thought to be able to ligate an iron-sulfur cluster located within the membrane (Weiss et al., 1991). The ND5 subunit contains the same consensus sequence around its conserved myristoylated lysine that the proton conducting COX subunit 1 does, although the exact purpose of the myristoylation is not yet known (Plesofsky et al., 2000). Work done on a mouse A9 cell line has shown that a frameshift mutation in the mitochondrial gene for the ND6 subunit, resulting in a
complete absence of the subunit caused a loss of assembly of the mtDNA-encoded subunits and a serious impairment of oxidative phosphorylation function (Bai et al., 1998). Similar loss of assembly was seen in the E35 stopper mutants of *N. crassa* which were deleted of the ND2 and ND3 subunits (Alves and Videira, 1998).

**b) The nuclear encoded subunits**

*NDUFA1* (MWFE), *NDUFA2* (B8), *NDUFA3* (B9), *NDUFA4* (MLRQ), *NDUFA6* (B14), *NDUFA7* (ASAT), *NDUFA8* (PGIV), *NDUFA9* (39 kDa), *NDUFA10* (42 kDa), *NDUFAB1* (SDAP), *NDUFBI* (MNLL), *NDUFB2* (AGGG), *NDUFB3* (B12), *NDUFB4* (B15), *NDUFB5* (SGDH), *NDUFB6* (B17), *NDUFB7* (B18), *NDUFB8* (ASHI), *NDUFB9* (B22), *NDUFB10* (PDSW), *NDUFC1* (KFY1), *NDUFC2* (MMTG), *NDUF S7* (PSST) and *NDUF S8* (TYKY) (*Table 1.2*) are nuclear genes that encode for subunits in the HP fraction of complex I. So far, the genomic organization has only been determined for *NDUFA1* (Zhuchenko et al., 1996), *NDUFA6* (Dunham et al., 1999), *NDUF B9* (Lin et al., 1999) and *NDUF S8* (de Sury et al., 1998).

Two consensus tetranuclear Fe-S clusters are present in the TYKY subunit (Dupuis et al., 1991a), and a consensus binuclear Iron-sulfur pattern is found in the PSST (Arizmendi et al., 1992) and PGIV (Dupuis et al., 1991b) subunits. The 10 kDa SDAP subunit is known to have a phosphopantetheine attachment site unique to acyl-carrying proteins (Runswick et al., 1991), as well as an EF-hand calcium binding domain (Triepels et al., 1999a).
Knowledge of the functionality of other subunits in the HP fraction is scarce. However, important information is emerging as more of these subunits are being characterized. For example, the expression of the B17 subunit was found to be highest in kidney, indicating that the composition of complex I may be tissue specific and that mutations in the \textit{NDUFB6} gene may cause distinctive phenotypes (Smeitink \textit{et al}, 1998b). This claim is also supported by the expression of subunits B8, B12 and B14, all of which were not found in the aorta, brain or kidney (Ton \textit{et al}, 1997). Au \textit{et al} (1999) demonstrated that the MWFE subunit is essential for mammalian complex I activity by showing that complex I activity was severely reduced (<10\%) in a MWFE mutant Chinese hamster cell line. The phosphorylation of MWFE as well as the 42 kDa subunit has been described (Sardanelli \textit{et al}, 1995). The B22 subunit is considered a candidate for BOR (branchio—oto-renal) syndrome because the \textit{NDUFB9} gene has been mapped to a 1-Mb deletion at chromosome 8q13 which also contains the gene for BOR syndrome (Gu \textit{et al}, 1996). Thus far however, mutational analysis of BOR families have yielded no association between the \textit{NDUFB9} gene and BOR syndrome (Lin \textit{et al}, 1999). The 39 kDa subunit related to hydroxysteroid reductase/isomerase was found to contain a NAD(P)-binding motif and is proposed to participate in intramitochondrial fatty acid synthesis, together with the acyl carrier protein (SDAP) of complex I (Friedrich \textit{et al}, 1995; Yarnaguchi \textit{et al}, 1998; Schulte \textit{et al}, 1999). The cDNA sequence of the 42\textsuperscript{nd} complex I subunit has recently been deduced from both bovine and human heart mitochondria and placed in the HP fraction (Skehel \textit{et al}, 1998; Triepels \textit{et al}, 2000). Electrospray mass spectrometry performed on complex I and two related subcomplexes in bovine heart,
identified a 17.2 kDa subunit (B17.2) with an acetylated α-amino group (Skehel et al, 1998). Since this protein shows 70% amino acid identity to a 13 kDa human protein associated with differentiation, it is hypothesized that this 13 kDa protein is part of the B17.2 kDa homologue in human complex I (Skehel et al, 1998).

Electrospray ionization mass spectroscopy experiments carried out on complex I have from time to time identified a 10566 Da protein (Fearnley et al, 1994). Sequencing of this protein would most likely result in the identification of the 43rd and final subunit of complex I (Skehel et al, 1998).

**Structural Model of Complex I**

Vital to the understanding of the operation of this large, multi-functional enzyme of the respiratory chain is the orientation and interaction of its subunits within the mitochondrial inner membrane. Most of the information pertaining to the composition, spatial and functional organization of complex I has been obtained from Bos taurus heart, bacteria and fungi (Grigorieff, 1999). Bovine heart complex I, like that of Neurospora crassa is an L-shaped enzyme with one arm (the extrinsic membrane domain) extending into the mitochondrial matrix and another arm (the intrinsic membrane domain), which remains in the membrane (Guenebaut et al, 1997; Grigorieff et al, 1998). The appearance of the bovine complex differs from that of N. crassa by possessing a significantly bigger membrane-bound globular domain and also by having a thin stalk region (30 diameter) linking this globular arm with the intrinsic membrane domain like bacterial complex I (Fig. 1.7) (Grigorieff et al, 1998). The stalk is thought to contain part of the electron transfer
Figure 1.7. Three dimensional models of complex I from *E. coli*, *N. crassa* and *B. taurus* as determined by electron cryo-microscopy. A reconstruction of complex I from (a) *E. coli* (b) *N. crassa* and (c) *B. taurus* is shown as determined by the electron cryo-microscopy studies of Guenebaut (1998) and Grigorieff (1998). All three models share the L-shaped structure with an intrinsic membrane arm extending into the lipid bilayer and a peripheral arm protruding into the matrix. Additional protein mass is observed around the mitochondrial complexes when compared to complex I from *E. coli*, especially at the junction between the arms and around the membrane domain. The *E. coli* and bovine structures both show a narrowing between their membrane matrix domains (the stalk).
pathway linking the NADH binding site in the globular arm with the ubiquinone binding site in the membrane domain (Grigorieff et al, 1998).

(i) Assembly of Complex I

The assembly of this enzyme in N. crassa occurs by the formation of a series of intermediates (Tuschen et al, 1990). The matrix arm and the membrane portion of complex I form independently and are joined in the course of assembly (Tuschen et al, 1990). The membrane arm is formed by the association of a 200 kDa and 350 kDa assembly intermediates (Kuffner et al, 1998). The larger 350 kDa assembly intermediate is associated with two extra proteins of 80 and 30 kDa which are not constituent parts of the mature complex I and are called complex I intermediate associated proteins. CIA84 and CIA30, respectively (Kuffner et al, 1998). These two proteins are considered single-target chaperones specific for complex I because they are integrated into the large membrane arm during an early stage of complex I assembly, but are set free when the complex has been assembled, only to be cycled once again to take part in the formation of further intermediates (Kuffner et al, 1998). Deletion of the cia genes coding for these proteins results in the severe disruption of the assembly process whereby the large membrane arm intermediate is not formed, even though the matrix arm and the small membrane arm intermediate are not affected (Kuffner et al, 1998). Homologues to the CIA proteins have not been detected in prokaryotic genomes and no such chaperones are yet known for other eukaryotic complex I enzymes.
(ii) Spatial organization and subunit interaction in human complex I

The use of strong but non-denaturing detergents such as LDAO (lauryldimethylamine oxide) has been used to resolve bovine mitochondrial complex I into two subcomplexes called 1α and 1β (Fig. 1.8) (Finel et al, 1992). Subcomplex 1α is composed of ~ 23 mostly hydrophilic subunits and contains all the redox centers. It is an active NADH:dehydrogenase which can transfer electrons to a water soluble ubiquinone and is likely to represent the peripheral arm and part of the membrane domain of the enzyme (Finel et al, 1992). Subcomplex 1β is made up of 17 mainly hydrophobic subunits corresponding to the intrinsic membrane domain of complex I and has no known enzyme activity (Finel et al, 1992).

Enzymatically active subcomplexes called 1λ, 1S and 1λS (Fig. 1.8) have also been purified by sucrose gradient centrifugation in the presence of detergents (Finel et al. 1994). All subcomplexes retained an NADH-oxidizing activity with ferricyanide or Q-1 as acceptors, but lost activity with decylubiquinone as acceptor along with a loss of rotenone sensitivity during Q-1 reductase activity (Finel et al, 1994). 1λ comprises 14 of the subunits in subcomplex 1α and represents most or all of the peripheral arm of complex I (Finel et al, 1994). The subcomplex 1λS contains only approximately 13 subunits and like 1λ, retains all the nuclear encoded subunits, whose homologues are present in the bacterial genome and which are known to bind FMN and Fe-S clusters (75-, 51-, 49-, 30-, 24- kDa, TYKY and PSST subunits) (Finel et al, 1994). The IS has more subunits including the very hydrophobic ND4, but there are also subunits present in
Figure 1.8. Model of the overlap in subunit composition between the different subcomplexes of complex I. Detergent treatment of complex I yields three major subcomplexes, the Iα fraction which contributes mainly to the peripheral arm, the Iβ fraction and the smaller Iγ fraction which mainly contribute to the membrane arm of the enzyme. The location of the smaller IS, Iα and IλS as well as the IβS and IβL subcomplexes are also shown. Subcomplexes that contain identical subunits are drawn above each other. IM refers to the inner mitochondrial membrane.

Subcomplex Iλ that are missing from IS (Finel et al, 1994). None of the mitochondrially encoded subunits have been found in subcomplexes Iα, Iλ or IλS (Finel et al, 1994). This suggests that most of the redox chemistry of complex I occurs in a hydrophilic domain on the inside (mitochondrial matrix) of the membrane, with special structural arrangements
required for proton translocation (Fine et al, 1994). The mitochondrionally encoded ND subunits may play a role in the formation of these structures (Fine et al, 1994).

Recently, a previously undetected fragment referred to as Iγ has been found by using LDAO to disrupt complex I (Sazanov et al, 2000). This subcomplex has been ascertained to contain the hydrophobic subunits ND1, ND2, ND3 and ND4L, all of which are mitochondrionally encoded and the nuclear encoded KFY1 subunit. Also, the Iγ subcomplex has been seen to release a fragment containing ND1 and ND2 and the IB subcomplex has been found to dissociate further to form IBL (containing ND4 and ND5) and IBS (containing the rest of the IB subunits) (Sazanov et al, 2000). Important associations illustrated by this study also include the somewhat loose association of the 42 and 39 kDa subunits with Iγ and the presence of the 39 kDa and 15 kDa (IP) subunits in both Iγ and Iα subcomplexes (Sazanov et al, 2000).

Other reports that give information on the special organization of complex I subunits include findings by Han et al (1989) that the 75 kDa subunit exists on both sides of the inner mitochondrial membrane and those by Patel et al (1988) which show that the 49 kDa and 30 kDa subunits are exposed to the intermembrane space. Determination of the crystal structure of complex I in Neurospora crassa and immunolabeling the 49 kDa subunit, has pinpointed it to the matrix-localized protruding arm of the complex (Guenebaut et al, 1997). This is in agreement with its position in the E.coli complex I. where it has been localized to the so-called connecting fragment which links the peripheral arm to the membrane arm (Friedrich et al, 1995; Leif et al, 1995).
Energy conversion in complex I

The main function of complex I is to transport electrons by oxidation of NADH followed by reduction of ubiquinone. This process is accompanied by the translocation of protons from the mitochondrial matrix to the intermembrane space.

(i) Iron-sulphur clusters, Flavin and Semiquinones

Several prosthetic groups catalyze these electron transport reactions, which are hypothesized to be at least one Flavin Mononucleotide (FMN), six to eight Iron-Sulfur clusters (Albracht et al, 1997) and one or two species of semiquinone (de Jong et al, 1994; Vinogradov et al, 1995).

The two types of iron-sulfur clusters encountered in complex I are \([2\text{Fe}-2\text{S}]^{(2-)}\) and \([4\text{Fe}-4\text{S}]^{(2+)}\) (Ohnishi et al, 1998). A binuclear cluster is made up of two iron atoms which are bridged by two acid-labile inorganic sulfides and ligated to four cysteinyl sulfur from the polypeptide chain of the apoprotein. Each iron is tetrahedrally coordinated to two acid labile sulfur and two cysteinyl sulfur atoms (Ohinshi, 1998). A tetranuclear cluster contains four iron atoms and four acid-labile sulfides arranged in a distorted cube structure with the iron atoms bound to the polypeptide via four cysteine sulfur ligands (Ohinshi, 1998). An important indication to the total predicted number of iron-sulfur clusters in complex I and their possible subunit locations comes from the fully conserved sequence motifs (Fearnley et al, 1992; Weidner et al, 1993). EPR analyses has led to the spectral resolution of only six distinct iron-sulfur clusters, namely clusters N1a, N1b, N2, N3, N4 and N5 associated with complex I so far (Ohinshi, 1998). While the 14 subunit
NADH:ubiquinone reductase of *E. coli* contains one FMN and up to 9 Fe-S clusters as prosthetic groups (Braun *et al.*, 1998), mammalian complex I contains at least four [4Fe-4S] clusters (N-2, N-3, N-4, N-5) and two [2Fe-2S] clusters (N-1a, N-1b). The 51 kDa FP subunit contains an FMN and a tetranuclear iron-sulfur cluster (N-3). The 24 kDa FP subunit contains a binuclear iron-sulfur cluster. The 75 kDa IP subunit contains a tetranuclear (N-4) cluster, a binuclear cluster and probably another tetranuclear cluster. The TYKY subunit contains eight conserved cysteine motifs for housing two tetranuclear clusters, while the PSST subunit has a motif of 3 cysteine residues that could ligate another tetranuclear cluster. However, the structural similarity of TYKY to bacterial low potential ferredoxins which carry two tetranuclear clusters in close vicinity, makes this subunit a less likely candidate for carrying the N-2 cluster (Brandt, 1997). Although other subunits with conserved cysteine residues such as the PGIV (Dupuis *et al.*, 1991). 49- and 30-kDa (Preis *et al.*, 1990; Fearnley and Walker, 1992) subunits have been suggested to bind to Fe-S clusters, more recent studies show that it is unlikely (Finel *et al.* 1994). Although, the exact subunit location of the various iron-sulfur clusters is not conclusively known, table 1.3 illustrating the two existing models, summarizes current knowledge on this issue.

The non-covalently bound flavin of complex I can assume three different redox states, namely, fully oxidized, semiflavlin and fully reduced (Leif *et al.*, 1995; Finel, 1993). The oxidized form of flavin has four orders of magnitude higher affinity to its specific binding site that its fully reduced form (Leif *et al.*, 1995). A strong spin-spin interaction exists
between the semiflavin and iron sulfur cluster N3, both of which are located in the 51 kDa subunit (Sled et al. 1994).

**Table 1.3. Current hypotheses on the subunit location of FMN and iron-sulphur (Fe-S) clusters in the minimal nuclear-encoded functional unit of bovine heart Complex 1**

<table>
<thead>
<tr>
<th>Name of subunit</th>
<th>Prosthetic groups</th>
<th>Prosthetic groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>75 kDa</td>
<td>1 (N1a), 2 (N4)</td>
<td>1 (N1b), 1 (N4), 1 (N5)</td>
</tr>
<tr>
<td>51 kDa</td>
<td>2 (FMN), 2 (N3)</td>
<td>1 (FMN), 1 (N3)</td>
</tr>
<tr>
<td>24 kDa</td>
<td>1 (N1b)</td>
<td>1 (N1a)</td>
</tr>
<tr>
<td>TYKY</td>
<td>2 (N2)</td>
<td>2 [4Fe-4S]</td>
</tr>
<tr>
<td>PSST</td>
<td>None or 1 [4Fe-4S]</td>
<td>1 (N2)</td>
</tr>
</tbody>
</table>

Discovered in activated bovine heart submitochondrial particles (SMP), semiquinones have also been a source of much debate (Suzuki and King, 1983). Although certain groups have concluded that there is only a single species of semiquinone (SQ) (de Jong et al. 1994), most other researchers agree that two distinct species of semiquinone (SQ\(_{Nf}\) and SQ\(_{Ns}\)) exist with fast and slow spin relaxation behaviour, respectively. SQ\(_{Nf}\) is strongly spin-coupled with cluster N2 and SQ\(_{Ns}\) is located closer to the cytosolic side membrane surface and is weakly spin-coupled with cluster N2. A third semi-quinone, undetectable by EPR has also been hypothesized (Ohnishi, 1998; Friedrich et al., 1998).

(ii) **Electron transfer in complex I**

NADH, the electron donor of complex I binds to the FMN-containing 51 kDa subunit (Deng et al., 1990). It is thought that FMN, because it can take up two electrons and
release them one at a time to one-electron acceptor such as an Fe-S cluster, is the immediate oxidant of NADH (Ragan, 1987; Weiss et al, 1991). A recent $^{32}$P photolabeling study by Yamaguchi and co-workers (2000) with purified beef heart complex I showed that the 30 kDa and 42 kDa subunits, in addition to the 51 kDa subunit were capable of binding NADH. The 39 kDa subunit as well as a 18-20 kDa subunit were shown to bind NADP(H) (Yamaguchi et al, 2000). However, since there is no evidence that complex I contains any flavin other than the FMN bound to the 51 kDa subunit, it is not likely that these other subunits with nicotinamide nucleotide binding capabilities are actually involved in binding reducing equivalents in vivo (Yamaguchi et al. 2000).

A strong homology between the 51- (Pilkington et al, 1991) and 24-kDa (Tran-Betcke et al. 1990) subunits of complex I with the $\alpha$-subunit of the NAD+-reducing hydrogenase of Alcaligenes eutrophus and high homology between the N-terminal half of the 75 kDa subunit of complex I and the $\gamma$-subunit of the same enzyme gives some insight into electron flow in this enzyme. The homology of the PSST subunit to the $\delta$ subunit of the hydrogen-hydrogenase complex of Alcaligenes eutrophus, suggests the association of the PSST subunit with the 75 kDa Fe-S protein (Robinson, 1993). Taken together, the following scheme of electron flow between the iron-sulfur clusters of complex I can be envisioned: NADH -> 51 and 24 kDa of FP -> 75 kDa of IP -> 23 (TYKY) and 20 kDa (PSST) of HP (Belogrudov and Hatefi, 1996). The proximity of the FP and IP subunits
of complex I was determined by Hatefi et al (1993) through cross-linking experiments (Fig. 1.9) and supports this model of electron transfer between the subunits.

It is known that electron transfer between FeS clusters can easily occur over distances of 1.0 to 1.5 nm (Onuchic et al, 1992). The pH dependent midpoint potential ($E_m$) value of cluster N2 is the highest among all iron-sulfur clusters in complex I and it's one electron reduction or oxidation is coupled with the binding and release of one proton in the physiological range (Ingledew and Ohnishi, 1980). Cluster N2 has also been found to be only 8-11 Å apart from one of the EPR-detectable ubisemiquinones (Vinogradov et al, 1995) and therefore has been assumed to serve as the electron donor to ubiquinone and to be intimately linked to the proton translocation mechanism of the enzyme (de Jong and Albracht, 1994).

The quinone binding site itself within complex I has been a matter of some debate. While it is now widely accepted that there are at least two ubiquinone binding sites (Vinogradov, 1993; Degli Esposti and Ghelli, 1994), their location within the complex has not been clearly defined. Darrouzet et al (1998) have shown through studies on piericidin-resistant mutants of the bacterium Rhodobacter capsulatus that the 49 kDa subunit is associated with the resistance and therefore involved in quinone binding. Their model proposes that the ubiquinone binding sites are located at the interface between the membranous and peripheral domains of complex I, whereby the polar cyclic head of the quinone is borne by the 49 kDa subunit and the hydrophobic isoprenyl tail is borne by the hydrophobic ND1 subunit (Darrouzet et al, 1998). This model attempts to accommodate the findings that ND1 and ND4 subunits may be involved in quinone
binding, as pathogenic mutations in these subunits causing complex I deficiency seem to also slightly alter the interaction of the complex with quinones and its sensitivity to rotenone (Majander et al, 1991; Degli Esposti and Ghelli, 1994). Earlier photolabeling studies on isolated complex I had actually pointed to the ND1 subunit as a candidate subunit for the isoprenyl tail binding site (Earley et al, 1987). The latest studies however, dispute this claim by showing through photolabeling studies on mitochondrial electron transport particles that it is the PSST subunit which couples electron transfer from its iron-sulfur cluster N2 to quinone (Schuler et al, 1999). Schuler and colleagues showed that the ND1 subunit is probably not directly involved in quinone binding because complex I inhibitors including rotenone were not able to prevent photoaffinity labeling of this subunit at the same levels that totally blocked labeling of the high affinity PSST site. This is substantiated by the fact that an isolated peripheral fragment of complex I, devoid of all the ND (mitochondrially encoded) subunits, is able to catalyze NADH-quinone oxidoreduction (Friedrich et al, 1989; Finel et al, 1992). Recent work based on a structural element containing a weak sequence motif that is common to the quinone sites of bacteria, mitochondrial bc1 and photosystem I, points to the ND4 and ND5 subunits as candidates for the locations of Q sites in complex I, either as a pair harbouring two separate sites or as a pair forming a single site (Fisher and Rich, 2000). The two likely candidates for quinone binding at this point in study seem to be the 49 kDa subunit of the IP fraction and PSST subunit of the HP fraction.
Figure 1.9. Structural model of complex I. The complex is portrayed as an elbow-shaped entity with a membrane bound intrinsic arm and an extrinsic arm which protrudes into the matrix. The numbers inside the subunits represent their molecular mass. NADH binds to the 51 kDa flavoprotein subunit containing a 4Fe-4S center and electrons are passed through the 24 kDa, 75 kDa, TYKY and PSST iron sulphur proteins. Two quinone binding sites are postulated. The mtDNA encoded (ND) subunits are all part of the membrane arm of the complex and some of them are possibly involved in proton pumping. * indicates the subunits that house the redox groups.
(iii) Models for coupling electron flow with proton translocation

The oxidation of one NADH by complex I is now considered to be linked to the translocation of $4H^+$ per $2e^-$ (Galkin et al., 1999). Electron transfer in proteins occurs through either tunneling over large distances (up to 20 Å) from one redox center to the next or transfer along covalent bonds (Moser et al., 1995).

Although many mechanistic models have been proposed over the years, none of them have fully taken into consideration factors such as the existence of complex I as a multiprotein enzyme (Models by Mitchell, 1966 and Lawford and Garland, 1972), its $4H^+/2e^-$ stoichiometry (Model by Suzuki and King, 1983), the $E_{m7}$ potential of cluster N-2 being $-150$ mV in bovine complex I and the $pK_a$ values of FMN and FMNH$_2$ (Model by Krishnamoorthy and Hinkle, 1988), thermodynamic requirements (Ragan, 1990), reoxidation of the internal ubihydroquinone by cluster N-2 (Model by Weiss et al. 1991), movement of ‘translocating’ semiquinones across the membrane and protonation/deprotonation occurring at opposite sides of the membrane (Model by Vinogradov, 1993), contributions to the proton translocation process by the electron input part of complex I containing FMN and the Fe-S clusters and other thermodynamic problems associated with the dual Q-gated pump (Model by Degli Esposti and Ghelli, 1994). The proposed hypotheses on the energy transducing mechanism of complex I can be roughly divided into two types of either direct (Brandt, 1997; Dutton et al., 1998) or indirect (Belogrudov and Hatefi, 1994; Yagi et al., 1998) coupling between electron transport and proton translocation. Therefore, only the two most recent hypotheses representing each model will be discussed.
The direct coupling model proposed by Dutton et al (1998) is based on the Q-cycle with a few variations to the already established mechanism for complex III. Their model proposes two quinone binding sites, \( Q_{nz} \) and \( Q_{nx} \), that can exchange \( Q/QH_2 \) with the membrane pool \( (Q_{pool}) \), with the \( Q_{nz} \) site having access to the protons on the matrix side of the membrane and the \( Q_{nx} \) having access to protons on the cytosolic side of the membrane. In addition, a non-pool exchangeable \( Q_{ny} \) occupies a site that can assume either of two different conformations between the other quinone binding sites and may even be covalently bound. One conformation provides access to the protons on the matrix side of the membrane, presumably through a channel or a pore. The other conformation provides access to protons on the cytosolic side of the membrane. This \( Q_{nx} \) site may very well be the novel redox group detected by Weiss’s group (1997) within the membrane arm of complex using ubiquinone-10 depleted complex I purified from \( N.\text{crassa} \). Only quinones are used in this model to manipulate proton motion (although flavins are equally capable of binding and releasing protons, their localization in the transmembrane domain seems to make it quite unlikely) (Dutton et al, 1998).

According to the model by Dutton et al (1998) (Fig. 1.10), electron transfer from NADH in the matrix to the N-2 Fe-S cluster close to the matrix-membrane interface, proceeds through the large number of flavin and Fe-S redox cofactors forming a long transfer chain. As N-2 is reduced, it injects a single electron into a \( Q \) drawn from the membrane pool into the nearby \( Q_{nz} \) site, generating an unstable transition state semiquinone (SQ) (Dutton et al, 1998). The SQ at the \( Q_{nz} \) site near the matrix acts as a strong oxidant to pull two electronic charges across the membrane via a \( Q \) at the \( Q_{ny} \) site.
Figure 1.10. A hypothetical model for direct energy conversion in complex I. Complex I is represented as a transmembrane protein with a NADH, FMN and iron-sulphur complex which delivers electrons to the [4Fe-4S] cluster N2. Quinone binding sites are Q\textsubscript{nz} and Q\textsubscript{nx}; Q\textsubscript{pool} is the membrane pool; Q\textsubscript{ny} is a non-pool exchangeable quinone; SQ represents semiquinone. The sequence proceeds as follows: 1) NADH arrives at Complex I, Q\textsubscript{nz} and Q\textsubscript{nx} sites can exchange with pool 2) Reduction of N2 occurs by the NADH subcomplex 3) N2 reduces a Q drawn from Q\textsubscript{pool} into Q\textsubscript{nz} forming an unstable, transition SQ 4) The SQ in the Q\textsubscript{nz} site oxidizes the QH\textsubscript{2} fixed in the Q\textsubscript{ny} site. SQ of the Q\textsubscript{nz} site is reduced to QH\textsubscript{2} drawing two protons from the matrix and QH\textsubscript{2} of Q\textsubscript{ny} site is oxidized to SQ releasing protons to the cytosol 5) QH\textsubscript{2} in the Q\textsubscript{ny} site adopt a geometry with access to cytosolic protons and the newly formed SQ assumes a geometry with access to protons in the matrix 6) QH\textsubscript{2} in the Q\textsubscript{nx} site can reduce the SQ in the Q\textsubscript{ny} site back to QH\textsubscript{2} and protons from the channel to matrix are bound and protons are released from the Q\textsubscript{nx} site. The top half of the figure shows that as the QH\textsubscript{2} is oxidized in the Q\textsubscript{nx} site two protons are released into the cytosol leaving a SQ anion. No further protons are released from Q\textsubscript{nx} in the bottom half of the figure. The reduced Q\textsubscript{ny} site assumes the original geometry and Q\textsubscript{nz} exchanges with the pool and steps 2-6 are repeated. Adapted from Dutton et al, 1998.
from the \( Q_{nx} \) site near the cytosolic side (Dutton et al., 1998). The \( QH_2 \) in the \( Q_{ny} \) site is therefore oxidized by the SQ in the \( Q_{nz} \) site. Thereby, as the SQ of the \( Q_{nz} \) site is reduced to \( QH_2 \), it binds 2 protons, ultimately drawn from the matrix (Dutton et al., 1998). As the \( QH_2 \) of the \( Q_{ny} \) site is oxidized to SQ, one (or two) protons are released to the cytosolic channel (Dutton et al. 1998). While \( QH_2 \) in the \( Q_{ny} \) site is favoured to adopt a conformation with access to cytosolic protons, the newly formed SQ rapidly assumes a conformation with access to protons in the matrix (Dutton et al., 1998). In this geometry, the \( QH_2 \) in the \( Q_{nx} \) site can reduce the SQ in the \( Q_{ny} \) site back to the \( QH_2 \) and in the process one or two protons from the channel to matrix are bound (Dutton et al. 1998). At the same time, one or two protons are released from the \( Q_{nx} \) site (Dutton et al. 1998).

The intervening \( Q_{ny} \) site therefore acts as a proton pumping element, quite like the proton pump of complex IV (Dutton et al., 1998). The overall reaction describes two electrons carried by complex I from NADH to Q catalyzing the translocation of 4 or 6 proton charges from the mitochondrial matrix to cytosol (Dutton et al., 1998). The experimentally observed SQ states \( Q_{Nf} \) and \( Q_{Ns} \) mentioned before (Vinogradov et al. 1995) are said to correspond to \( Q_{ny} \) and \( Q_{nz} \), respectively (Dutton et al., 1998). This model also supports the idea of two clear classes of inhibitors corresponding to the diffusible quinone binding sites \( Q_{nx} \) and \( Q_{nz} \), as has already been suggested by some experimental studies (Friedrich et al., 1994).

The indirect coupling hypothesis put forth by Belogrudov and Hatefi (1994) is based on analogy to the ATP synthase complex in which there is no proton carrier and energy coupling between the catalytic and the proton-translocating sectors appears to take place
via conformational changes of the subunits. This mechanism is supported by their findings that the proximities of the three subunits of FP to one another, the proximity of the 51 kDa subunit of FP to the 75 kDa subunit of IP and the proximities of all of the IP subunits to one another and to some of the HP subunits are altered when the catalytic sector of complex I is reduced by NADH or NADPH (Belogrudov and Hatefi, 1994). That is, the proximity changes observed from NAD(P)H treatment of complex I involved not only those subunits that contain electron carriers but also those that are devoid of them (Belogrudov and Hatefi, 1994). Conformational changes beginning with the reduction of FP, transmitted to subunits of IP and finally to some of the subunits of the membrane sector of complex I (HP) serve as the device by which the energy derived from electron transfer through the catalytic components of complex I is transduced and conveyed to the subunits of the membrane sector (Belogrudov and Hatefi, 1994). The resulting changes in the pKₐ of appropriate residues induced by these conformational changes lead to proton uptake and release on opposite sides of the membrane (Belogrudov and Hatefi, 1994).

Recent cryo-electron crystallization of two subcomplexes of bovine complex I lends support to both mechanisms of energy transduction (Sazanov and Walker, 2000). The large distances seen between subunits ND5 and ND4 (where proton pumping is thought to take place) and the Iα subcomplex (where NADH oxidation and electron transport reactions occur) argue in favour of the indirect mechanism of proton translocation involving long-range conformational changes in the enzyme (Sazanov and Walker, 2000).
On the other hand, the ND2 (proton-translocating) and ND1 (ubiquinone-binding) subunits seen in the Iγ subcomplex could participate in a direct coupling mechanism (Sazanov and Walker, 2000).

**Complex I Inhibitors**

The sixty or so different families of compounds known to inhibit complex I include those of both natural and commercial origin (reviewed by Degli Esposti. 1998). The two most potent natural inhibitors of complex I are rotenone and piericidin. Rotenone, the classical inhibitor of complex I belongs to a family of isoflavonoids extracted from *Leguminosae* plants (Degli Esposti, 1998). Rotenone inhibition is time-dependent with a $K_i$ as low as 1 nM being reported recently (Grivennikova *et al*. 1997). Piericidins, a group of potent complex I inhibitors are 2,3-dimethoxy-4-hydroxy-5-methy-6-polyprenyl-pyridine antibiotics produced by some *Streptomyces* strains (Miyoshi, 1998). Their close similarity to ubiquinones renders Piericidin A to be an effective inhibitor of complex I with a $K_i$ ranging between 0.6 and 1 nM (Degli Esposti, 1998). Complex I inhibitors are thought to bind at or close to quinone binding sites and kinetic studies have suggested that these inhibitors can be grouped into three classes, namely rotenone, piericidin A and capsaicin (Degli Esposti *et al*. 1994). Capsaicin which is part of the vanilloid family is a pungent extract from red peppers and although it is a relatively weak inhibitor of complex I compared to other natural products, it shows the rare property of competitive inhibition versus quinone substrates in complex I (Degli Esposti, 1998). A study by Okun *et al* in
1999 showed through two independent approaches that all tested hydrophobic inhibitors shared a common binding domain with partial overlapping sites. While the rotenone site overlaps with both the piericidin A and capsaicin site, the latter two sites were found not to overlap (Okun et al., 1999). Other natural inhibitors of complex I are annonaceous acetogenins such as rolliniastatin, myxobacterial antibiotics such as stigmatellin, and antibiotics such as cochlioquinone B (Degli Esposti, 1998). Synthetic and commercial inhibitors of complex I include various short-chain ubiquinones, pesticides, drugs, neurotoxins and fluorescent dyes (Degli Esposti, 1998).

**Complex II: The Succinate-ubiquinone oxidoreductase complex**

Complex II refers generally to succinate-ubiquinone oxidoreductase (SQR) or menaquinol-fumarate oxidoreductase (QFR), membrane-bound enzyme complexes which are alike in both structure and function (Maklashina et al., 1999). While SQR couples the oxidation of succinate to fumarate in the Krebs cycle and the subsequent reduction of ubiquinone in aerobic organisms, QFR oxidizes menaquinol and reduces fumarate to succinate as part of an anaerobic respiratory chain (Maklashina et al., 1999). The SQR complex does not translocate protons, and therefore only feeds electrons to the electron transport chain (Hagerhall, 1997; Ackrell et al., 2000). It consists of a peripheral domain, exposed to the matrix in mitochondria and a membrane-integral domain that spans the membrane (Hagerhall, 1997). The peripheral part, which contains the dicarboxylate binding site, is composed of a flavoprotein (FP; 64-79 kDa) subunit, with one covalently bound FAD, and an iron-sulfur protein (IP; 27-31 kDa) subunit containing three iron-
sulfur clusters, a [2Fe-2S] cluster denoted S1, a [4Fe-4S] cluster denoted S2 and a [3Fe-4S] cluster denoted S3 (Ackrell et al., 2000). The membrane-integral domain functions to anchor the FP and IP subunits to the membrane and is required for quinone reduction and oxidation (Ackrell et al., 2000). The anchor domain shows variability in composition and primary sequence. It consists of one larger or two smaller hydrophobic peptides and either one or two protoheme IX groups, with hexa-coordinated iron, or no heme at all (Ackrell et al., 2000). Inhibitors of this complex include Theonyltrifluoroacetone (TTFA), 3'-methyl carboxin and 2-n-heptyl-4-hydroxyquinoline N-oxide (HQNO) (Hagerhall, 1997). Complex II is the only one of the respiratory chain enzymes which has the same general composition in pro- and eukaryotes (Ackrell et al., 2000). This is in contrast to complexes I, III and IV of higher organisms which have various numbers of subunits present in addition to the “minimal functional units” found in their prokaryotic counterparts.

**Complex III: The Ubiquinol-ferricytochrome c oxidoreductase complex**

The second proton translocation site within the electron transport chain is at complex III or the QH$_2$-cytochrome c reductase complex. The reaction catalyzed by complex III is:

\[
QH_2 + 2H^+ + 2\text{cyt } c_{\text{ox}} \rightarrow Q + 4H^+ + 2\text{cyt } c_{\text{red}},
\]

where Q represents ubiquinone (Voet and Voet, 1990). Recently, the complete crystal structure of the cytochrome bc$_1$ complex from bovine heart mitochondria was determined (Iwata et al., 1998). Mammalian complex III is composed of eleven subunits with known amino acid sequences (Yu et al., 1998). Bacterial homologs are found of only the three of the subunits and these are the
subunits that carry redox centres (Saraste, 1999). The key redox components are
subunits with two $b$ cytochromes ($b_{565}$ and $b_{562}$), which are mitochondrially encoded
and have eight transmembrane helices with two heme $b$ groups sandwiched between
helices B and D; a membrane-anchored cytochrome $c_1$ and a membrane-anchored FeS
protein carrying a Rieske-type center ($Fe_2S_2$) and a ubiquinone (Saraste, 1999). Most of
the other eight subunits are small proteins that surround the metalloprotein nucleus, but
the two largest membrane spanning proteins in the complex, subunits I and II termed
“core proteins” face the mitochondrial matrix and are homologous to mitochondrial
processing peptidases which function in protein import (Iwata et al. 1998). Thus,
complex III, which exists as a macromolecular dimer within the membrane in bovine heart
mitochondria, may be multifunctional (Iwata et al., 1998).

Coupling the redox reaction to the generation of a proton gradient across the
membrane is performed by a mechanism called the Q cycle (Yu et al., 1998). According to
this model, one electron is transferred from the ubiquinol at the $Q_o$ site (near the
cytoplasmic side of the inner mitochondrial membrane) to the Rieske iron-sulfur centre
and then to cytochrome $c$ via cytochrome $c_1$ (Yu et al., 1998). The ubisemiquinone which
is generated then reduces cytochrome $b_{566}$ heme which in turn transfers an electron to the
cytochrome $b_{562}$ heme (Yu et al., 1998). A ubiquinone or ubisemiquinone bound at the $Q_i$
site (near the matrix side of the membrane) then oxidizes the reduced $b(562)$ heme (Yu et
al., 1998). Proton translocation is therefore the result of deprotonation of ubiquinol at
the $Q_o$ site and protonation of ubiquinol at the $Q_i$ site (Yu et al., 1998)
Specific inhibitors of the complex include antimycin, myxothiazol and stigmatellin (Yu et al, 1998). Patients with complex III deficiency often show lactic acidosis, muscle weakness, ataxia, exercise intolerance, ocular myopathy or a multisystem disorder (Kennaway, 1988; Slipetz et al, 1991).

**Complex IV: The cytochrome c oxidase complex**

Cytochrome c oxidase (COX) or complex IV is a Y-shaped complex which exists as a large macromolecular dimer (Saraste, 1999). The enzyme catalyzes the following reaction:

\[
4 \text{cyt}_\text{red} + 4 \text{H}^+ + \text{O}_2 \rightarrow 4 \text{cyt}_\text{ox} + 2\text{H}_2\text{O}
\]

It is the only non-equilibrium, irreversible, step in oxidative phosphorylation (Voet and Voet, 1990). The crystal structure of bovine heart COX at 2.8 Å resolution was first obtained by Yoshikawa and co-workers (1998). This complex consists of 13 subunits, with the three largest subunits which constitute the catalytic core of the complex, encoded by mtDNA and the ten remaining subunits encoded by nuclear DNA (Robinson, 1998a). The core of this complex is made up of the products of the COX I, II and III genes which are mitochondrially encoded (Robinson, 1998a). Complex IV contains four redox centers, two a-type hemes (a and a₃) and two copper ions (Cuₐ and Cuₐ) (Saraste, 1999). The complex IV substrate, cytochrome c, is a water-soluble hemoprotein that donates electrons on the cytoplasmic side of the mitochondrial inner membrane (Saraste, 1999). Subunit II has a dinuclear, mixed valence copper center (Cuₐ), which is the first site to receive electrons from cytochrome c (reviewed by Capaldi, 1990). Subunit I contains the active site which contains the heme iron and copper that are used to reduce O₂ into two water molecules.
Electrons from subunit II are transferred to a low-spin heme (cytochrome a) in subunit I, and then to the bimetallic cytochrome a3/CuB active site (reviewed by Capaldi, 1990). The protons needed for this reaction are taken from the mitochondrial matrix side through two channels (reviewed by Calhoun et al, 1994). The same channels are used to pump one proton per electron across the membrane (reviewed by Calhoun et al, 1994). Although subunit III has no prosthetic groups, it is vital to the functioning of the other two core subunits (Robinson, 1998a). Inhibitors of complex IV include cyanide, carbon monoxide, sulfide and azide which interact with the heme redox centers (Tyler, 1992). Isolated cytochrome oxidase deficiency can be divided into five basic categories, namely, Fatal infantile lacticacidemia, Classical Leigh’s disease, Saguenay Lac St Jean Leigh’s disease, cardiomyopathy and mitochondrially encoded defects such as Kearns-Sayre, LHON and CPEO (Robinson, 1998a).

**Complex V: The ATP synthase**

The proton motive force generated by the electron transport chain is utilized by the H+-ATP synthase to generate ATP from ADP and inorganic phosphate (Leslie et al, 1999). This 16 subunit mammalian complex appears as a mushroom shaped structure with three parts designated \( F_o \) (membrane associated), \( F_1 \) (water soluble) and a stalk (Leslie et al, 1999). This general structural organization is in congruence with the crystal structure of F1-ATPase obtained at 2.8 Å resolution from bovine heart mitochondria (Abrahams et al, 1994). The \( F_o \)-membrane associated component forms a channel through which protons can traverse (Leslie et al, 1999). The electrochemical potential energy is
then utilized by the F₁ component, which faces the interior of the inner mitochondrial membrane to form ATP (Leslie *et al.*, 1999). The F₁ headpiece is composed of five subunits with the following stoichiometry: $\alpha_3\beta_3\gamma\delta\epsilon$ (Leslie *et al.*, 1999). The isolated F₁ component which contains the catalytic nucleotide binding-site ($\beta$-subunits) is as active as an ATPase and is therefore referred to as F₁-ATPase (Leslie *et al.*, 1999). It is now widely accepted that the three catalytic sites alternate between three different states - open, loose and tight - which have differing affinities for nucleotides (Leslie *et al.*, 1999). Conformational changes required for inter-conversion between the three types of catalytic sites is achieved by the rotation of the three catalytic $\beta$-subunits relative to the single copy subunits $\gamma$, $\delta$ and $\epsilon$ (Leslie *et al.*, 1999). Efrapeptins and aurovertins are potent inhibitors of the F₁-ATPase catalytic sites (Leslie *et al.*, 1999). The F₀ and stalk domains of the mammalian structure are composed of eight subunits: a, $b_2$, c₁₀, d, e, OSCP (Oligomycin Sensitivity Conferring Protein), factor A6 and A6L (Pedersen and Amzel, 1993). Subunits A and A6L are mitochondrially encoded while the remaining subunits are nuclear encoded (Papa *et al.*, 1992). Regulation is achieved by an inhibitory protein IF₁, which binds to the b-subunit in the absence of a proton gradient, thereby preventing the ATPase complex from running in reverse and hydrolyzing ATP (reviewed by Pedersen and Amzel, 1993). Clinical consequences of defects in the ATP synthase complex are best represented by NARP and Leigh's syndrome (reviewed by Schapira and Cock, 1999).
Table 1.4. Composition and genetic origin of mitochondrial respiratory chain (OXPHOS) subunits

<table>
<thead>
<tr>
<th>Complex</th>
<th>Subunits</th>
<th>Nuclear Encoded</th>
<th>mtDNA Encoded</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>43</td>
<td>36</td>
<td>7 (ND1, ND2, ND3, ND4, ND4L, ND5, ND6)</td>
</tr>
<tr>
<td>II</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>III</td>
<td>11</td>
<td>10</td>
<td>1 (cytochrome b)</td>
</tr>
<tr>
<td>IV</td>
<td>13</td>
<td>10</td>
<td>3 (cytochrome oxidase I, II and III)</td>
</tr>
<tr>
<td>V</td>
<td>14</td>
<td>12</td>
<td>2 (ATPase 6 and 8)</td>
</tr>
</tbody>
</table>

Part III. Mitochondrial disorders

**Typical symptoms of defects in energy metabolism**

A failure of mitochondria to produce ATP at normal rates because of genetic defects or hypoxic situations leads to a rapid acceleration of glycogenolysis and glycolysis brought about by increased ADP, AMP and inorganic phosphate concentrations (Robinson, 1998a). This increase is brought about through the initial defense of ATP levels through the creatine phosphokinase and adenylyl kinase equilibria (Robinson, 1998a). The glycolytic response to a shortage of mitochondrial ATP production results in an increase of lactic acid production (Robinson, 1998a). While the lactate production is increased, inhibition of respiratory chain function alters the redox state of the cell to make it more reduced in both the cytosolic and mitochondrial compartments (Robinson, 1998). The net result of this is a change in the ratio of lactate to pyruvate (L/P ratio) both in the intracellular and extracellular locations (Robinson, 1998a). Depending on the severity of the defect, L/P ratios will be elevated from 10:1 - 20:1 in normal individuals to as much as 25:1 - 40:1 in affected individuals and to as much as 100:1 in extreme cases (Robinson.
Collectively, patients with lacticacidemia represent about 1 in 5000 of the population (Robinson, 1993).

One of the most common features seen in patients with defects in energy metabolism is an increase in blood or CSF (cerebrospinal fluid) lactate (Robinson, 1998a). If lactic acid levels are elevated there appears to be some correlation between the observed elevation and the severity of symptoms in the patient (Robinson, 1998a). The three major disorders of energy metabolism, namely, deficiency of the pyruvate dehydrogenase complex, deficiency of NADH-ubiquinone oxidoreductase (complex I) and deficiency of cytochrome c oxidase (complex IV), seem to undertake the following progression of severity: Fatal Infantile Lactic Acidosis>Leigh disease>Psychomotor Retardation>Ataxia>Retinal Degeneration (Robinson, 1993). These disorders are usually nuclear in origin (Robinson et al, 1987; Robinson et al, 1990; Tatuch et al, 1992). Defects involving the oligomycin-sensitive ATPase (complex V) also follow this pattern. although there are no patients with fatal neonatal lactic acidosis (Tatuch and Robinson, 1993). A second set of symptoms that are common and characteristic of maternally inherited disorders (discussed below) include sensorineural hearing loss, muscular weakness, dementia and stroke-like episodes (Shoffner et al, 1995; Robinson, 1993). Other symptoms include external ophthalmoplegia, ptosis, exercise intolerance, myoclonus epilepsy and hyperventricular cardiomyopathy (Robinson, 1998).
**Mitochondrial Respiratory Chain Diseases**

Mitochondrial defects occur in a wide variety of degenerative diseases, errors in metabolism, aging and cancer. Abnormalities in mitochondrial metabolism encompass defects of fatty acid oxidation, tricarboxylic acid cycle enzymes as well as the respiratory chain and oxidative phosphorylation (OXPHOS) systems. Respiratory chain defects represent the most common biochemical deficiency of mitochondrial metabolism causing diseases.

(i) MtDNA associated diseases

Various pathogenic mtDNA abnormalities have been identified, the most common being (i) large-scale rearrangements (single deletions, multiple deletions, duplications), (ii) point mutations, insertions and deletions in transfer RNA (tRNA), (iii) point mutations in ribosomal rRNA (rRNA) and (iv) point mutations in protein encoding genes for complex I, cytochrome b, Complex IV and Complex V (Hanna and Nelson, 1999). Almost all of these disorders are characterized by mtDNA heteroplasmy. This is a state characterized by mutated and wild-type genomes co-existing in variable proportions and with tissue-to-tissue and cell-to-cell differences within the same individual. The severity of symptoms has been found to correlate to the percentage heteroplasmy in affected individuals with tissue heteroplasmy also being observed (Robinson, 1998a). The frequency of mtDNA point mutations is known to be approximately ten times that of nuclear mutations. This has been attributed to the high level of oxygen free radicals present in the mitochondria resulting from oxidative metabolism, the lack of adequate
DNA repair enzymes in the mitochondria and the absence of protective histone proteins (Brown and Wallace, 1994). The various primary defects in mtDNA described above can be found in association with many different clinical phenotypes such as Kearns-Sayre syndrome (KSS), chronic progressive external ophthalmoplegia (CPEO), mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (MELAS), myoclonus epilepsy with ragged red fibres (MERRF), mitochondrial neurogastrointestinal encephalomyopathy (MNGIE), neurogenic muscular weakness, ataxia and retinitis pigmentosa (NARP), Lebers hereditary optic neuropathy (LHON) and Leigh's syndrome (Brown and Wallace, 1994). While some mtDNA defects associate with multisystem disorders such as KSS, MELAS or NARP, others are surprisingly tissue specific such as that seen in LHON where the disease is confined to just the optic nerve (Brown and Wallace, 1994).

(ii) Nuclear DNA associated diseases

Little is known about the nuclear gene defects that cause respiratory chain dysfunction. The first nuclear gene defect reported to cause respiratory chain dysfunction was in the flavoprotein subunit of complex II, identified in two siblings with Leigh's syndrome and severe deficiency of complex II activity (Bourgeron et al, 1995). Precise gene defects have also been identified in the A1DQ, TYKY, PSST and the 51 kDa subunits of complex I (van den Heuvel et al, 1998; Loeffen et al, 1998c; Triepels et al. 1999b; Schuelke et al, 1999). Linkage analysis in pedigrees with autosomal dominant CPEO (mtDNA deletion disorders) has identified two loci, one on chromosome 10q23.3-
24.3 and the other on chromosome 3p14.1-21.2 (Suomalainen et al., 1995; Kaukonen et al., 1996), but the gene has not yet been identified. Complementation studies have shown that mtDNA depletion syndrome (clinical phenotypes include fatal infantile hepatopathy, fatal infantile myopathy and a more benign infantile myopathy) and isolated COX deficiency (Leigh's syndrome), are nuclear gene disorders (Carrozzo et al. 1998; Munaro et al., 1997).

Diseases are also caused by mutations of nuclear genes encoding non-OXPHOS mitochondrial proteins such as Frataxin in Friedreich's ataxia (Priller et al., 1997; reviewed by Leonard and Schapira, 2000), Surf-1 (an assembly factor of COX) in COX-deficient Leigh's syndrome (Zhu et al., 1998), ATP 7B in Wilson's disease (Lutsenko and Cooper, 1998) and others.

(iii) Mitochondrial respiratory chain disorders associated with neurological diseases

Huntington's disease is inherited in an autosomal dominant fashion and brain tissue from patients with this disease has shown a 53-59% decrease in complexes II/III activity, a 32-38% reduction in complex IV activity as well as elevated L/P ratios in cerebrospinal fluids (Gu et al., 1996). An anatomically specific complex I defect in the pars compacta has also led to the association of Parkinson's disease (autosomal dominantly inherited) with mitochondrial complex I deficiency (Bindoff et al., 1989; Schapira et al., 1990). COX deficiency has also been described in Alzheimer's disease (Simonian et al. 1994).
**Human Complex I deficiencies**

Human NADH:ubiquinone oxidoreductase deficiency can be present in a wide spectrum of biochemical and symptomatic phenotypes. The observed range of severity can range from fatal infantile lactic acidosis to adult onset exercise intolerance or optic neuropathy (Robinson, 1998). Complex I deficiency can be caused by mutations in the nuclear genome as evidenced by Mendelian inheritance or mutations in the mitochondrial genome as demonstrated by maternal inheritance (Loeffen et al, 2000). Isolated complex I deficiency is now known to be the most common mitochondrial respiratory chain defect (Loeffen et al. 2000).

**(i) mtDNA encoded defects in complex I**

Mitochondrial DNA defects can cause complex I deficiency either by disturbing the process of mitochondrial protein synthesis or by interfering at specific sites in the process of oxidative phosphorylation (Robinson, 1998a). The majority of mtDNA disease causing mutations are present in the heteroplasmic state, eventually leading to depressed rates of electron transport and ATP synthesis (Robinson, 1998a).

Diseases of the first kind have been associated with a characteristic red staining of convoluted fibres around the mitochondria in muscle when viewed with the Gomori trichrome stain (Robinson, 1998a). These fibres are called “ragged red fibres” and complex I diseases associated with them are: MELAS, MERRF, KSS, CPEO, myopathy and cardiomyopathy (MMC) and mitochondrial myopathy (MM) (Wallace, 1992). These diseases involve mutations in the mitochondrially encoded tRNA species or
mtDNA deletions and many also show a decrease in complex IV (COX) activity (Wallace, 1992).

Diseases of the second type are caused by a number of point mutations in mtDNA as well as in mtDNA encoding complex I polypeptides such as ND1, ND4, ND5 and ND6 (Shoffner et al, 1995). These mutations are known to cause LHON (with no ragged red fibres) and LHON with dystonia (basal ganglia disease) in most cases and MELAS in some (Morgan-Hughes et al, 1999).

(ii) Nuclear encoded defects in complex I

Isolated complex I deficiency in children seems in majority to be caused by mutations in the nuclear DNA (Loeffen et al, 2000). Examination of the pattern of inheritance in families with complex I deficiency has shown that that the disease appears to be autosomal, or possibly in a few cases X-linked, in which the patient can experience a wide variety of symptoms (Robinson, 1988a). The clinical presentation of isolated complex I deficiency is summarized in table 1.5.

(iii) Free radical generation in complex I deficiency

One to two percent of all electrons passing down the electron transport chain are diverted into the formation of the superoxide radicals O$_2^-$ (Wong et al, 1989). Superoxides are formed by the interaction of molecular oxygen with a species of semiquinone (UQ$_{10}$) which is a natural intermediate in one electron reduction events in electron transport through complexes I and III (reviewed by Robinson, 1998b). Superoxide is removed by
Table 1.5. The clinical presentation and incidence of isolated complex I deficiency attributed to nuclear encoded defects

<table>
<thead>
<tr>
<th>Incidence</th>
<th>Major Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most common</td>
<td>Leigh's disease (LD) +/- cardiomyopathy (WPW) or progressive deterioration - slow/fast</td>
</tr>
<tr>
<td>Fairly common</td>
<td>Fatal infantile lactic acidosis (FILA) cystic changes in white matter</td>
</tr>
<tr>
<td>Uncommon</td>
<td>Hepatopathy and tubulopathy (HT)</td>
</tr>
<tr>
<td></td>
<td>Cardiomyopathy and cataracts (CC)</td>
</tr>
<tr>
<td></td>
<td>Cataracts and developmental delay (CD)</td>
</tr>
<tr>
<td></td>
<td>Neonatal lactic academia followed by mild symptoms (MS)</td>
</tr>
<tr>
<td></td>
<td>Normal development (some with exercise intolerance and congenital Cataracts)</td>
</tr>
</tbody>
</table>

Severe cases of complex I deficiency have been shown not to assemble complex I (Robinson et al, 1990).

The enzyme Mn$^{2+}$ superoxide dismutase (MnSOD) (mitochondrial) and Cu$^{2+}$Zn$^{2+}$ superoxide dismutase (CuZnSOD) (extramitochondrial but also exists in an extracellular form) (Pitkanen et al, 1996). These enzymes are able to convert superoxides to the stable hydrogen peroxide which is diffusible across biological membranes and removed by catalase or glutathione peroxidase (Pitkanen et al, 1996). Free radicals were seen to be produced in excess in patients with complex I deficiency and MnSOD production was also seen to increase to combat the production of superoxide (reviewed by Robinson, 1998b).

Free radical production from the respiratory chain has been implicated in selective cell death or apoptosis (Korsmeyer et al, 1995). The bcl-2 family of proteins is known to serve as a protective agent against apoptosis and bcl-2 levels are increased in cells with respiratory chain defects involving complex I (Korsmeyer et al, 1995). Bcl-2 induction is
seen to proceed through a similar but different route as MnSOD induction (Korsmeyer et al. 1995). Inborn errors of OXPHOS may therefore render a cell susceptible to death by apoptosis and may be important in the pathogenesis and progression of respiratory chain deficiencies (reviewed by Robinson, 1998b).

**Objectives and Rationale**

Defects of mitochondrial metabolism result in a wide variety of human disorders which can present at any time from infancy to late adulthood and involve virtually any tissue either alone or in combination. Abnormalities of the electron transport and oxidative phosphorylation (OXPHOS) system, encoded by both mitochondrial and nuclear genomes, are probably the most common cause of mitochondrial diseases. Of these, isolated complex I deficiency is a major cause of abnormalities encountered in respiratory chain enzymes. Only a limited number of enzymatic deficiencies of one or more of the complexes of the mitochondrial electron transport chain are associated with mutations in the mtDNA. For complex I deficiency in particular, the several mtDNA mutations that have been well characterized cannot explain all cases of isolated complex I deficiencies, especially the most frequently observed phenotype which presents in infancy and expresses a highly progressive lethal course (LD). Theoretically, about equal numbers of mutations are predicted for the mitochondrial and nuclear genomes, because even though nuclear genes are subjected to a somewhat slower mutation rate, there are a greater number of nuclear genes involved in complex I. It is therefore of utmost importance that the contribution of the numerous complex I subunits to isolated complex I deficiency be
established. In the ensuing chapters, the NADH:ubiquinone oxidoreductase complex (Complex I) will be the initial object of focus. In particular, the nuclear encoded MLRQ subunit of this complex will be examined at both gene and protein levels. The strategy used to clone, characterize and identify the chromosomal locus of \textit{NDUFA4} is presented in chapter 2. Chapter 3 focuses on the biochemical characterization of this subunit by looking at its expression in human tissues and its association with other subunits in the complex. This chapter also looks at strategies that were attempted in order to study the structure and function of this subunit in complex I through bacterial protein expression and mammalian anti-sense expression studies, respectively. Finally, chapter 4 brings all these concepts together and outlines the possible role of a supernumerary subunit such as MLRQ in the context of complex I function and dysfunction in mitochondria. Thus, the objectives of this study are to characterize the gene and gene product of a nuclear encoded subunit of complex I and illustrate its role in the functioning of this first enzyme of the OXPHOS system of mitochondria.
Chapter 2

Cloning, molecular characterization and chromosomal localization of the MLRQ subunit of human NADH:ubiquinone oxidoreductase

Abstract

The genomic DNA and cDNA sequences encoding the MLRQ subunit of human NADH:ubiquinone oxidoreductase (complex I of the mitochondrial respiratory chain) have been determined. The cDNA clones contain an open reading frame of 243 bp, 53 bp of 5’-untranslated sequence (UTR) and a 3’ UTR of 181 bp. MLRQ cDNAs obtained from a number of different human tissue types such as brain, heart, liver and kidney as well as fibroblasts and cardiomyocytes were sequenced and found to be identical. Two clones namely, 2F23 and 96E24 were isolated from a P1-artificial chromosome (PAC) library by using a MLRQ cDNA probe. Fluorescence in situ hybridization (FISH) was then used to map 2F23 to chromosome 1 p21 and 96E24 to chromosome 7 p21-p22. Searching the High Throughput Genome Sequence (HTGS) database with this information revealed that the NDUFA4 gene is situated on chromosome 7, spanning approximately 6.5 kb and consisting of 4 exons. Partial sequencing of the MLRQ-like gene product from PAC clone 2F23 revealed that in contrast to NDUFA4, it is an intronless gene with several basepair changes in its coding and non-coding regions and therefore is a pseudogene.

Introduction

Among the inborn errors of metabolism, disturbances in mitochondrial energy metabolism occur in 1/10,000 live births (Bourgeron et al, 1995). Of these, isolated
complex I deficiency is the most common mitochondrial respiratory chain defect
(Loeffen et al, 2000). The clinical presentation and course of complex I deficiency can
vary greatly due to the involvement of both the mitochondrial and nuclear genomes, as
well as factors related to oxygen free radical generation in the complex (Robinson et al.
1998). Complex I defects encoded in mitochondrial DNA have been well characterized
at the molecular level (Morgan-Hughes et al, 1999). However, underlying genetic
defects can be ascribed to mtDNA in only ~ 5% of complex I patients (Liang and Wong,
1998). In many isolated complex I deficiencies, no mutation in the mitochondrially
encoded subunits of complex I have been found. These cases have been attributed to the
nuclear genome and nuclear encoded complex I deficiency appears to be an autosomal, or
possibly in a few cases, an X-linked disease in which the patients can suffer from a wide
spectrum of symptoms (Smeitink et al, 1998a).

Although the cDNA characterization of human complex I subunits has recently been
completed (Loeffen et al, 1998b; Triepels et al, 2000), the genomic structure and
chromosomal localization of many of these nuclear genes are yet to be determined.
Particularly in the hydrophobic (HP) fraction, the complete genomic DNA sequence has
only been elucidated for the NDUFA1 (MWFE), NDUFA6 (B14), NDUFS8 (TYKY) and
NDUFB9 (B22) genes (Zhuchenko et al, 1996; Dunham et al, 1999; de Sury et al, 1998;
Lin et al, 1999).

AQRQ, TYKY, PSST and the 51 kDa subunit are the only nuclear encoded subunits
in which mutations have been characterized thus far (van den Heuvel et al, 1998; Loeffen
et al, 1998c; Triepels et al, 1999b; Schuelke et al, 1999). Table 2.1 characterizes the
mutations that have been found in patients with isolated complex I deficiency. A recent
study has also for the first time reported two mutations in the AQDQ subunit associated
with combined complex I and III deficiency (Budde et al, 2000).

Table 2.1. Nuclear gene mutation in patients with isolated complex I deficiency

<table>
<thead>
<tr>
<th>AFFECTED GENE (MUTATION)</th>
<th>SYMPTOMS AT PRESENTATION</th>
<th>LACTIC ACID CONCENTRATION</th>
<th>MRI FINDINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Blood</td>
<td>Urine</td>
</tr>
<tr>
<td>(1) NDUFVI (51 kDa)</td>
<td>Vomiting, strabismus, hypotonia</td>
<td>I'</td>
<td>N'</td>
</tr>
<tr>
<td>(R59X, T423M)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) NDUFVI (51 kDa)</td>
<td>Vomiting, strabismus, hypotonia</td>
<td>I'</td>
<td>N'</td>
</tr>
<tr>
<td>(R59X, T423M)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) NDUFVI (51 kDa) Infantile myoclonic epilepsy</td>
<td>N'</td>
<td>N'</td>
<td>I'</td>
</tr>
<tr>
<td>(A341V)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) NDUFS4 (AQDQ)</td>
<td>Vomiting, failure thrive, hypotonia</td>
<td>N'</td>
<td>N'</td>
</tr>
<tr>
<td>(5-bp duplication)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) NDUFS7 (PSST)</td>
<td>Feeding problems, dysarthria, ataxia</td>
<td>N'</td>
<td>N'</td>
</tr>
<tr>
<td>(V122M)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) NDUFS7 (PSST)</td>
<td>Vomiting</td>
<td>N'</td>
<td>N'</td>
</tr>
<tr>
<td>(V122M)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) NDUFS8 (TYKY)</td>
<td>Feeding difficulties, hypotonia, episodic apnea and cyanosis</td>
<td>I'</td>
<td>I'</td>
</tr>
<tr>
<td>(P79L, R102H)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*N - normal; I - increased; Sl.I - slightly increased
Patients (1, 2 & 3) (Schuelke et al, 1999); Patient (4) (van den Heuvel et al, 1998);
Patients (5 & 6) (Triepels et al, 1999b); (7) (Loeffen et al, 1998c)

The characterization of new human genes such as NDUFA4 (MLRQ) has greatly
benefited from bioinformatics' tools and databases such as dbEST, UniGene and HTGS,
which are divisions of GenBank. GenBank is a public sequence database that
incorporates all known nucleotide and protein sequences, primarily through the direct
submission of sequence data from individual laboratories and from large-scale sequencing projects. GenBank ensures comprehensive international coverage by data exchange with the DNA DataBank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL) and GenBank at the National Center for Biotechnology Information (NCBI) (Benson et al, 1999). There are approximately 7,376,000,000 bases in 6,215,000 sequence records as of April 2000 in GenBank (www.ncbi.nlm.nih.gov/Genbank).

Expressed Sequence Tags (ESTs) are short (usually about 300-500 bp), single-pass sequence reads from mRNA (cDNA) which are produced in large batches and represent a snapshot of genes expressed in a given tissue or cDNA library by "tagging" them (Boguski et al, 1993). ESTs are a major source of new sequence records and genes in GenBank and the discovery of new members of a gene family is facilitated by retrieving ESTs homologous to a known cDNA sequence of a different organism, from the EST database (dbEST) (www.ncbi.nlm.nih.gov/dbEST). In order to organize the EST data in a useful fashion, NCBI created the UniGene collection of unique human genes and mouse genes (www.ncbi.nlm.nih.gov/schuler/UniGene). UniGene starts with entries in the primate or rodent division of GenBank, combines these with ESTs of that organism and creates clusters of sequences that share virtually identical 3' untranslated regions (3' UTRs). In this manner, the millions of human ESTs in GenBank are reduced ~ 20 fold to sequence clusters, each of which may be considered as representing a single human gene (Benson et al, 1999). The High Throughput Genomic Sequences (HTGS) are unfinished large-scale genomic records that are in transition to a finished state, after which they will be placed in the appropriate organism division. Search and retrieval of sequence data over the world wide web (www) is performed most commonly by using the BLAST
family of search programs. Each BLAST search can be customized according to the required information such as blastn for aligning nucleotide query sequences against nucleotide sequences and searching databases such as nr (non-redundant), HTGS, dbEST, epd (eukaryotic promoter database) etc; blastx for aligning nucleotide query sequences translated in all frames against protein sequences and blastp for aligning amino acid query sequences against protein sequences and searching databases such as nr or swissprot.

The rationale behind the characterization and molecular screening of the human homolog of MLRQ was driven by a couple of factors. Linkage and complementation studies performed by Scheffler and Day (Day et al, 1982) had long suggested the existence of two X-linked subunits. The MWFE subunit has already been established as one of these possibilities (Zhuchenko et al, 1996). MLRQ warranted investigation as the other candidate subunit for X-linkage, when a search of the UniGene database revealed that researchers at the Whitehead institute had mapped a MLRQ-like gene to the human X-chromosome through EST sequencing. These reasons, along with the observation that there is a preponderance of affected males in certain cases of complex I deficiency (Orstavik et al, 1993), raised the possibility that the MLRQ gene may be situated on the X-chromosome.

The nuclear encoded MLRQ subunit was found to coprecipitate with bovine complex I as a 9 kDa protein (Walker et al, 1992). It’s hydropathy profile suggests that it can be folded into a membrane spanning α-helix of the HP fraction and that it interacts with other subunits of complex I through the more hydrophilic parts of its sequence (Walker et al, 1992).
Characterization and mapping of the human genes for nuclear-encoded complex I subunits is essential to the understanding of the genetic mechanism of the disease as well as to the understanding of the structure and function of complex I.

**Materials and methods**

**Part I. Molecular Characterization of MLRQ cDNA in various tissues, cells and patient cell lines**

**Tissue culture of cardiomyocytes and fibroblasts**

Human fibroblasts were cultured in Eagle's α-minimal essential medium supplemented with 10% (v/v) fetal calf serum and glucose to a final concentration of 4.5 mM. Cardiomyocytes were cultured in Iscoves medium supplemented with 10% (v/v) fetal calf serum.

**RNA isolation from cells and tissues**

Total RNA was isolated from human tissues and cultured cell lines using TRIzol Reagent (Gibco BRL) according to the manufacturer's protocol.

**cDNA synthesis and PCR**

First-strand cDNA was produced by reverse transcription of mRNA isolated from human fibroblasts, cardiomyocytes and from heart, brain, kidney and liver autopsy tissue using an oligo (dT$_{12-18}$) primer and Superscript II reverse transcriptase (Gibco BRL) with appropriate reverse transcription controls.

Oligonucleotide primers forward (5'-TTTAGCTTAGGGCCTGGTGCT-3') (corresponding to residues 8-30 of the bovine cDNA sequence) and reverse
(5'-ACAATACAGATCTCAACATG-3') corresponding to residues 405-426 of the bovine cDNA sequence) designed from a human EST (Expressed Sequence Tag) sequence (accession number H65454) were used to amplify the cDNA sequence of the 5' UTR and the coding region. PCR was carried out for 35 cycles (1 min denaturation at 94°C, 1 min annealing at 41°C, 1 min primer extension at 72°C, and a 10 min primer extension in the final cycle).

In order to amplify the 3' region of the gene, RACE (Rapid Amplification of cDNA Ends) PCR was employed using primer MLRQ 249 (5'-CCAATGATCAATACAAAG TTC-3') and Oligo dT₁₂₋₁₈ with an XbaI site (5'-TTTTTCTAGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT-3'). PCR was carried out under the same conditions as before except with an annealing temperature of 52°C.

The full-length MLRQ cDNA sequence was amplified from the cultured cell lines, tissues and a commercial human heart cDNA library by using forward primer MLRQ (5'-CCTGGTGCTAGGTCGG-3') and reverse primer MLRQ 478 (5'-AAGTTTCAGTTATTTATTGATTAA-3'). PCR conditions were similar to those outlined above except the annealing temperature was 52°C.

cDNA cloning and sequencing of MLRQ

The PCR products were gel purified and subcloned into TA™ cloning vector (Invitrogen) and competent DH5α E.coli cells were transformed with the construct. Blue/White screening was used to pick the colonies with the insert of interest and LB media supplemented with ampicillin (100 µg/ml) was inoculated with these colonies. Cells were then harvested after overnight growth at 37°C and plasmid minipreps were
carried out on the cultures according to Maniatis et al (1989). The resulting DNA was sequenced using the T7 DNA sequencing kit (Amersham Pharmacia) by the dideoxy chain termination method with universal forward and reverse primers according to manufacturer’s instructions.

**Mutational screening of complex I deficient patients**

Patients with various complex I clinical presentations were screened for possible mutations in the MLRQ cDNA sequence. These patients had abnormally high L/P ratios and were diagnosed as complex I deficient. Sequencing was initially carried out with cycle-sequencing kit (Amersham Pharmacia), but due to a series of false positive mutations, the MLRQ cDNA was amplified using PFU (Stratagene), subcloned and sequenced as described above using T7 polymerase. The following patients were screened for mutations from cultured fibroblasts.

**Table 2.2. Summary of clinical information on patients screened for mutations in the cDNA of NDUFA4**

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Clinical Presentation</th>
<th>L/P ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>5624</td>
<td>Hepatomegaly with Renal Tubulopathy (HT)</td>
<td>93.132 ± 21.781 (5)</td>
</tr>
<tr>
<td>7937</td>
<td>Fatal Infantile Lacticacidosis (FILA) and Cardiomyopathy</td>
<td>71.776 ± 13.449 (5)</td>
</tr>
<tr>
<td>8479</td>
<td>Leigh’s Disease (LD) with cardiomyopathy</td>
<td>173.143 – 19.24 (6)</td>
</tr>
<tr>
<td>8693</td>
<td>Lacticacidemia; Myopathy; Optic atrophy</td>
<td>117.736 ± 48.41 (8)</td>
</tr>
<tr>
<td>8768</td>
<td>Familial megalencephaly; Leukodystrophy; Basal ganglia disease</td>
<td>130.462 ± 27.54 (4)</td>
</tr>
<tr>
<td>8889</td>
<td>Fatal Infantile Lacticacidosis (FILA)</td>
<td>113.924 ± 26.74 (5)</td>
</tr>
</tbody>
</table>

** L/P ratios for control cell lines used to compare each of these measurements range from 14.705 – 2.190 (6) to 19.355 – 2.55 (8)
Part II. Genomic characterization and localization of the *NDUFA4* (MLRQ) gene and pseudogene

**Screening the PAC library**

Full-length MLRQ cDNA was used as a probe to search a P1-artificial chromosome (PAC) genomic DNA library, namely pCYPAC-1 (RPCI-1). Screening of the library was done following the protocols outlined by Osborne *et al* (1996) through the Toronto Centre for Applied Genomics.

**Southern blot analysis of PAC clones**

PAC clones 2F23 and 96E24 cultures were grown overnight at 37°C in LB media supplemented with kanamycin (50 μg/ml). Plasmid DNA was isolated according to Maniatis *et al* (1989).

Overnight restriction was performed using enzymes *EcoRI*, *MboI*, *BamHI* and *NotI* on 5 μg of plasmid DNA from both 2F23 and 96E24 clones, and using *AvaII* on plasmid DNA from clone 2F23. Restricted DNA was loaded onto 0.7% agarose gels and electrophoresed overnight. Alkaline transfer and southern hybridization was performed as per Maniatis *et al* (1989) with 200 μg/ml of Heparin as the blocking agent. 400 ng of MLRQ (full-length) cDNA probe was spotted onto the nylon membrane (Amersham). The DNA was UV cross-linked for 5 minutes. 50 ng of the MLRQ probe was labeled with P32 using the oligolabeling kit (Pharmacia). The probe was purified on a Sephadex G-50 column washed with STE (0.1 M sodium chloride, 10 mM TrisHCl, pH 8.0 and 1 mM EDTA, pH 8.0) and then mixed with salmon sperm DNA and denatured. The denatured mix was added to the membrane and shaken overnight at 65°C. The
membrane was washed twice in 2X SSC (sodium chloride and sodium citrate) and 1% SDS at room temperature for 5 mins. This was followed by two 15 minute washes in the same buffer at 65°C and then a single wash with 0.2X SSC at 65°C. The membrane was exposed to Kodax Biomax film to obtain an autoradiographic image.

**Northern analysis of MLRQ expression**

A nylon membrane blot (Human Normal Blot I, Invitrogen) with 20 μg total RNA from human tissues such as heart, brain, kidney, liver, lung, spleen and skeletal muscle was probed with 0.5 μg of the 515 bp MLRQ cDNA. 60 ng of a DNA fragment from the 3’ end of the β-actin gene was also used as a positive control probe. Probes were labeled with 32P using the nick translation system (Gibco) and purified on a QIAquick PCR purification column (Qiagen). Hybridization and subsequent washes were carried out in accordance with the manufacturer’s protocol. The autoradiograph was developed after exposing the membrane to Kodax Biomax film.

**Chromosomal localization**

Biotinylated 96E24 and 2F23 probes were prepared for FISH mapping to normal human lymphocyte chromosomes according to methods employed by Ling et al (1998) through the services of the Toronto Centre for Applied Genomics.

**Amplification, cloning and sequencing of MLRQ from genomic and PAC DNA**

Genomic DNA isolated from cardiomyocyte and fibroblast lines using alkaline lysis methods, as well as DNA isolated from PAC clone 2F23 (Maniatis et al, 1989) was used as template for PCR using primers MLRQ 1 and MLRQ 478 under the same conditions.
described above. Amplified fragments were cloned into TA\textsuperscript{II} vectors (Invitrogen) and sequenced also as described above.

**YAC library screening**

A chromosome 7-specific (CEPH) YAC library was screened using full-length MLRQ cDNA as probe, according to methods outlined by Kunz *et al* (1994) through the Toronto centre for applied genomics.

**Results and Discussion**

**Part I. Isolation and characterization of MLRQ cDNA**

**MLRQ cDNA structure**

The human MLRQ homologue was cloned from cultured fibroblast and cardiomyocyte cell lines as well as from heart, brain, liver and kidney tissues and a heart cDNA library. The full-length MLRQ cDNA sequence (accession no. AF301077) was compiled from the overlapping sequences of the two PCR fragments (Fig. 2.1). Full-length cDNA for MLRQ was found to be 515 bp in length with an open reading frame of 243 bp encoding a protein 81 amino acids long (predicted molecular mass of 9370 Da). This sequence was found to be identical to the human MLRQ cDNA clone from a Korean fetal liver tissue library (accession number U94586) (*Kim et al*, 1997), with the exception of an A to G transition at position 39 (with the start codon as pos. 1) and an additional 7 bp between the last polyadenylation site and the poly A tail. These polymorphisms can be attributed to the fact that these two clones are from the mRNA of individuals from different ethnic populations. The compiled human MLRQ cDNA sequence shows 89% (331/371) identity to its bovine counterpart (EMBL accession no. X64897) as well as
Figure 2.1. The nucleotide sequence of the human MLRQ subunit cDNA and its deduced amino acid sequence. The numbers at the right of each line denote the position of the last nucleotide and amino acid. The stop codon is indicated by an asterisk. The polyadenylation signals are underlined.
85% (193/226) identity to the *Mus musculus* MLRQ-like protein mRNA (GenBank accession no. U59509). At the deduced amino acid level, the human MLRQ subunit protein exhibits similarities of 93% (89% identity, 4% favoured substitutions) and 85% (80% identity, 5% favoured substitutions) with the bovine and mouse MLRQ subunits, respectively (Fig. 2.2).

```
hMLRQ: MLRQIGQA[KHPSLIPLFVFIGTGATGATLYLLRLALFNPDVWDR 47
bMLRQ: MLRQIGQA[KRHPSLIPLFIFIGAGGTGAALYVTRLALFNPDVSWDR 47
mMLRQ: --------QA[KKHPSLIPLFVFIGAGGTGAALYVMRLALFNPDVSWDR 40
```

**Figure 2.2.** Alignment of the predicted human MLRQ subunit protein sequence with that of bovine and mouse. Perfectly conserved residues between the human (hMLRQ), bovine (bMLRQ) and mouse (mMLRQ) sequences are highlighted.

The deduced polypeptide sequence of human MLRQ shows a relatively hydrophobic N-terminal region (amino acid positions 1-39) and a relatively hydrophilic carboxyl-terminus (amino acid positions 40-81). As with the bovine polypeptide sequence, the hydrophobic segment of about 25 amino acids (positions 15-39) is sufficiently long enough to form a transmembrane α-helix. It is interesting to note that the amino acid substitutions seen between the human and bovine/mouse MLRQ polypeptide do not affect the two domain structure of the protein, thereby suggesting that this feature is critical for the subunit’s structure and function. Based on the predicted sequence of the human MLRQ protein, it is safe to conclude that this subunit has no consensus phosphorylation sites and owing to a lack of cysteine-rich motifs (there is only one cysteine at amino acid residue 44), it is highly unlikely that the MLRQ subunit is an iron-sulfur protein. It is also interesting to note that the bulky hydrophobic amino acid
tryptophan, which has the lowest average occurrence in proteins (about 1.1%) (Voet and Voet, 1990) is seen twice (at amino acid residues 45 and 53) within the 81 amino acid MLRQ sequence.

**Mutational analysis of MLRQ cDNA in complex I deficient patients**

Since the contribution of mutations in the supernumerary subunits of complex I, such as MLRQ, to complex I deficiency had never been fully studied, mutational analysis of this subunit was undertaken. A group of six young children in whom isolated complex I deficiency had been confirmed in skin tissue (by virtue of the fact that other respiratory chain enzymes such as COX and succinate cytochrome c reductase were found normal) and in whom common mitochondrial DNA mutations had been excluded, were examined for mutations in MLRQ cDNA. However, no disease causing mutations were found in any of these patients. Recently, Triepels et al (2000) conducted an extensive study in which 19 genes of the HP fraction, including that of MLRQ, were screened for molecular aberrations in 14 young patients with isolated complex I deficiency. No disease-causing mutations or polymorphisms in the cDNA sequence of the *NDUFA4* (MLRQ) gene were detected by Triepels' group either. Therefore, mutations in the ORF of human *NDUFA4* probably represent a minimal contribution to complex I deficiency.

**Part II. Chromosomal localization and characterization of the *NDUFA4* gene and pseudogene**

**Library screening and FISH mapping**

Screening of the pCYPAC-1 human genomic DNA library with the full-length MLRQ cDNA as probe, generated two PAC clones, namely 2F23 and 96E24, with the latter
showing a much stronger hybridization signal. FISH mapping of the 2F23 and 96E24 clones to normal human lymphocyte chromosomes assigned 2F23 to chromosome 1 p21 and 96E24 to chromosome 7 p21-p22 (Fig. 2.3a, b). Another PAC clone (69E11) carrying a sequence with great similarity to the MLRQ cDNA sequence discovered through a BLAST search of the non-redundant (nr) database, was also FISH mapped to chromosome 1 q24 (Fig. 2.3b). The fact that the two clones 2F23 and 96E24 obtained from the PAC library were mapped to two different chromosomes, suggested the existence of either a pseudogene (non-transcribed) or an isoform of MLRQ.

In order to validate this possibility, primers used to amplify the cDNA sequence of MLRQ were used in a PCR with PAC 2F23 as the template. The amplified sequence (accession no. AF206638) (Fig. 2.4) revealed that there were changes in its nucleotide composition when compared to the sequence obtained from the cDNA of cardiomyocyte and fibroblast cell lines (Fig. 2.1). 6 basepair changes in the coding region; 2 basepair changes in the 5' region and 5 basepair changes in the 3' region were noted (Fig. 2.4). The nucleotide changes seen in the PAC 2F23 sequence give rise to the substitutions R3C, I5L, T27S, R35H, Y65C and R78C. If this isoform of MLRQ is expressed, the presence of the three cysteines would result in a protein that is considerably different in structure from that of the known MLRQ subunit.

Confirmation that PAC 2F23 carries a pseudogene and not an isoform of MLRQ was also done with the sequencing of the cDNA obtained from RNA isolated from heart, brain, kidney and liver tissues. The MLRQ sequences amplified from these tissues were identical to the sequences amplified from cardiomyocyte and fibroblast lines as well as a heart cDNA library. The fact that the NDUFA4 gene maps to chromosome 7 was also
Figure 2.3. Schematic of *NDUFA4* structure and chromosomal allocation of the gene and its pseudogene. (a) The idiogram of human chromosomes 7 is shown illustrating the distribution of labeled sites determined with PAC 96E24 as probe. A schematic representation of the genomic organization of the *NDUFA4* gene is also provided. (b) The idiogram of human chromosome 1 is presented illustrating the distribution of labelled sites at p21 and q23-24 determined with PACs 2F23 and 69E11 as probes, respectively.
confirmed by primers designed from the 3' region of the MLRQcDNA sequence which were used to probe a YAC genomic library. The two YAC clones that were isolated, namely C745f8 and C932h3, were both FISH mapped to the p21-p22 region of chromosome 7 (Fig. 2.3a). Sequencing the 5' region of the two PAC clones, a complete divergence in sequence between 2F23 and 96E24 was discovered ~75 bp upstream of the start codon (Fig. 2.4).

Some diseases associated with chromosome 7 mutations are cystic fibrosis, Palister-Hall syndrome, Pendred syndrome, GCK diabetes, Split hand/foot malformation type I and Williams-Beuren syndrome. However, none of these disease genes have been localized in proximity to the NDUFA4 gene loci 7 p21-22. Other genes that have been mapped to the area 7p21-22 are those encoding carniosynostosis and interleukin 6 at p21 as well as those encoding cell division protein FtsJ, G protein-coupled receptor 30, islet cell autoantigen, nucleotide diphosphate linked moiety X-type motif 1, v-maf avian musculoaponeurotic fibrosarcoma oncogene family-protein K, platelet-derived growth factor alpha polypeptide and replication protein A3 at p22. Other genes coding for complex I subunits that have been localized to chromosome 7 are NDUFB2 (DDGD) and the pseudogene possessing NDUFA5 (B13), both of which have been mapped to the q arm of the chromosome.

**MLRQ expression at the transcriptional level**

To determine the size of MLRQ transcripts as well as to shed some light on the question of an MLRQ pseudogene, northern analysis was performed on a blot containing total RNA isolated from human heart, brain, kidney, liver, lung, spleen and skeletal
Figure 2.4. Sequence of the MLRQ pseudogene on PAC 2F23. The structure of the MLRQ pseudogene on chromosome 1p21 as determined by sequencing PAC 2F23 is presented. The 5' region in bold from pos. 1 to 130 represents sequence which differs from that of NDUFA4. The single bases in bold represent the nucleotide changes in the pseudogene compared to MLRQ cDNA. The asterisks denote putative TATA and CAAT sequences while arrows point to the putative start and stop codons. The underlined sequences show the direct repeats at both ends of the coding region of MLRQ which are characteristic of Type II pseudogenes.

muscle. Probing with $^{32}$P labeled full-length MLRQ cDNA, revealed a single hybridizing band of about 500 bases in all lanes (Fig. 2.5a). The DNA fragment from the β-actin gene served as a positive control probe and indicated that RNA isolated from these tissues was not degraded (although some degradation can be observed with skeletal muscle RNA) (Fig. 2.5b, lane 7). Expression of the MLRQ transcript is seen highest in heart tissue (Fig. 2.5a, lane 1), which has the highest requirement for ATP, followed by
Figure 2.5. Northern blot analysis of *NDUFA4* transcripts in normal human tissues. A blot containing 20 μg of total RNA from normal human tissues (Invitrogen) was hybridized with $^{32}$P-labeled (a) Full length MLRQ cDNA probe and (b) DNA fragment from 3' end of β-actin gene at 42°C. Washes were performed as follows: once with 2X SSC/0.05% SDS at room temperature, three times with the same buffer at 42°C and twice with 0.1X SSC/0.1% SDS at 50°C. **Lane 1**, heart; **lane 2**, brain; **lane 3**, kidney; **lane 4**, liver; **lane 5**, lung; **lane 6**, spleen; **lane 7**, skeletal muscle. Sizes of the hybridizing fragments are indicated on the right.
skeletal muscle and brain tissues (Fig. 2.5a, lanes 7 and 2). In comparison to the amount of β-actin transcripts, MLRQ mRNA levels in liver tissue seem to be quite high. This is not surprising since liver is the metabolic clearing house of the body, degrading fatty acids and amino acids whose byproducts ultimately enter the OXPHOS system. Northern analysis performed by Kim et al (1997) agrees with this pattern of expression seen with the different tissue types. A more prudent positive control for this experiment might have been the use of a DNA probe for transcripts of a mitochondrial protein such as citrate synthase, which would have allowed comparison with protein expression as shown in the next chapter (Fig. 3.1).

Amplification of the *NDUFA4* gene from genomic DNA

Primers designed from the full-length MLRQ cDNA sequence were used to amplify the *NDUFA4* gene from genomic DNA isolated from fibroblasts and cardiomyocytes. Products of varying sizes were amplified, including one that was of similar size to the MLRQ cDNA. Sequencing of this amplified product showed that it was identical to the MLRQ pseudogene sequence on PAC 2F23. The larger sized products including one that was about 6.5 kb in size (corresponding to *NDUFA4*) (Fig. 2.6, lane 2) were not cloned or sequenced.

Southern blot analysis

Another strategy that was approached simultaneously in order to determine which PAC clone held the true genomic sequence of MLRQ was Southern hybridization (Fig. 2.7). Restriction products of the pCYPAC2 vectors 2F23 and 96E24 were probed with the MLRQ cDNA sequence. A strong hybridizing signal with an insert of
Figure 2.6. PCR amplification of *NDUFA4* from genomic DNA. Primers MLRQ 1 and MLRQ 478 were used to amplify the identified bands from genomic DNA isolated from a normal human fibroblast line. PCR conditions were as follows: 35 cycles, $T_{\text{Annealing}} = 52^\circ\text{C}$. **Lane 1**, negative control; **Lane 2**, products amplified from genomic DNA. PCR products were run on a 1.5% agarose gel. The largest product (about 6.5 kb) probably corresponds to the *NDUFA4* genomic sequence) while the smaller fragments represent other MLRQ-like sequences in the human genome. The 478 bp product was sequenced and found to be identical to the MLRQ-like sequence from PAC 2F23.
Figure 2.7. Southern analysis of PAC clones 2F23 and 96E24. Five micrograms of plasmid DNA isolated from PACs 2F23 and 96E24 were digested with the indicated restriction enzymes and electrophoresed through a 0.7% agarose gel. The DNA was transferred onto a Hybond-N+ (nylon) membrane and hybridized with a $^{32}$P-labeled MLRQ cDNA probe. The blot was washed twice in 2X SSC/1% SDS at room temperature, followed by two 15 minute washes in the same buffer at 65°C and a single wash with 0.2X SSC at 65°C. Lane 1, $\lambda$ HindIII marker; lanes 2 and 8, EcoRI; lanes 3 and 9, MboI; lane 4, AavII; lanes 5 and 10, BamHI; lanes 6 and 11, NotI; lane 7, 50 bp marker. Lanes 2-6 contain digested DNA from PAC 2F23 and lanes 8-11 contain digested DNA from PAC 96E24. Arrows point to hybridizing fragments, however, band sizes are only provided for the strongly hybridizing signals.
approximately 4 kb was seen in NotI restricted 2F23 (lane 6), while no signal was detected in the case of 96E24 restricted by the same enzyme (lane 11). This is surprising especially because NotI is the enzyme that is supposed to release the genomic insert from the pCYPAC2 vector. 2F23 cut with AvaII gives two fragments approximately 1.5 kb and 2 kb in size (lane 4) and 2F23 restricted with BamHI gives a strong signal at 4kb (lane 5). The only signals detected for PAC 96E24 were in lanes with EcoRI restricted 96E24 (lane 8), which had a 1 kb fragment and MboI restricted 96E24 (lane 9) where a very large fragment was detected. Cloning and sequencing of these hybridizing DNA fragments would have provided answers as to the sequence of genomic MLRQ. However, since the 96E24 PAC clone was mapped to chromosome 7, this sequence was first used in a High Throughput Genomic Sequence (HTGS) database search for a genomic clone. Fortuitously, search results revealed that a clone DJ0855F16 (accession no. AC007029) contained the NDUFA4 genomic sequence. Therefore, characterization of the hybridizing fragments from PACs 2F23 and 96E24 was not pursued.

Genomic organization of NDUFA4

The NDUFA4 gene is composed of 4 exons that range from 58 bp (exon III) to 237 bp (exon IV). The three introns are estimated to range from 0.66 kb (intron 2) to nearly 4.4 kb (intron 3) in size with the entire NDUFA4 gene spanning approximately 6.7 kb (Fig. 2.3a). The splice donor and acceptor sites in each of the four introns follow the consensus GT/AG splicing sequence (Mount, 1982) (Table 2.3). Analysis of the exon/intron structure of NDUFA4 indicates that the predicted transmembrane domain is encoded primarily by exon II, while exons III and IV carry the hydrophilic domain.
Genomic and cDNA sequences of MLRQ reveal that it has three in-frame UAG stop codons upstream of the start site. As with the bovine cDNA sequence, this suggests that the import sequences for MLRQ lie within the mature protein and not at the N-terminus (Walker et al, 1992). Many of the subunits in the hydrophobic fragment possess this feature and are brought into the mitochondrial IM by binding to the soluble Tim9 and Tim10 proteins and imported through the Tim22/Tim54 membrane protein complex via Tim12 (see Fig. 1.4). In the PAC 2F23 sequence however, while the UAG in-frame stop codon immediately preceding the start codon is preserved, two additional UGA in-frame stop codons are also seen at different upstream sites. Three in-frame AUG codons are also present in the 5' UTR of the PAC 2F23 sequence. Two of these are between the two UGA stop codons, while one is found further upstream (Fig. 2.4). Interestingly enough, the 5' UTR of the presumed pseudogene on PAC 2F23 does show the classical TATA and

Table 2.3. Exon-intron splice junctions of the human NDUFA4 gene

<table>
<thead>
<tr>
<th>Exon</th>
<th>Size (bp)</th>
<th>5' Splice donor</th>
<th>Intron Size (bp)</th>
<th>3' Splice donor</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>&gt;95</td>
<td>CCGAGC gtaagt</td>
<td>1</td>
<td>1119 gtttag TTGATC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ser 14</td>
<td></td>
<td>Leu 15</td>
</tr>
<tr>
<td>II</td>
<td>89</td>
<td>TGTTTG gtaagt</td>
<td>2</td>
<td>660 tttag TTGGGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cys 44</td>
<td></td>
<td>Cys 44</td>
</tr>
<tr>
<td>III</td>
<td>58</td>
<td>TACAAG gtaaac</td>
<td>3</td>
<td>4397 tccaag TTCTAC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lys 63</td>
<td></td>
<td>Phe 64</td>
</tr>
</tbody>
</table>

IV 237

Exon and intron sequences are shown in upper-case and lower-case letters, respectively. The numbering of the amino acids corresponds to that of the mature MLRQ protein while the nucleotide sequence coding for these amino acids is italicized within the exon sequence.
CCAAT boxes, but at positions -96 to -99 (107 to 110 on Fig. 2.4) and -21 to -25 (181 to 185 on Fig. 2.4), respectively, whereby they are rendered non-functional. Although AT-rich boxes are found in the 5' regulatory region, no such promoter motifs are seen in the 130 bp sequence preceding the start codon of PAC 96E24 (Fig. 2.8).

Although AT-rich boxes are found in the 5' regulatory region, no such promoter motifs are seen in the 130 bp sequence preceding the start codon of PAC 96E24 (Fig. 2.8).

The absence of TATA and CAAT boxes is not however exclusive to the NDUFA4 gene. The promoter regions of genes NDUFA1 and NDUFV2 do not contain these classical promoter motifs either (Zhuchenko et al, 1996; de Coo et al, 1995; Hattori et al, 1995).
The translational initiation codon for the *NDUFA4* gene occurs in the context GCAAACatgC, which deviates a little from the optimal GCCACCatgG as defined by Kozak (1996). However, one of the two positions said to exert the strongest effect for translation within this consensus, an A at position −3, is preserved. The 5' end of *NDUFA4* is embedded in a CpG-island which is particularly apparent extending from nt -597 to -338 (G+C content 76%). The G-C rich region could well extend into the first intron to about nt +383 with the G+C content becoming 58%. This is a trend particularly noticeable on housekeeping genes and other nuclear respiratory genes such as *NDUFV2*, *NDUFA1* and *NDUFA5* (Tensing *et al.*, 1999). Nucleotide sequence analysis of this region has revealed potential binding sites for transcription factors such as NRF2 (nuclear respiratory factor 2) at positions -245 to -242; −96 to −93 (GGAA), NF1 at positions -357 to −354 (TGGCA), AP1 at positions −586 to −583 and −333 to −330 (TGAC) as well as a binding site for the ubiquitous transcription factor Sp1 (-552 to −549). However, search for the promoter region using an algorithm called PromoterInspector ([http://genomatix.gsf.de/cgibin/promoterinspector/promoterinspector.pl](http://genomatix.gsf.de/cgibin/promoterinspector/promoterinspector.pl)) did not identify this region as a putative eukaryotic polymerase II promoter. Instead, this algorithm which is capable of identifying highly specific localizations of promoter regions in large genomic sequences (Scherf *et al.*, 2000), identified a 192 bp region between nt positions 19,881 and 20,072 of the 95,345 bp clone RP5-855F16 (accession number AC007029) which carries the genomic sequence of *NDUFA4* as the only putative promoter region in the clone. Binding sites for transcription factors IK1, GC, SP1 and AP2 as well as two sites for MZF1 and IK2 have been identified in this 192 bp region. If
this is indeed the promoter modulating transcription of the *NDUFA4* gene, it is acting from a distance of approximately 34 kb.

**The pseudogene on chromosome 1**

While the human somatic cytochrome c gene (HCS) is represented by a single expressed gene and 11 processed pseudogenes (Evans *et al.*, 1988), pseudogenes have thus far only been found associated with complex I subunits such as the 24 kDa subunit (de Coo *et al.*, 1995), the B13 subunit (Pata *et al.*, 1997), the 39 kDa subunit ([www-dsv.cea.fr/thema/MitoPick/ImagesCxHumain/Xsome_Map_Human.html](www-dsv.cea.fr/thema/MitoPick/ImagesCxHumain/Xsome_Map_Human.html)) and possibly the DDGD (13 kDa) subunit (Loeffen *et al.*, 1998a). It is now widely accepted that pseudogenes can arise from two main mechanisms (Mighell *et al.*, 2000). Type I pseudogenes arise through gene duplication of a functional, parental gene and may remain transcriptionally active (Dover, 1989; Wilde, 1986). Retrotransposition on the other hand, gives rise to type II pseudogenes as a result of reverse transcription of the processed parental transcript (Dover, 1989; Wilde, 1986). Type II pseudogenes are not closely linked to their parental genes, but instead are found on different chromosomes (Mighell *et al.*, 2000). This is because the cDNA is reinserted into the genome at arbitrary breaks in a chromosome and as a result, the pseudogene is normally transcriptionally silent (Mighell *et al.*, 2000). The fact that the two PAC sequences were localized to two different chromosomes suggests that in the case of the *NDUFA4* gene, retrotransposition may have been the cause of pseudogene generation. The fact that none of the MLRQ cDNA sequences determined from tissues such as heart, brain, liver and kidney corresponded to the pseudogene sequence on PAC 2F23, further suggests that this MLRQ pseudogene may be transcriptionally silent. Processed pseudogenes are also known to
have a lack of introns, precise boundaries coinciding with the transcribed regions of the
gene and short direct repeats at both ends (Mighell et al., 2000). These features are
characteristic of the pseudogene in question. The pseudogene on PAC 2F23 is intronless,
is very similar to the transcribed gene except for the described basepair substitutions and
has two pentanucleotide repeats (Fig. 2.4). The first direct repeat TGCAT starts at pos.
103 in the 5' region and at nucleotide pos. 550 in the 3' region. The second direct repeat
at ATGTT starts at nucleotide pos. 110 in the 5' UTR and at pos. 561 in the 3' region. In
fact, the entire nucleotide sequence starting from pos. 94 through to pos. 114
(TTCTTTAAGTGCATATATGTT) in the 5' UTR shows a great degree of similarity to
the nucleotide sequence starting from pos. 541 through to pos. 565
(TTCTCTAAATGCATGAAATCATGTT) in the 3' UTR. The origin of the MLRQ
pseudogene should become apparent after a more thorough analysis of the upstream and
downstream flanking DNA sequences on PAC 2F23. This may in fact be possible with
recent results from the search of the HTGS database with the sequence from PAC 2F23.
An almost identical sequence (99%) was found in two separate clones, Homo sapiens
close RP11-24J14 (accession number AC016085) and Homo sapiens chromosome 1
close RP11-270C12 (accession number AL160002). These search results confirm the
fact that this pseudogene is indeed an intronless sequence.

Other MLRQ-like sequences in the genome

Also important to note is that other MLRQ-like sequences (in addition to the
previously mentioned one, mapped to the X-chromosome by researchers at the
Whitehead Institute) have been reported. One sequence from PAC 69E11 (accession
number AL021397) (Fig. 2.9) mapped to chromosome 1 q24 (Fig. 2.3b), is 88% identical
Figure 2.9. Comparison of the MLRQ cDNA sequence with the sequences from PACs 2F23 and 69E11 in the region showing greatest similarity. Identical DNA sequences are shaded. The start and stop codons of the cDNA open reading frame are shown in bold. The polyadenylation signals are underlined. The bold base 92 (asterisk) represents an adenine/guanine polymorphism within the human NDUFA4 coding sequence. The last 7 basepairs of the 3' UTR (with asterisks) denote another polymorphic site. Although numbered consistently for convenience, the sequence from PAC 69E11 is actually reverse complementary to the MLRQ cDNA and PAC 2F23 sequences.
to the coding and 3' region of MLRQ cDNA. However, this sequence is intronless and prematurely truncated and reverse complementary to the actual MLRQ cDNA sequence. Another sequence from clone 1189B24, mapped to chromosome Xq25-26.3 (accession number AL030996) has been reported to show 84% identity to the coding and 3' region of the MLRQ cDNA sequence. This intronless sequence too gives rise to a premature stop codon resulting in a protein with only 18 amino acids. Recently, another sequence from PAC clone 219d7 (accession number AF225899) with 88% identity to the MLRQ cDNA sequence, but truncated after the first three amino acids, as well as a sequence mapped to chromosome 12p11-37.2-54.4 (accession number AC012156) showing 86% identity, but without the start codon at the expected position have also been reported. In comparison to the aforementioned MLRQ-sequences however, clone 2F23 shows a greater similarity (97%) to the NDUFA4 cDNA sequence (Fig. 2.9). Even though the NDUFA4 gene does not code for one of the 14 core complex I subunits found in bacteria (Weidner et al, 1993), the existence of many MLRQ-like sequences seems to suggest that this subunit may have appeared quite early on in the evolutionary course of complex I, in order to have undergone such a high degree of changes.
Abstract

Expression of the MLRQ protein in human tissues such as heart, brain, liver, kidney, placenta and muscle showed it to be of uniform size (9 kDa) with no evidence of post-translational modifications that significantly affect its molecular mass. Overexpression of this subunit in a patient with hepatomegaly with renal tubulopathy (HT) who had decreased amounts of other complex I subunits seems to suggest a differential regulation of MLRQ from that of other subunits in the enzyme. Bacterial expression and purification of MLRQ as a GST-fusion protein was successful. Although transfection of the MLRQ gene in the anti-sense orientation into transformed fibroblasts was successful, downregulation of the subunit was not observed by immunodetection with the MLRQ antibody. Solubilization studies demonstrated that MLRQ extracted differently from the other complex I subunits examined, leading to the proposition that the region proximal to its N-terminus is buried very deeply in the membrane arm. The C-terminus being quite hydrophilic is the exposed region that interacts with other subunits as evidenced by immunoprecipitation studies with the MLRQ antibody directed towards epitopes at its C-terminus. These studies also implicate MWFE as well as the 49 kDa and B18 subunits to a lesser extent, as being subunits that are in close proximity to MLRQ. Cross-linking of bovine heart mitochondria with reagents DST and EGS support these findings and suggest the association of MLRQ with MWFE through an as yet unidentified subunit.
Introduction

The biochemical characterization of MLRQ was necessitated by the fact that limited information was available on its structure, spatial organization or function. MLRQ was assigned as a subunit of complex I from studies by Walker et al (1992) who found that it was a peptide that eluted at 44% acetonitrile on a HPLC column after being isolated in a 2-step chromatographic procedure from complex I that had been dissociated in guanidinium-hydrochloride. Even though Walker et al (1992) first identified it as a 9 kDa subunit in B. taurus, the position of this subunit which was isolated from urea-solubilized complex I was not known with certainty. To date, the MLRQ subunit has never been visualized on acrylamide or polyacrylamide gels from studies on the resolution of complex I subunits (Walker et al, 1992; Finel et al, 1992; Fearnley et al, 1994; Sazanov et al, 2000). This inadequacy may be attributed to the fact that there are other complex I subunits that migrate in the 8 to 10 kDa MW range such as the 10 kDa (FP) subunit as well as the B8, B9, SDAP and AGGG subunits of the HP fraction. Regardless, the first question that needed to be addressed was whether the MLRQ subunit is indeed an integral part of complex I, or whether it is simply a polypeptide that co-precipitates with complex I during extraction studies.

The molecular mass of MLRQ in B. taurus was determined from electrospray ionization mass spectrometry (e.s.i.-m.s) by Fearnley et al (1994) to be 9324 (+/- 1.0) Da, agreeing very well with the expected molecular mass of 9324.7 Da as calculated from the sequence. However, no groups have thus far examined the expression of this subunit through immunodetection, thereby creating a need for such studies from which valuable
information on the expression of MLRQ in various human tissues as well as complex I deficient patients can be obtained.

MLRQ has also been detected by N-terminal sequencing or electrospray mass spectrometry, as a subunit in subcomplex Iα, which is an active NADH dehydrogenase and represents the predominantly hydrophilic (globular) domain of intact complex I (Finel et al, 1992). The fact that subcomplex Iλ, which is more water-soluble than Iα was found not to contain MLRQ (Finel et al, 1992) is also in agreement with structural information from the MLRQ amino acid sequence which predicts it to have one hydrophobic segment with the potential to be folded into a membrane-spanning α-helix. Given this information, a few other questions regarding MLRQ such as the structure of the subunit, its location in complex I and its interaction with other subunits of the enzyme also arise.

Answers to these questions are essential in order to determine not only the role of the MLRQ subunit but ultimately to augment existing knowledge on the structural organization and function of human complex I which is made up of many such supernumerary subunits.

**Materials and methods**

Part I. MLRQ expression in human tissues and cells

**Antibody Generation**

High titer polyclonal MLRQ antibody was developed against a 14 amino acid sequence at the C-terminus (amino acid residues 69-82: NVDYSKLKKEGPDF) of bovine MLRQ. The human MLRQ protein sequence is identical in that region except for
arginine (R) in place of glycine (G) at amino acid position 79. The KLH-peptide was emulsified by mixing with an equal volume of Freund's Adjuvant and injected into three to four subcutaneous dorsal sites in 3-9 month old New Zealand white rabbits, for a total volume of 1.0 mL (0.1 mg of peptide) per immunization (Research Genetics Inc.). The animals were bled from the auricular artery. The blood was allowed to clot and serum was collected by centrifugation. The anti-peptide antibody titer was determined with enzyme linked immunosorbent assay (ELISA) with free peptide on the solid phase. The antibody titer at third bleeding was determined to be 51,700.

**Western blot analysis of MLRQ expression**

25 μg of mitochondrial protein as quantified by the Lowry assay (Lowry *et al.* 1951) and prepared from human tissues such as brain, heart, liver, kidney, placenta and muscle (Pitkanen *et al.*, 1996) were electrophoresed on 16% SDS polyacrylamide gels. The protein was transferred onto a nitrocellulose membrane (Xymotech) and subjected to Western blot analysis. Primary antibodies targeted to the C-terminus of complex I-MLRQ and the citrate synthase enzyme (Research Genetics) were used to probe the blot. The MLRQ antibody was used to probe the blot after the citrate synthase antibody was stripped using a solution with a final concentration of 6M urea and 50mM Tris-HCl (pH 7.4). α rabbit IgG horseradish peroxidase was used as the secondary antibody. Development of the autoradiograph was carried out using the ECL kit (Amersham-Pharmacia) according to manufacturer's instructions.

Expression of MLRQ in 50 μg of mitochondrial protein from patient cell line 5624 and control cell line 4212 was also examined using antibodies against the MLRQ, ASHI, B18 and 49 kDa subunits. The blots were developed using goat anti-rabbit IgG alkaline
phosphatase conjugated secondary antibody and NBT/BCIP (Biorad) (McEachern et al. 2000).

Part II. Bacterial expression of MLRQ protein

**Design of MLRQ-fusion protein construct**

Primers GST3 5' CAGTCGGGGATCCCTGCC 3' and GST 304 5' TGAATTCGAAATCTGGACGTTC 3' (flanking the start and stop codons of MLRQ) were used to amplify the MLRQ sequence from a human heart cDNA library. * indicates the bases that were mutagenized in order to engineer the restriction sites EcoRI and BamHI into the amplified sequence. The engineered MLRQ cDNA fragment was subcloned into the TA" cloning vector (Invitrogene) and then sequenced following the protocols described in chapter 2. The MLRQ insert of interest was obtained by restricting the recombinant vector using enzymes BamHI and EcoRI. The pGEX-3X vector (Pharmacia) was also restricted with the same enzymes and dephosphorylated according to the manufacturer’s protocol. T4 DNA ligase (Pharmacia) was used to ligate the MLRQ insert into pGEX-3X. Competent DHα cells were transformed with the pGEX-3X-MLRQ construct and isolated clones were sequenced to ensure that the Factor Xa cleavage site (“Ile Glu Gly Arg”) in the fusion construct had been preserved. BL21 cells made competent using RbCl (Maniatis et al, 1989) were then transformed with the sequenced pGEX-3X-MLRQ construct.
Induction and purification of MLRQ-GST fusion protein

After successful small-scale expression of the fusion protein, the protocol was optimized to get a greater degree of solubility for the fusion protein during large scale expressions. 100 mL LB media supplemented with ampicillin was inoculated with BL2I cells carrying the pGEX-3X-MLRQ construct and induced with 0.1 mM IPTG overnight at 30°C when the OD₆₀₀ reached between 0.6 to 1.0. Pilot experiments examined expression of the fusion protein at 0, 1, 2, 3, 4 hrs and overnight after induction. Cells were sedimented by centrifugation at 7,700g in an SS-34 rotor (Sorvall). The ensuing pellet was resuspended in 5ml of Lysis buffer (20 mM Tris HCl, pH 8.0; 150 mM NaCl; 1 mM EDTA; 1 mM DTT; 10% glycerol, 0.1 mg/ml Lysozyme and 10 μg/ml leupeptin). Cells were lysed by freeze/thawing three times. 0.15g urea was added to the cells over 30 minutes at room temperature to give a final concentration of 0.5M. This was followed by the addition of Triton x-100 over 1 hour to give a final concentration of 1%. Cells were spun at 12,000g for 30 minutes in an SS-34 rotor and the “soluble” supernatant was applied to a prepared 1 ml Glutathione Sepharose 4B column (Pharmacia). The column was washed with 10 bed volumes of 1X PBS. Fusion protein bound to the column was eluted in 1ml batches by incubating the column in Glutathione elution buffer (10 mM reduced glutathione in 50 mM Tris-HCl, pH 8.0) for 10 minutes at room temperature. The pooled eluates were then analysed by SDS-PAGE.

Factor Xa cleavage of fusion protein

The pooled eluate from the Glutathione Sepharose 4B column was injected into a dialysis slide (Pierce) and dialysed overnight in 50 mM Tris HCl, pH 7.5 and 150 mM
NaCl buffer. Cleavage was performed in a Factor Xa to fusion protein ratio of 1:10 (w/w) with the addition of 10 μg Factor Xa in a final concentration of 1mM CaCl₂ to the dialysed solution and left overnight at room temperature. After digestion, MLRQ protein was retrieved in the flow-through by column purification on Glutathione Sepharose 4B as described before and visualized on a SDS-PAGE gel by Coomassie staining.

Part III. Anti-sense expression of MLRQ in mammalian cells

**Design of sense and anti-sense oriented pREP9 constructs**

MLRQ was cloned into pREP9 vectors (a mammalian episomal vector conferring neomycin resistance) in both the sense and anti-sense orientation. Primers Hind4 5’

```
** * 
TGGTGGCTAGGAAGCTTCTCT 3’ and Xho339 5’ ACCTTCATTCTCGAGCAGCGT
```

3’ were used to amplify a 311 bp MLRQ fragment from human fibroblast cDNA through PCR. * indicates the bases that were mutagenized to engineer the HindIII and XhoI sites. The resulting fragment encompassed 40 bp of the 5’ region, the start and stop codons and 26 bp of the 3’ region and was cloned in the sense orientation using the engineered restriction sites. The anti-sense construct contained a 284 bp fragment that was also

```
** 
amplified from cDNA using primers Bam30 5’ AGTCGGATCCCTCTGCG 3’ and
```

```
** 
Hind339 5’ ACCTTGATTCTAAACTTCGT 3’. * indicates the bases that mutagenized to engineer the BamHI and HindIII sites. The resulting fragment was cloned into the same vector in an anti-sense orientation to the CMV promoter using these restriction sites. The anti-sense fragment covered 24 bp of the 5’ region, the start and stop codons and 26 bp of the 3’ region. These fragments were first subcloned into TA II vectors.
(Invitrogen) and then sequenced using the T7 DNA sequencing kit (Amersham Pharmacia) before being cloned into the pREP9 vectors.

Optimization of transfection conditions

A vector with a colourometric tag, pCMVβ was used to determine the DNA versus ExGen 500 (MBI Fermentas) ratio required for optimal transfections. Transfections were carried out with DNA:ExGen 500 ratios of 2μg:4μl; 2μg:8μl; 2μg:10μl; 2μg:12μl; 4μg:4μl; 4μg:8μl; 4μg:10μl; 4μg:12μl; 4μg:16μl. The desired volume of ExGen 500 diluted with 150 mM NaCl was added dropwise to an equal volume of the desired quantity of pCMVβ vector, also diluted with 150 mM NaCl. The mixture was incubated for 10 minutes at room temperature. Fibroblasts transformed with SV40 were collected by centrifugation at 6000 rpm and washed twice in serum-free α-MEM + pyruvate + uridine media. Cells were resuspended in 0.5 mL of the same media and the ExGen 500 + plasmid mix was added to the cells. After a 5 minute spin at 1500 rpm to increase transfection efficiency, cells were plated in 6-well plates and transfection was allowed to proceed for 4-5 hours. After washing twice with PBS, 2 mL of the same media with serum was added and culture growth was monitored for two days. Cells were fixed with 2% formaldehyde/PBS and 1 mg/mL X-gal (Gibco BRL) was added. Cells were observed under the microscope for blue staining as an indicator of transfection efficiency.

Transfection and selection

Both lymphoblasts and transformed fibroblasts were transfected with sense and anti-sense pREP9 constructs following the optimization protocol outlined above. Selection on these cell lines transfected with the sense and anti-sense pREP9 constructs was begun
three days after transfection by treating the cells with increasing concentrations of the neomycin analog G-418 sulfate (Gibco BRL), starting with 100 µg/mL.

**Transfection and selection with the linearized vector pCDNA 3.1+**

Sense and anti-sense fragments were amplified as described above, cloned into pCDNA 3.1+ vector and subsequently linearized using *Pvu*I. 10 µl Superfect (QIAGEN) was added to 2 µg of the DNA (sense, anti-sense or control pCDNA 3.1 +) in 100 µl of serum-free α-MEM + pyruvate + uridine media and left for 5-10 minutes at room temperature. SV40 transformed cells in 6-well plates were washed twice with PBS. 600 µl of growth media (α-MEM + pyruvate + uridine with sera) was added to the Superfect/DNA mixture and applied to the cells. 2 mL of the same media was added 6 hours after transfection. Two days after transfection, selection was started on the transfected cells with 100 µg/mL G-418 sulfate (Neomycin analog). Three weeks after the start of selection, viable clones were isolated using cloning rings and grown on individual 100 x 20 mm plates. Selection was gradually increased upwards to 800 µg/mL of G418. Selection on untransfected SV40 transformed fibroblast cells with varying concentrations of G-418 was also concurrently performed.

**Western blot analysis of sense and anti-sense expression in transfected cells**

Mitochondria was isolated from pCDNA 3.1 + (control) as well as sense and anti-sense construct transfected SV40 cells lines (Pitkanen et al, 1996). 100 µg of mitochondrial protein as assayed by the Lowry method (Lowry et al, 1951), from various clones transfected with sense and anti-sense constructs as well as the control fibroblast line immortalized with SV40 were run on a 16% SDS-PAGE gel and western blotting
was performed as described by McEachern et al (2000). Expression levels of MLRQ and the 49 kDa subunit was detected in these clones using the respective antibodies.

**Amplification of MLRQ from transfected cells**

Primers PCDNA-R 5' AGAAGGCACAGTCGAGGC 3' (designed from within the pCDNA 3.1 + sequence) and either PCSense 5' GCCGCAAACATGCTCCGC 3' or PCAnti 5' GAAATCTGGACGTT CCTTCTT 3' were used to amplify the MLRQ sequence from sense and anti-sense pCDNA 3.1 + transfected fibroblast clones. PCR was carried out at an annealing temperature of 50°C for 35 cycles.

**Part IV. Association of MLRQ with other complex I subunits**

**Solubilization of beef heart mitochondria**

Beef heart mitochondria were isolated as per Pitkanen et al, 1996 and initial studies on solubilization involved extracting 1 mg/ml of mitochondria for 1 hour at 4°C and centrifuging the sample at 10,000 rpm for 30 minutes in an Eppendorf centrifuge. Detergents used included 0.2%, 0.5%, 1%, 2% dodecyl maltoside (DDM), 0.5%, 1%, 2% Triton X-100, 1% octylglucoside, 1% CHAPS, 1% sodium deoxycholate, 1% Nonidet P40 with 0.5% sodium deoxycholate, 1% SDS, 1% DDM with 1.5% sodium deoxycholate.

Further extraction studies were performed by resuspending beef heart mitochondria in phosphate buffered saline (PBS) with 0.2% azide to a final concentration of 5 mg/mL in the presence of 1 mM aprotinin, 2 mM leupeptin and 1 mM benzamide protease inhibitors. Either 1% DDM or 2% Triton-X 100 or 2% Triton-X 114 were added to the mitochondria and the sample was shaken for 1 hour at room temperature. The sample was spun at
16,500 rpm using the ultracentrifuge rotor SW41 for 45 minutes. The pellet was resuspended in 0.2% azide PBS. Two further extractions were performed on the resulting supernatants following the procedure outlined above. Aliquots were taken from mitochondria before DDM addition and also from the pellets and supernatants resulting from the three extractions, for western blot analysis. The supernatants were concentrated using Centricon tubes (Amicon) and the Lowry assay was used to quantify protein (Lowry et al, 1951).

Immunoprecipitation of MLRQ, MWFE and 49 kDa subunits

1 ml of the concentrated supernatant (1 mg/ml) from the first extraction with 1% DDM was treated with 10 μl of pre-bleed (non-immune) rabbit serum for 1 hour at 4°C on a rocking platform. 50 μl of the homogeneous protein A-agarose suspension was added to the sample and incubated for at least 3 hours at 4°C on a rocking platform. The beads were pelleted by centrifugation for 20 s at 12,000 g in a microfuge. Either 20 μl of MLRQ antibody or 30 μl of 49 kDa antibody or 40 μl of MWFE antibody were added to the supernatant and incubated for 1 hour at 4°C on a rocking platform. 50 μl of the homogeneous protein A-agarose suspension was added to the mixture and incubated again at 4°C on a rocking platform for at least 3 hours. The complexes were collected by centrifugation for 20 s at 12,000 g in a microfuge. The pellet was washed twice in 1ml of washing buffer 1 (50 mM Tris-HCl, pH 7.5, 150 mM sodium chloride (NaCl), 1% Nonidet P40 and 0.5% sodium deoxycholate), twice in washing buffer 2 (50 mM Tris-HCl, pH 7.5, 500 mM NaCl, 0.1% Nonidet P40 and 0.05% sodium deoxycholate) and lastly in washing buffer 3 (50 mM Tris-HCl, pH 7.5, 0.1% Nonidet P40 and 0.05% sodium deoxycholate). The washes were all performed at 4°C for 20 minutes on a
rocking platform and the pellets were collected by centrifugation as described above. After the last traces of the final wash were removed from the agarose pellet, it was resuspended in 40 μl of SDS gel loading buffer and heated to 100°C for 3 minutes before being centrifuged for 20 s at 12,000 g in an Eppendorf centrifuge. The immunoprecipitate was then analysed on a western blot with ECL (Amersham-Pharmacia) development of the autoradiograph according to manufacturer’s instructions. The identity of the protein bands that appear during immunoprecipitation was confirmed through MALDIToF sequencing (Keough et al, 2000) of these fragments excised from the SDS-polyacrylamide gels.

Cross-linking with DST and EGS

4.0 mg/mL of purified bovine heart mitochondria were dialyzed against 50 mM Triethanolamine pH 8.0, containing 0.25 M sucrose at 4°C for 6 hours. 0.2 mM EGS (Ethylene glycolbis(succinimidylsuccinate)) and 1 mM DST (Disuccinimidyl tartrate) were used to cross-link samples of these mitochondria for 1 hour as described by Yamaguchi and Hatefi (1993). However, Triton-X 100 was not added to the cross-linked samples and the reaction was quenched by the addition of 5 mM glycine for 30 minutes. Subsequent experiments used 0.2M or 0.5M of the cross-linker EGS and the cross-linking reactions were carried out for 2 hours at which point they were quenched with 20 mM glycine for 30 minutes.

Immunoprecipitation of cross-linked bovine heart mitochondria

The mitochondria cross-linked with 0.5 M EGS were solubilized with 1% DDM as described in previous sections and the resulting supernatant diluted in 1 mL PBS-azide (1
mg/mL) was used in immunoprecipitation reactions using both MLRQ and MWFE antibodies, also as outlined in previous sections.

**SDS-PAGE and western blot analysis of MLRQ during extraction, immunoprecipitation and cross-linking**

For visualization of protein, cross-linked samples as well as cross-linked and immunoprecipitated samples were run on 5-20% Tris-glycine gradient gels (Novex) and stained with Coomassie blue. Quantitation of protein samples was performed by the Lowry assay (Lowry et al, 1951). For western blot analysis, samples were electrophoresed similarly but transferred onto nitrocellulose membranes (Xymotech) and developed using ECL (Amersham-Pharmacia) reagents as mentioned above.

**Results and Discussion**

**Part I. Determining MLRQ form, function and expression**

**Tissue expression of MLRQ**

Western blot analysis of mitochondria isolated from various human tissues (Fig. 3.1) revealed that there is little variation in the apparent molecular mass of the MLRQ subunit. The citrate synthase enzyme which is part of the TCA cycle was used as a positive control and is reflective of the number of mitochondria in the tissue. Expression of MLRQ was seen to be highest in heart (lane 2), an organ where cellular energy production is most essential. The lowest expression of MLRQ was seen in mitochondria from placenta (lane 5). Also, when compared to the expression of citrate synthase, greater expression of MLRQ is seen in liver and kidney tissues (lanes 3 & 4). The high level of MLRQ expression in liver is supported by results from Northern analysis (Fig.)
Figure 3.1. Tissue specific expression of the MLRQ subunit of complex I. 25 μg of total mitochondrial protein was separated by SDS-PAGE using a 16% polyacrylamide gel. The separated proteins were subjected to Western blot analysis. The blot was developed using horseradish peroxidase linked to anti rabbit IgG. Lane 1, brain; lane 2, heart; lane 3, liver; lane 4, kidney; lane 5, placenta; lane 6, muscle.
which also showed increased amounts of \textit{NDUFA4} transcripts in this tissue. However, the magnitude of MLRQ expression in tissues such as brain and skeletal muscle (lanes 1 and 6) seems to be lower than that indicated by Northern analysis of \textit{NDUFA4} transcripts in the same tissues. The level of MLRQ expression was also seen to parallel that of the 49 kDa iron-sulfur subunit in other western blots. Detection of the MLRQ subunit with the expected molecular mass of 9 kDa suggests two possibilities. One possibility is that this subunit undergoes no post-translational modifications that significantly influence its molecular mass. The other is that the MLRQ subunit in all tissues is post-translationally modified.

Although expression of the MLRQ subunit was examined in many cases of complex I deficiency, noteworthy information about this subunit was obtained from the examination of only one patient (5624), diagnosed with hepatomegaly with renal tubulopathy (HT). No mutations were identified in the cDNA of this patient (see mutational screening in previous chapter) and the size of the MLRQ protein as detected by western blotting was at the expected MW. However, the magnitude of MLRQ expression in 5624 (lane 1) seemed to be two-fold that of the control 4212 fibroblast cell line (lane 2) (Fig. 3.2d). This is of significance, especially considering the fact that amounts of other complex I subunits such as 49 kDa, ASHI and B18 were underexpressed in this complex I deficient cell line when compared to the control (Fig. 3.2a,b,c). These immunoblotting experiments as well as the fact that no mutations were detected in any of the subunits screened (49 kDa, ASHI, B18, MLRQ, MWFE, PGIV, AQDQ, TYKY and PSST), suggest a problem in the assembly of complex I. In light of these findings, it seems likely that the expression of the MLRQ subunit is regulated by a different mechanism than that
Figure. 3.2. Western blot analysis on mitochondria isolated from cultured skin fibroblasts of a patient (5624 - HT) and control (4212) using various complex I antibodies. Mitochondria (100 μg of protein) from cultured skin fibroblast cultures of (1) 5624 - a patient with Hepatomegaly with renal tubulopathy (HT) and (2) 4212 - a control subject, were run on a 16% polyacrylamide gel, electrobotted onto a nitrocellulose matrix and immuno-detected with antibodies to the following complex I subunits (a) 49 kDa (b) ASHI (c) B18 and (d) MLRQ. Immunoreactive proteins were visualized with alkaline phosphatase conjugated anti-rabbit IgG.
responsible for other complex I subunits. Alternatively, it is also a possibility that MLRQ is one of the first subunits to be incorporated during the assembly of complex I and therefore not affected by the titre of the other subunits that follow.

**Bacterial expression of MLRQ: Attempts at defining subunit structure**

In order to determine the secondary structure of the MLRQ protein through CD spectroscopy and NMR, bacterial expression of the MLRQ protein was undertaken. Initial efforts to express the MLRQ protein in bacteria included using pProEX Hta (Gibco BRL) as well as pET-21d (+) (Novagen) systems, both using a Histidine tag for purification. Attempts to induce the protein with IPTG however proved unsuccessful with both systems. Rationalizing that this was probably because proteins less than 10 kDa in size like MLRQ are generally difficult to express stably in _E. coli_ because they cannot fold correctly and are often subject to proteolytic degradation, expression of MLRQ as a Glutathione-S-transferase (GST) fusion protein was attempted.

Preliminary trials revealed that bacterial expression of the MLRQ-GST fusion protein became optimal at a temperature of 30°C and that greater expression of the protein occurred during overnight growth at that temperature. Expression was seen to increase proportionally from 1 to 4 hours of induction and level off at overnight induction (Fig. 3.3). Purification of the MLRQ-GST protein on GST Sepharose columns was also successful, yielding pure samples of the fusion protein as determined on SDS-polyacrylamide gels (Fig. 3.4a,b,c). As seen in Fig. 3.4a (lanes 2, 3 and 4), a considerable amount of the fusion protein remains insoluble. Inducing the BL21 cells carrying the pGEX-3X-MLRQ vector at a temperature below 30°C may have increased
Figure 3.3. Bacterial expression of the MLRQ subunit as a GST-fusion protein. Temporal induction of BL21 cells carrying the pGEX-3X-MLRQ vector with 0.1 mM IPTG at 30°C. Cells were induced when OD$_{600}$ reached a value of 0.7. **Lane 1**, uninduced sample at OD$_{600}$ = 0.7; **lanes 2-6**, induced samples after 1, 2, 3, 4 hours and overnight, respectively. Arrow indicates the MLRQ-GST fusion protein with a combined molecular mass of 35 kDa.
Figure 3.4. Purification and cleavage of GST-MLRQ fusion protein. 

a) Purification of GST-MLRQ fusion protein on a Glutathione Sepharose 4B column. Lane 1, E.coli BL21/pGEX-3X-MLRQ lysed (soluble fraction); lane 2, lysed insoluble fraction; lanes 3 and 4, flow-through from column upon applying soluble fraction; lanes 5 and 6, first and last washes of column.

b) Lanes 1 and 2, fusion protein eluates upon addition of glutathione elution buffer.

c) Digestion products from overnight cleavage of GST-MLRQ fusion protein in a Factor Xa to fusion protein ratio of 1:10.

d) Lanes 1 and 2, Flow-through from Glutathione Sepharose 4B column following application of Factor Xa cleavage products. The 35 kDa protein represents the GST-MLRQ fusion protein; the GST protein is 26 kDa and the MLRQ protein is 9kDa. All protein samples were loaded onto 16% SDS-polyacrylamide gels and stained with Coomassie blue dye.
solubility of the fusion protein. Cleavage of the fusion protein with Factor Xa proved to be quite difficult and successful cleavage was achieved only once (Fig. 3.4d).

**Anti-sense expression of MLRQ: Attempts to determine subunit function**

The anti-sense strategy was based on the principle that the anti-sense MLRQ mRNA made by the expression vector introduced into a cell can undergo Watson-Crick hybridization to MLRQ mRNA or pre-mRNA and may via a variety of mechanisms inhibit translation of that mRNA into MLRQ protein. Overexpression of anti-sense mRNA can modulate the transfer of information from the gene to the protein by interfering with various levels of the process which includes transcription, RNA splicing, polyadenylation, translation or termination by preventing the binding of necessary regulatory proteins, interfering with the export of mRNA from nucleus to cytosol as well as by inducing structural changes in RNA by binding to it that would result in its degradation. It was postulated that by specifically blocking synthesis of the MLRQ protein, its role in the assembly and function of complex I could be determined.

The “sense” MLRQ construct encompassed 40 bp of the 5' region, the start and stop codons and 26 bp of the 3' region. The “anti-sense” MLRQ construct somewhat mirrored the “sense” construct covering 24 bp of the 5' region, the start and stop codons and 26 bp of the 3' region but cloned in an anti-sense orientation to the CMV promoter on the expression vector. Transfection efficiency was determined using different DNA to transfection reagent (carrier) ratios and was found to be optimal at 2 μg DNA with 10 μl of the reagent. Transfection efficiency was determined using pCMVβ which carries the β-galactosidase gene and changes colour upon addition of X-gal. After the optimization described in the Materials and methods section, the pREP9 system was used to deliver
MLRQ in both “sense” and “anti-sense” orientations into both transformed fibroblast and lymphoblast control cell lines. Selection studies showed that non-transfected control lymphoblast and transformed fibroblasts were both killed at a G418 concentration of 100 μg/mL. However, cell lines transfected with “sense” and “anti-sense” constructs did not do much better, failing to thrive past G418 concentrations of 200 μg/ml. Using lymphoblasts as a recipient cell line also proved to be a problem, as it was hard to readily distinguish the dead cells (non-transfected) from the live ones.

A new vector pCDNA 3.1+ was used as a shuttle to deliver these “sense” and “anti-sense” MLRQ constructs into transformed fibroblasts, as this linearized vector was capable of integrating into the cell’s chromosomes, thereby making for stable transfection. With the new vector and Superfect as the carrier, both “sense” and “anti-sense” transfected cells were now able to survive G418 concentrations up to 600 μg/ml. At a concentration of 800 μg/ml G418 however, clones transfected with the “sense” vector survived, whereas those transfected with the “anti-sense” vector showed poor growth. Mitochondria isolated from these clones at this point were subjected to immunodetection with MLRQ antibody, but no difference in MLRQ protein expression was detected between the “sense” and “anti-sense” clones or clones transfected with pCDNA 3.1 + as control (Fig. 3.5).

Explanations for the inefficacy of the anti-sense strategy became somewhat apparent when the MLRQ sequence in the respective orientation was successfully amplified using primers designed from within the pCDNA3.1 + vector from cells transfected with “anti-sense” (Fig. 3.6, lane 4) and “sense” (Fig. 3.6, lane 6) constructs but not from the control clones (Fig. 3.6, lanes 1 and 2). Since it was known that the “sense” and “anti-
Figure 3.5. Western blot analysis of MLRQ expression in SV40 immortalized fibroblasts transfected with sense and anti-sense pCDNA 3.1+ constructs. (a) 100 μg of mitochondria isolated from fibroblasts lane 1, control untransfected; lane 2, transfected with pCDNA 3.1+ vector; lane 3, transfected with anti-sense MLRQ pCDNA 3.1+ construct; lane 4, transfected with sense MLRQ pCDNA 3.1+ construct were subjected to western blot analysis with antibodies to the 49 kDa, B18 and MLRQ subunits of complex I. The transfected cells were harvested at 200 μg/μl of G418 resistance. (b) 50 μg of mitochondria isolated from fibroblasts lane 1, transfected with pCDNA 3.1+; lane 2, transfected with anti-sense MLRQ pCDNA 3.1+ construct; lane 3, transfected with sense MLRQ pCDNA 3.1+ construct were immunodetected with antibodies to the 49 kDa and MLRQ subunits. Transfected cells were harvested at 800 μg/μl of G418 resistance.
Figure 3.6. PCR amplification of sense and anti-sense MLRQ sequences to confirm transfection of fibroblasts. Lanes 1 and 2 are negative controls. Lanes 3 and 5 are positive controls. PCR amplification of Lane 1, anti-sense MLRQ sequence from pCDNA 3.1+ transfected fibroblasts; in 2, sense MLRQ sequence from pCDNA 3.1+ transfected fibroblasts; in 3, anti-sense MLRQ sequence from anti-sense MLRQ pCDNA 3.1+ vector; in 4, anti-sense MLRQ sequence from anti-sense MLRQ transfected fibroblasts harvested at 400 µg/µl G418 resistance; in 5, sense MLRQ sequence from sense MLRQ pCDNA 3.1+ vector; in 6, sense MLRQ sequence from sense MLRQ transfected fibroblasts harvested at 400µg/µl of G418 resistance. PCR reactions were carried out 35 cycles at an annealing temperature of 50°C. PCR products were run on a 1% agarose gel.
"sense" MLRQ constructs had been successfully delivered to the nucleus of the cell by the carrier, this suggested a couple of other factors that were at play here. Firstly, due to high-order mRNA structure, the construct designed to target the mRNA of MLRQ may not have been effective in inhibiting its expression. Secondly, the presence of other MLRQ-like mRNAs may have interfered with the anti-sense RNA. Lastly, the transcription rate of the MLRQ sequence in the "anti-sense" orientation may not have been high enough to match the transcription of the native MLRQ mRNA in the cell. This may have also been compounded by the slow rate of degradation or chemical half-life of the native MLRQ mRNA in the cell compared to that of the anti-sense RNA. With steady-state levels of MLRQ mRNA not affected, an anti-sense effect at the protein level would not be seen. The fact that clones with the expression vector containing the MLRQ sequence in the "sense" orientation did not lead to increased levels of MLRQ protein may also be attributed to the faster rate of degradation of any MLRQ mRNA produced by the expression vector. Protection of mRNA by formation of ribonucleoproteins is also a well-documented phenomenon which has been reported to impair binding of anti-sense RNA (Strickland et al., 1988). Conclusive statements about the expression of MLRQ mRNA in either sense or anti-sense orientations can only be made by examining levels of the respective RNAs on a Northern blot.

Part II. Subunit interactions of MLRQ within complex I

**Detergent solubilization of complex I subunits**

Varying concentrations of many detergents and detergent cocktails were used to extract MLRQ from bovine heart mitochondria. Non-ionic detergents such as Octyl glucoside, DDM, Triton X-100, Triton X-114 and Nonidet-P40, anionic detergents such
Figure 3.7. Extraction of the MLRQ subunit from bovine heart mitochondria. Mitochondria isolated from bovine heart were extracted with a variety of detergents and detergent cocktails and the solubilized protein was immunodetected with the MLRQ and/or 49 kDa antibodies (as indicated by arrows) to determine extraction efficiency. All lanes contain 20 μg of extracted protein from single extractions (unless otherwise specified) with the indicated detergent(s) S, supernatant fraction; P, pellet fraction; suffix of 1 or 2 indicate the number of extractions that resulted in the fraction (a) Lane 1 contains purified bovine heart mitochondria. Lanes 2-8 represent corresponding pellets and supernatants resulting from extractions with: In 2, 1% Octyl glucoside; ln 3, 1 % DDM; In 4, 1% CHAPS; ln 5, 1% Deoxycholate; In 6, 0.5% Deoxycholate; In 7, 0.5% Triton x-100; In 8, 1% Triton x-100 (b) Extraction with: lanes 1-3, 1% DDM/1.5% sodium deoxycholate; lanes 4-6, 1% Nonidet-P40/0.5% sodium deoxycholate (c) Extractions with: lanes 1-2, 1% DDM; lanes 3-4, 1% CHAPS; lanes 5-6, 1% Octylglucoside. (d) Extractions with: lanes 1-2, 0.2% DDM, ln 3-4, 1% SDS; lanes 5-6, 1% Triton x-100
as SDS and sodium deoxycholate as well as zwitterionic detergents such as CHAPS were used in order to determine the detergent solubility of MLRQ as compared to other complex I subunits (Fig. 3.7). Preliminary experiments with detergents demonstrated that DDM (Fig. 3.1a, lane 3) and Triton X-100 (Fig. 3.7a, lanes 7,8) were the most effective at solubilizing complex I subunits. This can be noted by observing relative amounts of the MLRQ subunit present in the pellet versus the supernatant after extraction with various detergents (Fig 3.7a). Although DDM has become the detergent of choice in many experiments involving solubilization of other respiratory chain enzymes, only a recent study corroborates our findings on the usefulness of this detergent in the purification of complex I (Okun et al, 2000). The non-denaturing properties of DDM make it a favourable detergent for selective solubilization, so that close subunit-subunit interactions within the complex are preserved for examination (Okun et al, 2000). n-dodecyl-β-D-maltoside (DDM) is an analogue of octyl glucoside that possesses small, uniform micelles, a simple and chemically well-defined structure as well as a high critical micelle concentration that permits easy removal by dialysis (Rosevear et al, 1980). The conformation of the sugar moiety on DDM has been implicated as the critical factor in influencing the micelle forming abilities of this detergent (Rosevear et al, 1980). However, even though subunits such as 49 kDa, B18, B17 and MWFE were successfully removed from the pellet and into the supernatant with a single extraction with DDM, solubilization of MLRQ proved to be more difficult, with a good portion of it still remaining in the pellet fraction even after three extractions with the detergent (Fig. 3.8). This result is quite interesting taking into consideration the fact that other subunits belonging to the hydrophobic (HP) fraction such as B18, B17 and MWFE were easily
Figure 3.8. Solubilization of the MLRQ subunit compared to other complex I subunits. Three extractions of bovine heart mitochondria were performed with 1% DDM and the supernatant and pellet fractions collected each time. 20 µg of each fraction was loaded onto a 16% SDS-polyacrylamide gel and western blot analysis was performed with antibodies to the complex I subunits MLRQ, 49 kDa, B17, AQDQ/18IP, B18 and MWFE. C, (control) purified bovine heart mitochondria; S1, supernatant from 1st extraction; P1, pellet from first extraction; S2, supernatant from 2nd extraction; P2, pellet from 2nd extraction; S3, supernatant from 3rd extraction; P3, pellet from 3rd extraction.
Figure 3.9. Hydropathy profile and membrane orientation of the MLRQ polypeptide. (a) Hydropathy plot for the MLRQ subunit based on the hyrophobicity indices of Kyte and Doolittle (1982) averaged across six residues. (b) Predicted membrane topology of MLRQ based on its amino acid sequence.
extracted compared to MLRQ. While the 49 kDa subunit and MWFE have been assigned to the α-subcomplex of the enzyme, the B18 and B17 subunits are actually part of the mainly hydrophobic β-subcomplex with B18 in the IβS fraction and B17 in the IβL fraction associated with the ND4 and ND5 subunits. The Kyte-Doolittle hydropathy profile indicates that the human MLRQ subunit has the potential to be folded into a membrane spanning α-helix (Fig. 3.9). The first 39 amino acids are relatively hydrophobic and it is thought that this protein is anchored in the inner membrane and interacts via more hydrophilic parts of the sequence with other subunits of complex I. However, the extremely low detergent solubility of the MLRQ subunit compared to the aforementioned hydrophobic subunits of the complex suggests that it is perhaps embedded more deeply and heavily bound to inner membrane phospholipids than originally hypothesized.

**Proximity and association of MLRQ with other complex I subunits**

Immunoprecipitation after extraction with detergents was carried out in order to determine the association of MLRQ with other subunits of complex I. Subunits that are in close association with MLRQ should immunoprecipitate along with it. Antibodies to either the 49 kDa, MLRQ or MWFE subunits were added to 1% DDM solubilized bovine heart mitochondria followed by pelleting the immunoprecipitate with protein A agarose and analyses of the immunoprecipitates through SDS-PAGE and western blotting. The MLRQ antibody (produced in rabbit) proved to be quite effective in immunoprecipitating the subunit from bovine heart mitochondria (Fig. 3.10a, lane 4). The MLRQ subunit at 9 kDa and the IgG at 55 kDa and 30 kDa (doublet) were visible upon western blotting with the MLRQ antibody. The identity of IgG was confirmed through MALDIToF protein
Figure 3.10. *Immunoprecipitation studies using protein A agarose*. Bovine heart mitochondria was extracted with 1% DDM and immunoprecipitated with MLRQ antibody as detailed in the materials and methods section. (a) **Lane 1**, supernatant from immunoprecipitation after addition of protein A-agarose and MLRQ antibody; lane 2, first wash with washing buffer 1 (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% Nonidet P-40 and 0.5% sodium deoxycholate); lane 3, last wash with washing buffer 3 (50 mM Tris-HCl, pH 7.5, 0.1% Nonidet P-40 and 0.05% sodium deoxycholate); lane 4, MLRQ immunoprecipitate. Immunodetection with the MLRQ antibody (in 4) gives 4 bands: 9 kDa (MLRQ), 55 kDa (IgG) and a doublet at around 30 kDa. (b) 25 μg of the immunoprecipitated sample was therefore run on a 5-20% Tris-glycine gradient gel, stained with coomassie blue and the bands excised and sequenced through MALDI-ToF peptide mass fingerprinting. The double band at 30 kDa was found to be part of the IgG.
sequencing of these bands (Fig. 3.10b). Ease of immunoprecipitation with the MLRQ antibody which is directed at epitopes at the C-terminus of the subunit, suggests that the hydrophilic C-terminus is a region of MLRQ that is exposed. As a positive control, the supernatant from bovine heart mitochondria after 3 extractions with 2% Triton X-100 which was previously shown to have solely the MLRQ subunit, also showed similar results upon immunoprecipitation. In order to confirm that the band appearing at 9 kDa was not just a ubiquitous fragment, the same sample immunoprecipitated with MLRQ, treated with 1% SDS did not produce the 9 kDa MLRQ band upon immunoblotting with the MLRQ antibody.

Immunoprecipitation of MLRQ and immunoblotting the resulting pellet with the respective complex I antibodies revealed that neither the 39 kDa, 15 kDa, 18 IP (AQQD) or ASHI subunits have an affinity for MLRQ in order to be immunprecipitated along with it (Fig. 3.11a, c, d, lane 2). MWFE and the B18 subunit to some extent were the only subunits detected in the MLRQ immunoprecipitate (Fig. 3.11b, lane 2). In order to confirm this, the MWFE and 49 kDa immunoprecipitates were probed with the MLRQ antibody. The 9 kDa MLRQ subunit was indeed detected in the MWFE immunoprecipitate (Fig. 3.11a, lane 3) and to a lesser extent in the 49 kDa immunoprecipitate (Fig. 3.11a, lane 4). In addition, immunoblotting detected B18 (Fig. 3.11b, lane 3) and the 49 kDa subunit (Fig. 3.11c, lane 3) in the MWFE immunoprecipitate and B18 (Fig. 3.11b, lane 4) and MWFE (Fig. 3.11b, lane 4) in the 49 kDa immunoprecipitate, again in small amounts. Neither AQQD (Fig. 3.11a, lanes 3 and 4) nor ASHI (Fig. 3.11c, lanes 3 and 4) were detected in the MWFE or 49 kDa immunoprecipitates. These results indicate association between the 49 kDa, B18, MLRQ
Figure 3.11. Immunoblotting of the immunoprecipitated MLRQ, MWFE and 49 kDa subunits of complex I to determine subunit association. Bovine heart mitochondria solubilized with 1% DDM were immunoprecipitated with MLRQ, MWFE and 49 kDa antibodies (see materials and methods section), run on 5-20% Tris-Glycine gradient gels and transferred to nitrocellulose membranes for western blot analysis. 20 μg of protein sample were loaded in each lane. Lane 1, control bovine heart mitochondria; lane 2, MLRQ immunoprecipitant; lane 3, MWFE immunoprecipitant; lane 4, 49 kDa immunoprecipitant. Blots were immunodetected with antibodies to the following subunits (a) MLRQ and AQDQ/18IP (b) MWFE and B18 (c) 49 kDa and ASHI (d) 39 kDa and 15 kDa.
and MWFE subunits. Of these, the association between MLRQ and MWFE seems to be the most significant. The fact that MWFE was able to co-immunoprecipitate MLRQ but not vice-versa might be explained by the fact that the region of interaction between MLRQ and MWFE may have been masked by the interaction of the MLRQ antibody with MWFE. Another possibility is that the binding of the MLRQ antibody to the MLRQ subunit may have in fact distorted the protein such that it no longer interacted with MWFE. These findings support the emerging structural model of complex I which purports that MLRQ, MWFE and the 49 kDa subunit belong to the α-subcomplex of the enzyme. Even though the 49 kDa subunit belongs to the more hydrophilic Iα fraction within the Iα subcontext, it is known to form part of the pocket for quinone binding and is said to be in proximity to the ND1 subunit which is part of the membrane arm. Given this information, it is likely that both MLRQ and MWFE which have a transmembrane domain do interact with the 49 kDa subunit. The following table summarizes the information obtained from immunoprecipitation studies.

**Table 3.1. Subunit proximity in complex I as determined by immunoprecipitation studies**

<table>
<thead>
<tr>
<th>Subunits immunodetected</th>
<th>MLRQ</th>
<th>MWFE</th>
<th>49 kDa</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLRQ</td>
<td>Yes</td>
<td>Yes</td>
<td>SI</td>
</tr>
<tr>
<td>MWFE</td>
<td>SI</td>
<td>Yes</td>
<td>SI</td>
</tr>
<tr>
<td>49 kDa</td>
<td>No</td>
<td>SI</td>
<td>SI</td>
</tr>
<tr>
<td>18 IP (AQDQ)</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>39 kDa</td>
<td>No</td>
<td>Not ck</td>
<td>Not ck</td>
</tr>
<tr>
<td>15 kDa</td>
<td>No</td>
<td>Not ck</td>
<td>Not ck</td>
</tr>
<tr>
<td>ASHI</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>B18</td>
<td>SI</td>
<td>SI</td>
<td>SI</td>
</tr>
</tbody>
</table>

Not ck – Not checked; SI – slightly
Figure 3.12. Immunoblots of 1mM DST and 0.2mM EGS cross-linked bovine heart mitochondria with antibodies to the MLRQ and 49 kDa subunits. Bovine heart mitochondria cross-linked with 1mM DST and 0.2mM EGS for 1 hour and quenched with 5mM glycine, was analyzed by gel electrophoresis on 5-20% Tris-glycine gradient gels and immunoblotted as described under materials and methods. (a) Immunoblotting was done with antibody to the 49 kDa subunit. Lanes 1 and 5, 20 µg of non cross-linked bovine heart mitochondria; lanes 2 and 6, 40 µg of cross-linked mitochondria; lanes 3 and 7, 20 µg of cross-linked mitochondria. (b) Immunoblotting was done with antibody to the MLRQ subunit. Lane 1, 20 µg non cross-linked bovine heart mitochondria; lanes 2 and 5, 40 µg of cross-linked mitochondria; lanes 3 and 6, 20 µg of cross-linked mitochondria; lanes 4 and 7, 8 µg of cross-linked mitochondria. All lanes 2-4 contain 1mM DST cross-linked samples while all lanes 6-8 contain 0.2mM EGS cross-linked samples.
Cross-linking and immunoprecipitation studies

Chemical cross-linkers EGS and DST are both water-insoluble, homobifunctional $N$-hydroxysuccinimide esters (NHS-esters). Primary amines such as the $\alpha$-amine groups present on the $N$-termini of peptides and proteins as well as in some cases, the side chains of amino acids react with these NHS-esters to form covalent amide bonds. Preliminary experiments examined the cross-linking of the MLRQ, MWFE and 49 kDa subunits of bovine heart mitochondria with 1mM DST and 0.2 mM EGS by detecting with the respective antibodies. At such low cross-linker concentrations, MWFE was not found to cross-link with either DST or EGS. However, a slight band of 20 kDa is visible when the cross-linked mitochondria is immunodetected with the MLRQ antibody (Fig. 3.12b, lane 2). In addition to the possibility of MLRQ associating with other subunits approximately 11 kDa in size such as B8, B9 or AGGG (10 kDa), this band could also represent a dimer of MLRQ (9 + 9). The 49 kDa subunit was however found to clearly cross-link, yielding a band of approximately 60 kDa (Fig. 3.12a, lanes 2, 3, 6 and 7) and could in fact represent the association between the 49 kDa subunit and a 13 kDa subunit of IP (49 + 13). This is in agreement with studies performed by Yamaguchi and Hatefi (1993) who have shown the 49 kDa subunit to be linked to a 13 kDa subunit using an antibody made to a mixture of the two 13 kDa proteins, B13 and DDGD. The cross-linking patterns of DST and EGS were essentially the same, despite the considerable differences in the lengths of the reagents (molecular lengths of DST and EGS being 0.64 and 1.61 nm, respectively).

When cross-linker concentrations of EGS were increased to 0.2M and 0.5M, more obvious cross-linking patterns began to emerge. No differences between cross-linking
Figure 3.13. Immunoblotting of EGS cross-linked bovine heart mitochondria with complex I antibodies. 4 mg/ml bovine heart mitochondria was cross-linked for 2 hours and quenched with 20mM glycine and analysed on western blots as described under materials and methods. Lane 1, 20µl of 0.2M EGS cross-linked mitochondria (4 mg/ml); In 2, 20µl of 0.5 M EGS cross-linked mitochondria (4 mg/ml); In 3, 20 µg of non cross-linked mitochondria. Immunoblotting was performed with antibodies to (a) AQDQ/18 IP subunit (b) ASHI subunit (c) MLRQ subunit (d) MWFE subunit (e) 49 kDa subunit (f) B17 subunit (g) B18 subunit. Molecular masses are indicated for prominent cross-linked products.
patterns were observed using 0.2M or 0.5M EGS. Cross-linking at these higher concentrations and probing with 18 IP (AQDQ) revealed a double band approximately 36- and 40-kDa in size (Fig. 3.13a, lanes 1 and 2). The smaller of these unidentified bands could represent dimerization of the AQDQ subunit \((18 + 18)\) and was in fact a possibility proposed by Yamaguchi and Hatefi (1993), who observed a similar sized band when the EGS cross-linked IP fraction was probed with the 18IP antibody.

Probing the cross-linked mitochondria with the ASHI antibody, revealed a prominent double band at 38- and 43-kDa as well as less prominent bands at approximately 60- and 110 kDa (Fig. 3.13b, lanes 1 and 2). Again, the smallest of the bands may represent a dimer of the subunit \((18 + 18)\).

Cross-linked samples immunodetected with B17 revealed just one prominent band at 33 kDa (Fig. 3.13f, lanes 1 and 2), while the same samples immunodetected with B18 revealed two prominent bands at 30 kDa and 34 kDa as well as other bands of higher molecular weight and lesser intensity (Fig. 3.13g, lanes 1 and 2).

Cross-linking with 0.2M and 0.5M of EGS resulted in the appearance of many cross-linked bands when probed with the 49 kDa antibody, the most prominent ones being at 60 kDa, 79 kDa, 103 kDa and 135 kDa (Fig. 3.13e, lanes 1 and 2). As mentioned above, the 60 kDa band probably represents association between the 49 kDa subunit and one of the 13 kDa subunits. The 79 kDa band most likely represents the binding of the 30 kDa subunit \((49 + 30)\), also detected by Yamaguchi and Hatefi (1993). The myriad of other associations can only be speculated upon in a similar context. The fact that the 49 kDa subunit is not buried in the membrane is evidenced by the number of cross-linking associations formed by this subunit as compared to the other subunits examined here.
Immunoblotting these cross-linked samples with the MLRQ antibody, detected a very prominent band at approximately 20 kDa, a lighter band at approximately 31 kDa and a slight doublet at 40 kDa and 45 kDa (Fig. 3.13c, lanes 1 and 2). While the prominent band at 20 kDa may represent the aforementioned dimerization of MLRQ and the doublet at 40- and 45- kDa, dimeric or trimeric cross-linking to other subunits in the vicinity, the band at 31 kDa may in fact represent cross-linking of MLRQ to the MWFE subunit. This possibility arises from the presence of a single band of the same molecular mass detected in the cross-linked samples using the MWFE antibody (Fig. 3.13d, lanes 1 and 2). A molecular mass of 31 kDa however indicates that MLRQ and MWFE, in addition to their own association are also closely linked to other complex I subunits. Even so, this assumption leads to the question as to why no dimeric association between the two subunits was detected. One possibility is that there is no direct binding between MLRQ and MWFE and that any association between the two subunits, no matter how close, is through an intermediate subunit or subunits. A similar scenario is evident in the cross-linking experiments of Yamaguchi and Hatefi (1993) who detected trimeric cross-linking between the 75 kDa, 51 kDa and 24 kDa subunits, but no dimeric associations between 75- and 24- kDa or 75- and 9- kDa subunits. It has been established through other experiments as well, that the 9 kDa and 24 kDa subunits are linked to the 75 kDa subunit as well as to the rest of complex I, solely through their association with the 51 kDa subunit.

Assuming trimeric cross-linking between these subunits (9 + 6 + X), subunit X which is approximately 16 kDa in size, can be any one of B18, SGDH (18 kDa) or PGIV (19 kDa). ASHI (18 kDa) and B17 can be eliminated as probable candidates for subunit X.
because a band of approximately 31 kDa was not visualized in the cross-linked samples probed with either the ASHI (Fig. 3.13b, lanes 1 and 2) or B17 (Fig. 3.13f, lanes 1 and 2) antibody. A cross-linked product approximately 31 kDa in size was detected with the B18 antibody (Fig. 3.13g, lanes 1 and 2) and this subunit could very well be the intermediary subunit between MLRQ and MWFE, especially because it had been found previously to immunoprecipitate with both subunits (Fig. 3.11b, lanes 2 and 3).

However, the validity of this association is questioned by the fact that cross-linked products allowing for dimeric cross-linking between B18 and MLRQ or MWFE was not detected. While PGIV belongs to the Iα fraction of complex I, as do MLRQ and MWFE, the SGDH and B18 subunits belong to the β subfraction. A four subunit association including MLRQ and MWFE would involve only the smaller subunits such as B8, B9, SDAP (10 kDa), MNLL (5.5 kDa), AGGG (10 kDa) or KFYI (5 kDa). However, without antibodies to confirm the identity of these other subunits, this is merely conjecture at this point. Table 3.2 lists the cross-linked products detected by the complex I antibodies used. As expected, a greater number of cross-linkings were seen with subunits that are not embedded in the mitochondrial membrane.

Since complex I has 43 subunits and antibodies to most or all of them were not available to this study, immunoprecipitation of the cross-linked proteins was attempted in order to obtain enough of the cross-linked proteins on a SDS gel for protein sequencing. Immunoprecipitation of cross-linked samples with MLRQ was not very successful yielding only very minute quantities of MLRQ and none of any of the other subunits cross-linked to it (Fig. 3.14a, lane 3). Immunoprecipitation with MWFE on the other hand was very successful, yielding relatively large amounts of the MWFE subunit.
Figure 3.14. Immunoblotting of EGS cross-linked bovine heart mitochondria immunoprecipitated with MLRQ and MWFE antibodies. Mitochondria cross-linked for 2 hours with 0.5M EGS were solubilized with 1% DDM and subjected to immunoprecipitation with antibodies to MLRQ and MWFE as described under materials and methods. Immunoprecipitates from the cross-linked samples were run on 5-20% Tris-glycine gradient gels and then analysed by western blotting with antibodies to (a) the MLRQ subunit and (b) the MWFE subunit. 20 µg of protein was loaded in each lane. Lane 1, non cross-linked bovine heart mitochondria; lane 2, cross-linked bovine heart mitochondria; lane 3, cross-linked bovine heart mitochondria immunoprecipitated with MLRQ antibody; lane 4, cross-linked bovine heart mitochondria immunoprecipitated with MWFE antibody. Arrows point to immunoprecipitated products.
Table 3.2. Cross-linked products detected by immunodetection with various complex I antibodies

<table>
<thead>
<tr>
<th>Antibody Used</th>
<th>Cross-linked products</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLRQ</td>
<td>20 kDa 31 kDa 40 kDa 45 kDa</td>
</tr>
<tr>
<td>MWFE</td>
<td>31 kDa</td>
</tr>
<tr>
<td>49 kDa</td>
<td>60 kDa 79 kDa 103 kDa 135 kDa + **</td>
</tr>
<tr>
<td>AQDQ/18IP</td>
<td>36 kDa 40 kDa</td>
</tr>
<tr>
<td>ASHI</td>
<td>38 kDa 43 kDa 60 kDa 110 kDa + **</td>
</tr>
<tr>
<td>B18</td>
<td>30 kDa 34 kDa + *</td>
</tr>
<tr>
<td>B17</td>
<td>33 kDa</td>
</tr>
</tbody>
</table>

+** many other less prominent cross-linked products detected

(Fig. 3.14b, lane 4). However, none of the other subunits cross-linked to MWFE were found to co-immunoprecipitate with the subunit. This phenomenon may be linked to two factors: first, an excess of cross-linking may have resulted in the loss of material or reduced antigen availability in the MLRQ/MWFE subunits or both and secondly, the relative sensitivity of the antigen epitopes on these subunits to the EGS cross-linker.

Although the identity of the other subunits that cross-link with MLRQ can be deduced to some extent, positive identification needs evidence in the form of immunodetection with antibodies to the respective subunits or protein sequencing of the immunoprecipitated products.
Chapter 4

Postulating the role of a supernumerary subunit such as MLRQ in complex I function and dysfunction

Conclusions and Future Directions

Part I. Molecular structure of MLRQ

A 515 bp cDNA that codes for the human MLRQ subunit belonging to the NADH:ubiquinone oxidoreductase complex of the mitochondrial respiratory chain has been cloned. The NDUFA4 gene encoding this MLRQ subunit has been mapped to chromosome 7 p21-22 and a pseudogene has been localized to chromosome 1 p21. The pseudogene sequence was found to have 97% similarity to the MLRQ cDNA sequence and even though the start and stop codons are preserved, there are sufficient amino acid changes in the coding region of the gene itself, that if expressed would make it a very different protein from MLRQ. The pseudogene was found to be intronless and possibly not expressed in human tissues such as brain, heart, liver and kidney as well as fibroblasts and cardiomyocytes. A BAC clone containing sequence from chromosome 1 documented in the HTGS database has now been found to match this intronless pseudogene sequence. The NDUFA4 genomic sequence has been found to span approximately 6.5 kb, with three introns estimated to be 1.1, 0.66 and 4.4 kb in size. The existence of many other MLRQ-like sequences seems to suggest that this supernumerary subunit probably appeared quite early on in the evolutionary pathway of mammalian complex I, in order for its gene to have undergone the number of changes that it has. Also, the high level of conservation seen in MLRQ nucleotide and amino acid sequences
between the mouse, bovine and human species suggests a definite functional significance for the subunit in complex I.

Now that both the cDNA and genomic structure of MLRQ have been determined, an extensive mutational detection study of NDUFA4 is possible. Although our screening of MLRQ cDNA did not find any mutations in the ORF that are responsible for complex I deficiency, a greater subset of patients need to be examined before a definite statement can be made regarding the involvement of MLRQ in complex I deficiency. Future studies should include examination of the promoter region of this gene and subsequent screening for mutations. Also, further studies can answer whether mutations in the cis-acting transcription elements or in the trans-acting transcription factors for NDUFA4 are responsible for complex I deficiency. Even though mutations have now been found in a few core subunits of complex I encoded by the nuclear genome, nuclear-encoded genes responsible for complex I deficiency in numerous other cases are yet to be detected. It is particularly important at this point that molecular studies look further than abnormalities in structural complex I proteins. Molecular studies concerning aberrant complex I function should focus on other genes involved in the regulation, mitochondrial import and assembly of this large, multi-protein enzyme complex. The recent identification of the SURF1 gene in COX-deficient patients with Leigh syndrome (Tiranti et al., 1998; Zhu et al.; 1998) is a good example of the emerging class of gene defects which are responsible for the abnormal functioning of mitochondrial respiratory chain enzymes. The discovery of two novel chaperones involved in the assembly of complex I in *N. crassa* sets a precedent for such studies (Kuffner et al., 1998). Complementation of complex I activity using techniques such as micro-cell mediated chromosome transfer
(Zhu et al, 1998) in combination with linkage analysis using polymorphic microsatellite markers in patient groups could provide answers as to the underlying genetic cause of their deficiency by identifying chromosome regions containing essential candidate genes.

Part II. Biochemical characterization of MLRQ

Tissue expression of the MLRQ protein revealed that it is ubiquitously expressed. Higher expression of MLRQ in human tissues with high metabolic energy rates such as heart and muscle is indicative greater numbers of mitochondria in these tissues. Since bacterial expression and purification of the MLRQ subunit as a fusion protein was successful, future studies can look at fine tuning cleavage conditions in order to examine the structure of the purified protein using techniques such as CD spectroscopy, NMR and X-ray crystallography. Expression of the MLRQ fusion protein in minimal media labeled with $^2\text{H}$ (Deuterium) or doubly enriched by $^{13}\text{C}$ and $^{15}\text{N}$, followed by subsequent cleavage will facilitate NMR study of the subunit. Useful information on the secondary and tertiary structure of MLRQ can be obtained from CD spectra and X-ray crystal maps, respectively. The tryptophans present in MLRQ make fluorescence spectroscopy a technique that can be employed to obtain an emission spectra of the protein. However, although these techniques would be useful in characterizing the structure of the expressed protein, any real information on the conformation of MLRQ can only be discerned by examining this subunit in various lipid environments that mimic the inner mitochondrial membrane. Because MLRQ is one subunit in a huge complex made up of 43 subunits, even this approach can only give a rough approximation of the true nature of the subunit in complex I. In addition, the bacterially expressed MLRQ fusion protein itself can be
used for production of antibodies against the full human MLRQ protein. Generation of a
good MLRQ monoclonal antibody might be invaluable not only for screening patients but
also in future immunoprecipitation studies. The discovery of a patient with hepatopathy
and renal tubulopathy, who showed underexpression of other complex I subunits but an
overexpression of MLRQ suggests a differential mechanism of regulation for this subunit
and a process of incorporation into the mitochondria that is not dependant on the titre of
other complex I subunits.

Part III. Localization of MLRQ within complex I

A definite conclusion that can be drawn from our extraction studies is that the MLRQ
subunit has a great affinity for the inner membrane. Even though it has a considerable
hydrophilic domain which interacts with other subunits, the transmembrane domain of
the subunit anchors it very strongly to the membrane, thereby providing integrity and
stability to the entire complex. The fact that other subunits in the HP fraction were
solubilized easily (even subunits thought to be entirely in the membrane arm) indicates
that MLRQ or at least its transmembrane domain is embedded quite deeply in the
membrane with perhaps a high affinity for phospholipid membrane components.
Immunoprecipitation studies show that MLRQ is in close proximity to MWFE as well as
the 49 kDa and B18 subunits to a lesser extent. This reiterates the localization of the
subunit to the Iα subcomplex but also denotes an interaction with the hydrophobic Iβ
subcomplex. It is likely then that MLRQ makes up part of the bulky stalk region of
complex I (Fig. 4.1). Definitive statements cannot be made about the association of
Figure 4.1. Schematic representation of the proximity of MLRQ to other subunits of complex I. Circles represent individual subunits and the numbers inside the circles show their masses in kilodaltons. This model represents a summary of the information on the putative associations between MLRQ and other subunits in the Iα and Iβ fractions as gathered from extraction, immunoprecipitation and cross-linking studies. 49, 49 kDa subunit; 6, MWFE; 9, MLRQ; ~16, B18, PGIV or SGDH.
MLRQ to other complex I subunits without considerable proof of identity of the cross-linked products. However, by immunodetection with complex I antibodies, it seems likely that MLRQ is associated with MWFE through an intermediary subunit(s).

Deducing the identity of this subunit using molecular mass considerations suggests that in a trimeric cross-linking scenario, this subunit can be any one of B18, SGDH (18 kDa) or PGIV (19 kDa). The B18 subunit stands out in particular because it is one of the subunits immunoprecipitated by the MLRQ antibody. A four-subunit association would implicate many of the smaller subunits of complex I such as B8, B9, SDAP (10 kDa), MNLL (5.5 kDa), AGGG (10 kDa) or KFYI (5 kDa). Future studies need to focus on efficient immunoprecipitation of these cross-linked products. Modification of cross-linking conditions and times prior to immunoprecipitation would greatly enhance this endeavor.

Analytical ultracentrifugation of solubilized complex I is another approach that can be taken to resolve the subunit-subunit interactions of MLRQ. Sedimentation velocity can provide experimental evidence on the size and shape of MLRQ while sedimentation equilibrium can answer thermodynamic questions about the molar mass of MLRQ, its association constants and stoichiometries of association with other subunits. Two-dimensional blue native gel electrophoresis of complex I followed by western blotting can also yield beneficial information on the association of various subunits with MLRQ in complex I.

Part IV. Possible roles for MLRQ in complex I function

It is difficult to envisage the functionality of a subunit such as MLRQ from protein homology alone as database searches do not reveal significant sequence homology to any
other known protein. However, based on its biochemical and molecular characterization combined with current knowledge on the structure and subunits of complex I, certain conclusions about the nature of this protein can be drawn. Even though MLRQ is not one of the core proteins of complex I involved in electron transport and mutations have not been found in any of the supernumerary subunits of complex I, its role in the functioning of the enzyme may not necessarily be trivial. The only accessory proteins with known functions were the 39 kDa subunit which together with the acyl carrying SDAP protein was shown to be involved in the biogenesis of the complex (Friedrich et al. 1995; Walker et al, 1992). More recently however, deletion mutation studies on Chinese hamster cell lines demonstrated how another supernumerary subunit, MWFE, is essential for complex I activity (Au et al, 1999). This study gives relevant insight into the possible functioning of MLRQ because of the similarities shared between the MLRQ and MWFE subunits. They are both small MW integral membrane proteins with a hydrophobic N-terminus and hydrophilic C-terminus and have been localized to the Iα subcomplex. Like MWFE, MLRQ also lacks a signal sequence at the N-terminus of the protein also suggesting that it is imported into mitochondria and associated with complex I without requiring proteolytic processing. More significantly, our studies with immunoprecipitation and cross-linking have shown that the two subunits are in close proximity to one another and may be associated with each other, even if it is through another subunit. Given this, the MLRQ subunit may be functioning alongside with MWFE in the formation or integration of cofactors in complex I. The MWFE mutant cell line studied by Au et al (1999) showed a complete loss of complex I activity as well as decreases in the expression of the 75-, 51- and 24- kDa subunits which carry redox centres. Likewise, MLRQ may be
involved in the assembly or even in the import of these core subunits of complex I, similar to the subunits of complex III which serve as processing enzymes (Schulte et al, 1989). If the functionality of the MLRQ subunit is to be pursued through anti-sense strategies, certain changes need to be made to the attempted procedure. Firstly, since the anti-sense strategy is somewhat of an erratic procedure (for the reasons discussed in the previous chapter), constructs containing different portions of the MLRQ cDNA used to produce anti-sense RNA need to be tested. Previous studies have shown that the size of the anti-sense sequence has a great effect on the extent of mRNA inhibition (Scherczinger et al, 1992). However, a more useful approach might be to produce a MLRQ knockout in mouse and examine complex I activity in mutant cell lines derived from the organism if it proves viable. Expression of the MLRQ subunit in combination with other human complex I subunits in yeast would also give answers as to whether MLRQ is stable on its own or needs the other subunits to form part of the complex.

The MLRQ subunit is of particular importance because it possesses a transmembrane domain as well as a hydrophilic C-terminus which connects the membrane arm of complex I with the globular arm. It is likely then that this subunit forms part of the stalk which has been found to be thick enough to form an insulating layer around the electron pathway (Grigorieff, 1999). Ultimately, the MLRQ subunit together with the other supernumerary subunits of mammalian complex I probably forms a scaffold, a theme which is suggested by the position of the supernumerary subunits of COX, which package the core electron transferring subunits of the enzyme thereby reducing their contact with the surrounding phospholipids. Future studies looking at complex I dysfunction might also concentrate on oxygen free radical formation from supernumerary
subunits such as MLRQ and the role of these subunits in the assembly and function of complex I.

**The final word**

This work has answered questions and dispelled some of the uncertainties about the MLRQ subunit that were manifest at the beginning of this study. Contrary to what was originally thought, the MLRQ subunit has been localized to chromosome 7. We now know that it is not responsible for the X-linked mode of inheritance seen in certain cases of complex I deficiency. As foretold by the studies of Day and Scheffler (1982), the search is on for the other X-linked complex I subunit which can be confirmed through the chromosomal localizations of the B9, B12, B15, SGDH, MMTG and 17.2 kDa subunits or the discovery of a complex I assembly factor on the X chromosome. It can also be said that mutations in the coding region of NDUFA4 are not a common cause of complex I dysfunction. The availability of the genomic structure of NDUFA4 now makes it possible to screen for mutations in the promoter regions of the gene. The localization and confirmation of a pseudogene of MLRQ as well as the discovery of other MLRQ-like sequences has also helped explain the many incorrect localizations of the gene reported in GenBank databases. Lastly, the fact that MLRQ is part of complex I as opposed to merely coprecipitating with it has now been cemented by biochemical characterizations of the subunit through cross-linking and immunoprecipitation studies. This is supported by the ubiquitous expression of MLRQ in various human tissues, even though the regulation of the subunit may differ from that of some of the other complex I subunits. These investigations allow a better understanding of complex I function and dysfunction.
through the study of one of its supernumerary subunits and thereby fulfill the objectives set forth in this study.
REFERENCES


