Effects of missense mutations on protein function in human disease: tumor suppressor p53 and the AAA+ ATPase CLPB

by

Indhujah Thevarajan

B.Sc., University of Colombo (Sri Lanka), 2014

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Biochemistry and Molecular Biophysics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Abstract

The human genome consists of more than three billion base pairs, and an alteration in one base pair may lead to a disease condition. However, not all mutations have harmful effects and some mutations are beneficial to the organism. Mutations are important in evolution and allow a natural selection to occur. When the DNA sequence in a protein coding region is altered by point mutation, it may produce a partially or completely non-functional protein. In my study, we selected disease-associated missense mutations in two different human proteins: the p53 protein and the AAA+ ATPase CLPB. The goal of my research was to explore the effects of these missense mutations on protein function in disease.

Tumor suppressor *p53* is a frequently mutated gene in human cancers. Since previous studies mainly focused on mutations in the DNA binding domain (DBD) of p53 protein, very little is known about the structural and functional consequences of the cancer-associated mutations in the p53 transactivation domain (TAD). The level and activity of p53 are regulated by interactions of TAD with the negative regulator HDM2 and the general transcriptional co-activator CBP. Four cancer-associated mutations in p53-TAD were selected for this study and the interaction with its binding partners, HDM2 and CBP, were investigated using *in vitro* protein-binding assays utilizing Biolayer Interferometry (BLI). I found that the cancer-associated mutations can significantly perturb the balance of p53's interactions with the key activator (CBP) and the degradation regulator (HDM2).

Human caseinolytic peptidase B protein homolog (CLPB)/suppressor of potassium transport defect 3 (SKD3) belongs to the AAA+ family of ATPases associated with different activities. Mutations in the human *CLPB* gene cause 3-methylglutaconic aciduria type VII with cataracts, neurologic involvement, and neutropenia (MEGCANN), a rare autosomal recessive

genetic disease affecting children. The function of CLPB is unknown and its relation to the well-studied microbial disaggregase ClpB/Hsp104 has been controversial. The clinical manifestations of the mutations are currently challenging to explain due to the lack of understanding of the biological function of CLPB. To begin closing that gap in knowledge, I investigated the endogenous CLPB in mammalian cells and exogenously overexpressed CLPB. I found that endogenous human CLPB is found predominantly in the mitochondrial intermembrane space. I demonstrated that the selected disease-associated variants of human CLPB showed different levels of expression from WT CLPB. Overexpressed CLPB, while properly trafficked to the mitochondria, appeared to form large clusters/aggregates that were poorly extractable with non-ionic detergents and were readily visualized by immunofluorescence microscopy. Importantly, endogenous CLPB formed high molecular weight protein complexes in an ATP-dependent manner, indicating that ATP binding induces a conformational change in CLPB and controls its ability to self-associate or form complexes with other proteins. Collectively, these findings should inform future studies on the effect of pathogenic mutations on CLPB structure and function and on the mechanism of CLPB action in human cells.

Effects of missense mutations on protein function in human disease: tumor suppressor p53 and the AAA+ ATPase CLPB

by

Indhujah Thevarajan

B.S., University of Colombo (Sri Lanka), 2014

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Biochemistry and Molecular Biophysics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Approved by:

Co-Major Professor
Anna Zolkiewska

Approved by:

Co-Major Professor
Michal Zolkiewski

Copyright

Abstract

The human genome consists of more than three billion base pairs, and an alteration in one base pair may lead to a disease condition. However, not all mutations have harmful effects and some mutations are beneficial to the organism. Mutations are important in evolution and allow a natural selection to occur. When the DNA sequence in a protein coding region is altered by point mutation, it may produce a partially or completely non-functional protein. In my study, we selected disease-associated missense mutations in two different human proteins: the p53 protein and the AAA+ ATPase CLPB. The goal of my research was to explore the effects of these missense mutations on protein function in disease.

Tumor suppressor *p53* is a frequently mutated gene in human cancers. Since previous studies mainly focused on mutations in the DNA binding domain (DBD) of p53 protein, very little is known about the structural and functional consequences of the cancer-associated mutations in the p53 transactivation domain (TAD). The level and activity of p53 are regulated by interactions of TAD with the negative regulator HDM2 and the general transcriptional co-activator CBP. Four cancer-associated mutations in p53-TAD were selected for this study and the interaction with its binding partners, HDM2 and CBP, were investigated using *in vitro* protein-binding assays utilizing Biolayer Interferometry (BLI). I found that the cancer-associated mutations can significantly perturb the balance of p53's interactions with the key activator (CBP) and the degradation regulator (HDM2).

Human caseinolytic peptidase B protein homolog (CLPB)/suppressor of potassium transport defect 3 (SKD3) belongs to the AAA+ family of ATPases associated with different activities. Mutations in the human *CLPB* gene cause 3-methylglutaconic aciduria type VII with cataracts, neurologic involvement, and neutropenia (MEGCANN), a rare autosomal recessive

genetic disease affecting children. The function of CLPB is unknown and its relation to the well-studied microbial disaggregase ClpB/Hsp104 has been controversial. The clinical manifestations of the mutations are currently challenging to explain due to the lack of understanding of the biological function of CLPB. To begin closing that gap in knowledge, I investigated the endogenous CLPB in mammalian cells and exogenously overexpressed CLPB. I found that endogenous human CLPB is found predominantly in the mitochondrial intermembrane space. I demonstrated that the selected disease-associated variants of human CLPB showed different levels of expression from WT CLPB. Overexpressed CLPB, while properly trafficked to the mitochondria, appeared to form large clusters/aggregates that were poorly extractable with non-ionic detergents and were readily visualized by immunofluorescence microscopy. Importantly, endogenous CLPB formed high molecular weight protein complexes in an ATP-dependent manner, indicating that ATP binding induces a conformational change in CLPB and controls its ability to self-associate or form complexes with other proteins. Collectively, these findings should inform future studies on the effect of pathogenic mutations on CLPB structure and function and on the mechanism of CLPB action in human cells.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AAA	ATPases associated with different activities
AEBSF	4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride
AML	Acute myeloid leukemia
ANK	Ankyrin-repeat domain
APEX	Ascorbate peroxidase
ATP	Adenosine triphosphate
ATPγS	Adenosine 5'-(gamma-thio) triphosphate
Bcl-2	B-cell lymphoma 2
BLI	Biolayer interferometry
BN-PAGE	Blue-native polyacrylamide gel electrophoresis
BSA	Bovine serum albumin
CBP	Cyclic-AMP response-binding (CREB)-binding protein
CLPB	Caseinolytic peptidase B protein homolog
CD	Circular dichroism
CDKNIA	Cyclin dependent kinase inhibitor 1
CTD	C-terminal domain
DBD	DNA binding domain
DDB2	DNA damage-binding protein-2
DDM	n-dodecyl-β-D-maltoside
DDR	DNA damage repair
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic acid

DPBS	Dulbecco’s phosphate buffered saline
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EMS	Ethyl methanesulfonate
FBS	Fetal bovine serum
FDA	Food and drug administration
GST	Glutathione s-transferase
HAT	Histone acetyltransferases
HAX1	HCLS1-associated protein X-1
HDM2	Human double minute 2
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	Horseradish peroxidase
HRV 3C	Human rhinovirus 3C
HSP	Heat shock protein
IARC	International agency for research on cancer
IDR	Intrinsically disordered region
IMS	Intermembrane space
IPTG	Isopropyl-β-D-thiogalactoside
Kix	Kinase-inducible domain (KID) interacting domain
LB	Lysogeny broth
LFL	Li-Fraumeni-like syndrome
LFS	Li-Fraumeni syndrome
LOBSTR	Low background strain

MD	Molecular dynamics
MEGCANN	3-methylglutaconic aciduria with cataracts, neurological involvement and neutropenia
MGCA7	3-methylglutaconic aciduria, type VII
MPP	Mitochondrial processing peptidase
MTS	Mitochondrial targeting presequence
NBD	Nucleotide-binding domain
NCBD	Nuclear coactivator binding domain
NHEJ	Non-homologous end joining
Ni-NTA	Nickel nitrilotriacetic acid
NMR	Nuclear magnetic resonance
NRD	Negative regulatory domain
NTD	N-terminal domain
PAGE	Polyacrylamide gel electrophoresis
PARL	Presenilin-associated rhomboid-like protease
PHB	Prohibitins
PK	Proteinase K
PMSF	Phenylmethanesulfonyl fluoride
PNS	Post-nuclear supernatant
P-rich	Proline-rich domain
PTM	Post-translational modification
PVDF	Polyvinylidene difluoride
SDS	Sodium dodecyl sulfate

SKD3	Suppressor of potassium transport defect 3
TAD	Transactivation domain
TAZ1/2	Transcriptional adaptor zinc-binding 1/2
TIGAR	TP53-induced glycolysis and apoptosis regulator
UV	Ultraviolet

Acknowledgements

First, I would like to express my heartfelt gratitude to my advisors Dr. Michal Zolkiewski and Dr. Anna Zolkiewska for their continuous support throughout my graduate studies. I am thankful for their guidance, advice and encouragement that have helped me to grow as a researcher and a scientist. It has been a privilege to have them as my advisors.

I would like to extend my thanks to my committee members Dr. Jianhan Chen, Dr. Om Prakash and Dr. Kathrin Schrick for their guidance and advice during my study. I appreciate their time and effort to serve on my committee.

My sincere thanks go to my lab members Chathuranga Ranaweera, Dr. Sunitha Shiva, Dr. Przemyslaw Glaza and Yeni Romero for their encouragement and support throughout my studies. I would like to express my extra special thanks to Dr. Lynn Goss Schrag, fellow researcher on the p53 project, for his technical expertise and support during my course of study. All of them shared their experience and knowledge with me how to be a competent researcher and I enjoyed working with them.

I would like to extend my special thanks to Joel Sanneman from Confocal Core Facility of the College of Veterinary Medicine for teaching and helping me in acquiring confocal images. I wish to express my sincere thanks to Dr. Jeroen Roelofs from Division of Biology for serving on my committee until his departure from K-State and allowing me to access the BLItz instrument in his lab. I would like to thank each and everyone in Department of Biochemistry and Molecular Biophysics who helped me in one way or another during my study.

Finally, I would like to acknowledge my family and friends who always stood by me. Specially, I cannot thank my husband, Yasodarran enough. He is a constant pillar of support and I would not have been accomplished my Ph.D. studies without his encouragement and love.

Dedication

I wish to dedicate this dissertation to my husband, my family and my close friends for their love and support.

Chapter 1 - Literature review

1.1 Mutations

A mutation is an alteration in the DNA sequence of the genome. Human genome consists of more than three billion base pairs, and even a change in one base pair could lead to a disease condition. For example, sickle cell anemia is caused by a single nucleotide change in the hemoglobin-beta (HBB) gene. This single nucleotide change of the beta chain of the hemoglobin protein alters the sixth residue, glutamic acid to valine, and gives rise to sickle cell anemia [1]. Hundreds of diseases are caused by mutations of the genome, however not all mutations have noticeable effects. Some mutations are called “silent” because they do not have any effect on the amino acid sequence of proteins and on the biological phenotypes. Mutations could cause a harmful, or a beneficial, or no harmful or beneficial effect on the organism. Mutations are stated as a key factor in the evolution and lead to natural selection of the species. The mutations that give advantageous characters to changing environmental conditions are inherited by the offspring, thereby allowing a natural selection to occur [1].

1.2 Causes of mutations

The genetic information stored in the DNA can be altered by various reasons including, environmental factors, chemical mutagens, spontaneously arising changes, and errors that occur during DNA replication and DNA repair [2].

1.2.1 Environmental factors and mutagens

The mutations that are caused by environmental chemicals or radiations are called induced mutations. The following types of chemicals and radiations can cause induced mutations [2].

Chemicals:

1) Base analogs: This class of chemical mutagens contain the structures similar to standard nucleotide bases of DNA.

   Example: 5-Bromouracil (5BU): It is an analog of thymine. 5BU pairs with adenine as thymine, but it occasionally mispairs with guanine.

2) Alkylating agents: These chemicals donate alkyl groups to nucleotide bases.

   Example: Ethyl methanesulfonate (EMS) adds an ethyl group to guanine, producing O^6-ethylguanine, which pairs with thymine.

3) Mutagens that cause deamination

   Example: Nitrous acid deaminates cytosine and adenine, altering their structure, which leads to mispairing of the nucleotide bases.

4) Hydroxylamine: This specific base-modifying mutagen adds hydroxyl group to cytosine and, causes rare tautomer that pairs with adenine instead of guanine.

5) Oxidative species: Reactive oxygen species that are produced during aerobic metabolism, induce mutations.

   Example: Oxidation converts guanine into 8-oxy-7,8-dihydrodeoxyguanine which mispairs with adenine instead of cytosine.

6) Intercalating agents: This class of chemicals intercalate between adjacent bases in DNA, affect the three-dimensional structure of the helix and leads to single nucleotide insertions or deletions.

   Examples: Ethidium bromide, Dioxin

Radiation:

Exposure to ultraviolet radiation (UV) for a prolonged period of time can induce mutations in our skin cells and, cause skin cancer [3]. UV light induces covalent bonds between adjacent pyrimidine bases, produces pyrimidine dimers. In addition to UV light, exposure to ionizing radiation, such as cosmic rays, gamma rays, and x-rays induce double-stranded breaks in DNA. These types of radiation are more harmful as they can penetrate the tissues easily and can cause mutations anywhere in the body.

1.2.2 Spontaneous replication errors

Mutations that occur under normal conditions without any inducer is called a spontaneous mutation. The DNA replication process is highly accurate, however, spontaneous mutations do occur due to following reasons [4].

1) Tautomeric shifts: The change in the position of the hydrogen atom in the purine and pyrimidine bases leads to different chemical forms called tautomers. The presence of rare tautomeric forms of the bases incorporates incorrect base pairing during replication.
2) Depurination: The loss of a purine base creates an apurinic site in the template DNA, resulting in incorporation of an incorrect nucleotide base in the newly synthesized strand opposite to the apurinic site during replication.
3) Deamination: The loss of an amino group of a nucleotide base alters the nucleotide base, resulting incorrect base pairing.

1.2.3 Errors during DNA replication and DNA repair

Errors during DNA replication that are not corrected by DNA repair mechanism, lead to permanent mutation in the DNA. For instance, a mismatched nucleotide base in a newly synthesized strand during replication can serve as a template in the next round of replication. This incorporated error results as a permanent mutation, because the DNA repair system does not detect it, as the base pairing is correct in the next round of replication.

DNA repair mechanisms help to prevent mutations, however, there are instances that the mutations are introduced during this process. For example, Non-homologous end joining (NHEJ) pathway plays an important role in repairing double strand breaks [5]. During NHEJ pathway, few nucleotides are removed to align the two ends of the DNA, which are misaligned, and nucleotides are added to fill in gaps for rejoining. Sometimes, NHEJ introduces mutations as a result of this process.

1.3 Categories of mutations

Mutations are categorized into two groups, such as germline mutations and somatic mutations in the multicellular organisms. Germline mutations occur in cells that produce gametes. This type of mutations is passed to the offspring, and every cell in the developing embryo carries this mutation. Somatic mutations occur in cells of the body that do not produce gametes. This type of mutations is not passed to the offspring, but it is passed to the daughter cells and produce genetically identical cells. Hundreds of somatic mutations could accumulate during the life span, but most mutations have no effect on the phenotype of the organism. However, a somatic mutation which can affect a gene or regulatory element can produce cells with a selective advantage, resulting in preferential growth or survival, which often leads to cancers [6].

1.4 Types of mutations

The mutations can be classified based on different aspects. The following Table 1.1 [1] shows the mutations that are classified based on the molecular nature of the mutation and its effect on protein structure and phenotypic effect [1], [2].

Table 1.1:

Characteristics of different types of mutations [1]

Class of mutation	Type of mutation	Description
Molecular nature	Base substitution	Single nucleotide base is altered
	Insertion	One or more nucleotides are added
	Deletion	One or more nucleotides are removed
	Frame shift mutation	The reading frame is changed because of addition or deletion of one or more nucleotides
Impact on protein structure or function	Missense mutation	Alteration in the nucleotide base produces different amino acid in the protein
	Nonsense mutation	Alteration in the nucleotide base produces a stop codon, resulting in termination of the translation and an incomplete polypeptide
	Silent mutation	Alteration in the nucleotide base doesn't change the amino acid in the protein
	Loss of function mutation	This type of mutation causes complete or partial loss of protein function
	Gain of function mutation	This type of mutation produces a new trait or enhanced protein function
Phenotypic effect	Forward mutation	Alteration of the wild type phenotype
	Reverse mutation	Returns back to wild type phenotype from mutant phenotype

1.5 Links to disease

When the DNA sequence in a protein coding region is changed by mutation, it could cause alteration in the protein sequence that may produce non-functional protein. The correct functioning of thousands of proteins is important for the survival of the cells. A mutation can lead to a disease condition, if it alters the activity of a protein that plays an important role in cell survival. Missense mutations do change the protein sequence, but the biological function of the protein may or may not be affected depending on whether the change is conservative or nonconservative, and the importance of that amino acid on protein function and/or structure. The conservative amino acid substitution is a replacement of an amino acid with similar biochemical properties as the wild type. The nonconservative amino acid substitution result in an amino acid substitution that has different biochemical properties.

1.5.1 Inherited mutations

A mutation in a germ cell can be passed to the offspring and will then occur in all cells of the organism. The hereditary diseases can be caused by mutation in a single gene or multiple genes or by a chromosomal abnormality. DNA damage can be caused by many factors and it is usually repaired by DNA repair system. There are proteins, which recognize and repair the DNA damage to protect the cells from diseases. A mutation in a germ cell within a DNA repair gene could cause high risk of cancer. For example, the germline mutations in the *BRAC1* or *BRAC2* cancer susceptibility genes, associated with Hereditary Breast and Ovarian Cancer (HBOC), have up to an 85% lifetime risk of breast cancer and up to a 46% lifetime risk of ovarian cancer [7]. *BRAC1* and *BRAC2* are tumor suppressor genes and play an important role in DNA repair of DNA double-strand breaks [8].

1.5.2 Role in carcinogenesis

A mutation could occur in a somatic cell and it will be passed to the daughter cells when the cell divides. The accumulation of somatic cell mutations throughout life time can be a cause of malignant transformation [6]. The mutations, which can produce a selective growth advantage, often leading to cancer development, are termed as driver mutations. When the mutations do not produce phenotypic consequences or biological effects, they are termed as passenger mutations [9].

1.6 p53 mutations in cancer

The progression of a normal cell to a cancerous cell includes several genetic alterations and multiple genes are identified as contributing factors for tumorigenesis. p53 is a key tumor suppressor that regulates various signaling pathways [10]. In response to genotoxic stresses, p53 activates several genes that are involved in arresting cell cycle, repairing damaged DNA and inducing apoptosis [11]. The *TP53* (tumor protein p53) gene is located on chromosome 17 in the p13.1 arm [12]. p53 is a 393 amino acid long protein, which consists of 4 main domains: the intrinsically disordered N-terminal transactivation domain (TAD) (M1-P60); the proline-rich domain (P-rich) (D61-S94); the DNA binding domain (DBD) (T102-K292); and the C-terminal negative regulatory domain (CTD/NRD) (P301-D393), which includes the tetramerization domain (TD) (D324-A355) (Figure 1.1) [13]. p53 protein level and its activity in the cell are tightly regulated by post-translational modifications (PTMs) and interactions with several signaling proteins [14], [15].

The transactivation domain contains two sub-regions: AD1 helix; and AD2 helix, which are crucial for binding to key regulatory proteins (Figure 1.1) [16]. The human double minute 2 (HDM2) protein acts as a negative regulator of p53 and promotes degradation of p53 by poly-

ubiquitination of the C-terminus in the absence of cellular stress conditions [17]. When the cells are under genotoxic stress conditions, phosphorylation of p53-TAD recruits Histone acetyltransferases (HATs); cyclic-AMP response element-binding (CREB)-binding protein (CBP) and its paralog p300, which act as co-activators and promotes acetylation of the C-terminus, which leads to stabilization of p53 and inhibits p53 degradation [18]. The activation of p53 leads to tetrameric oligomerization state that controls growth arrest and apoptosis by inducing the expression of target genes [19]. The target genes that are activated span a multitude of biological processes, including DNA Damage Repair (DDR, DNA damage-binding protein-2 (*DDB2*)), cell cycle arrest (cyclin dependent kinase inhibitor 1, *CDKN1A*), apoptosis (*PUMA* and *BAX*), metabolism (TP53-induced glycolysis and apoptosis regulator, *TIGAR*) and post-translational regulators of p53 (*HDM2* and p53 induced phosphatase 1) [20].

TP53 is a highly mutated gene in human cancers and somatic mutations are the most common type of mutations linked to cancer [21]. The International Agency for Research on Cancer (IARC) TP53 mutation database maintains a list of somatic mutations and germline mutations, and it shows that around 97% of the mutations are located in the DNA binding domain and most of these are missense mutations [22], [23]. An example of germline mutations in *TP53* are those associated with Li-Fraumeni and Li-Fraumeni-like syndromes (LFS, LFL, respectively), which predispose to several types of early-onset cancers [24].

The tumor suppressor genes are mostly inactivated through frameshift or nonsense mutations, but *TP53* mutations are mostly missense, meaning that they are caused by single amino acid substitutions [25]. Functional analysis of p53 mutations from several systemic studies have found that some mutants produce a complete loss of p53 function while others lead to altered transactivation behaviors or retain function at lower temperatures [11], [26]–[28]. The

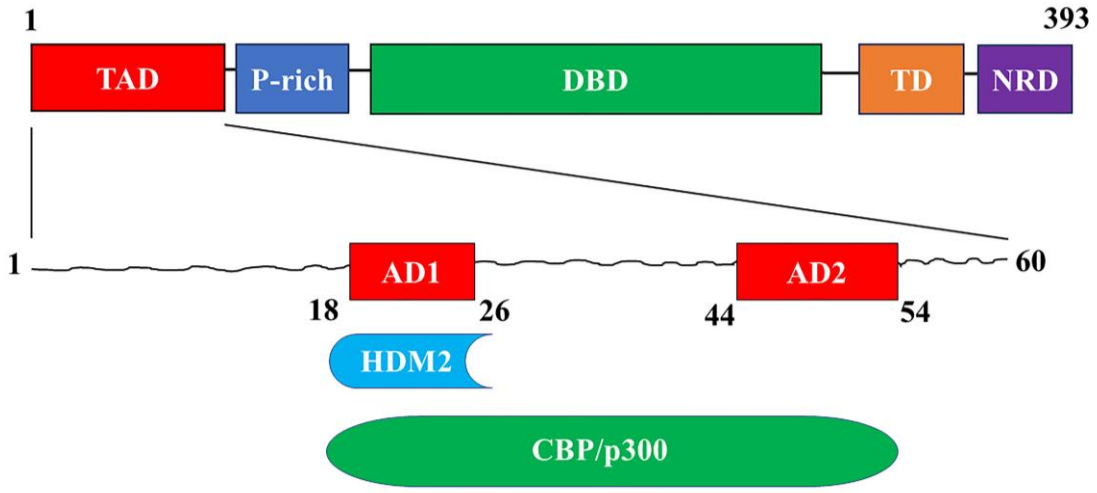
crystal structure of the p53 DBD in complex with a consensus DNA was solved in 1994, which paved the way for a better understanding of the effects of cancer mutations [29]. The most commonly mutated amino acid residues in p53 DBD are classified as 'contact' (Arg-248 and Arg-273) or 'structural' residues (Arg-175, Gly-245, Arg-249 and Arg-282), based on whether the mutations lead to a loss of DNA contact or alter the structure of the DND-binding surface [29].

Since previous studies mainly focused on mutations in DBD, there is a deficit in understanding the impacts of the cancer-linked mutations in TAD and other regulatory domains. The p53-TAD is an intrinsically disordered region (IDR), which adopts unique conformations only when interacting with binding partners (HDM2, CBP/p300) that control cellular responses to genotoxic stresses [30]. Although the prevalence of cancer-linked mutations in TAD is lower than in DBD [31], a number of TAD mutations are frequently associated with some cancers [32]. The E17D and K24N mutation in p53-TAD are documented in female genital cancers and some other TAD mutations are linked to the tumors in nasal cavity, tonsil, salivary gland and parotid gland [31]. All possible point mutations in *TP53* gene comes from a yeast-based functional assay, which reported the functionality of the currently known TAD cancer mutants [11]. However, the structural features and molecular interactions of TAD cancer mutants have not been studied extensively.

Most mutations in *TP53* inactivate p53 and exhibit defects in the p53 signaling pathway [33]. Clinical studies have shown that the type of p53 mutation can be connected to cancer prognosis, response to treatment, and patient survival [34]. Therefore, it is important to understand the underlying mechanism of p53 inactivation by missense mutations to assess the clinical outcome of potential treatment strategies.

Figure 1.1:

p53 domain structure and p53-TAD binding proteins



TAD: Transactivation domain; P-rich: Proline-rich domain; DBD: DNA-binding domain; TD: Tetramerization domain; NRD: Negative regulatory domain ; AD1, AD2: Short helical region of p53 TAD; HDM2: Human double minute 2; CBP/p300: CREB-binding protein and its paralog p300

1.7 Human CLPB/SKD3 mutations and the associated diseases

Human caseinolytic peptidase B protein homolog (CLPB)/suppressor of potassium transport defect 3 (SKD3) belongs to the AAA+ family of ATPases associated with different activities [35]. The biological function of this protein is unknown. The *CLPB* gene is located on the chromosome 11q13 and it is expressed in various human tissues [36]. Importantly, mutations in *CLPB* gene are associated with 3-methylglutaconic aciduria, type VII (MGCA7) with cataracts, neurological involvement, and neutropenia (MEGCANN) [36]–[41]. MGCA7 is an autosomal recessive metabolic disorder affecting young children [42]. In its severe form, the disease causes death at a few months of age as a result of significant neonatal neurological symptoms (absence of voluntary movements, hypotonia or hypertonia, swallowing problems, respiratory problems, and epilepsy) and severe neutropenia associated with life-threatening infections [42]. In moderate form of this disease, patients may have neurologic abnormalities in infancy and can develop spasticity, progressive movement disorder, epilepsy, and intellectual disability in later childhood [42]. The patients with mild severity of this disease have normal life expectancy, no neurologic involvement, but with mild neutropenia present [42]. The clinical manifestations of the disease-linked mutations are currently challenging to explain due to the lack of understanding of the biological function of CLPB. The CLPB disease-linked mutations that have been identified in clinical studies are listed in Table 1.2.

Human CLPB protein contains a single AAA+ module that is characteristic for all AAA+ proteins. The AAA+ domain in the C-terminal part of CLPB, shares similarity with the C-terminal AAA+ domain of the bacterial ClpB, yeast cytosolic Hsp104 and yeast mitochondrial Hsp78 family of proteins [35]. Unlike ClpB/Hsp104, human CLPB is not found in microorganisms, but instead it is found in vertebrates, including humans. The structure of the

human CLPB protein differs significantly from the bacterial ClpB/Hsp104/Hsp78 proteins by the presence of an ankyrin-repeat domain in the N-terminal region (Figure 1.2) [38]. Human CLPB contains a predicted N-terminal mitochondrial targeting presequence (MTS), followed by an ankyrin-repeat domain (ANK), an AAA+ nucleotide-binding domain (NBD), and a small C-terminal domain (CTD) (Figure 1.2). Bacterial ClpB, yeast cytosolic Hsp104, and yeast mitochondrial Hsp78 are well characterized molecular chaperones that cooperate with Hsp70 and Hsp40 during reactivation of aggregated proteins under severe cellular stresses [43]. Because of the similarity between the AAA+ domains of human CLPB and microbial ClpB/Hsp104/Hsp78, it has been postulated that human CLPB may also function as a molecular chaperone/disaggregase [44]. The microbial ClpBs associate into hexameric rings with a central pore, where the substrate proteins are translocated and unfolded, which is driven by ATP binding and hydrolysis [45], [46]. However, it is not clear whether human CLPB, with its single AAA domain, could engage in stable nucleotide-dependent protein complexes in cells.

Upregulation of the CLPB expression is observed in human acute myeloid leukemia (AML) upon acquisition of resistance to venetoclax, a selective BCL2 inhibitor and an FDA-approved drug for the treatment of AML [47]. Ablation of CLPB sensitizes AML cells to venetoclax [47]. It has been shown that CLPB interacts with the mitochondrial cristae-shaping protein OPA1 and maintains the mitochondrial cristae structure [47]. The loss of CLPB promotes apoptosis by induction of the mitochondrial stress response and cristae remodeling. The study further demonstrated CLPB interacts with the anti-apoptotic factor HAX1 that prevents mitochondrial-mediated apoptosis [47]. CLPB was also recently reported to bind with prohibitins complex (PHB) that are involved in antiviral innate immunity in mitochondria [48]. In spite of these published results, the biological function of CLPB in the cell remains elusive. Currently,

correlations are not evident between a severity of the disease and the mutation type or location. Therefore, understanding the biological role of human CLPB may shed new light in the clinical manifestations of the disease mutations that are currently challenging to explain.

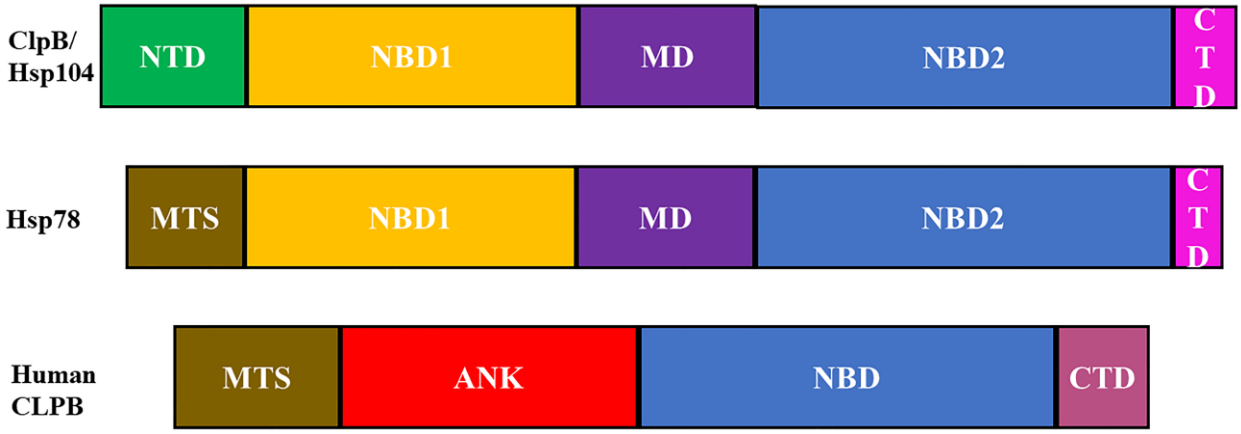
Table 1.2:

CLPB disease-linked mutations [36]–[41]

Mutation	Mutation type	Location of the mutation	Clinical severity
R250*	Nonsense	ANK	Not available
T268M	Missense	ANK	Moderate
A269T	Missense	ANK	Not available
Y272C	Missense	ANK	Severe
K321*	Nonsense	ANK	Severe
R408G	Missense	NBD	Moderate
M411I	Missense	NBD	Mild
R417*	Nonsense	NBD	Moderate
E435_G436delinsDP	Missense	NBD	Moderate
C486R	Missense	NBD	Moderate
E501K	Missense	NBD	Severe
Y567C	Missense	NBD	Severe
A591V	Missense	NBD	Severe
Y617C	Missense	NBD	Mild
R628C	Missense	NBD	Mild
E639K	Missense	NBD	Mild
G646V	Missense	NBD	Moderate
C647Lfs*26	Frameshift	NBD	Severe
I682N	Nonsense	CTD	Severe

Figure 1.2:

Domain map of microbial ClpB/HSP104, HSP78 and Human CLPB



NTD: N-terminal domain; NBD1: Nucleotide-binding domain 1; MD: Middle domain: NBD2: Nucleotide-binding domain 2; CTD: C-terminal domain; MTS: Mitochondrial targeting presequence; ANK: Ankyrin-repeat domain; NBD: Nucleotide-binding domain; CTD: C-terminal domain

Chapter 2 - Cancer-associated mutations perturb the structure and interactions of the intrinsically disordered p53 transactivation domain

Lynn G. Schrag[#], Indhujah Thevarajan[#], Xiaorong Liu[#], Om Prakash, Michal Zolkiewski and Jianhan Chen

[#] Equal Contributions

Part of the manuscript in preparation for journal submission

2.1 Abstract

Intrinsically disordered proteins (IDPs) are key components of regulatory networks that control crucial aspects of cell decision making. The transactivation domain (TAD) of tumor suppressor p53 mediates interactions with the regulatory pathways to control the p53 homeostasis during the cellular response to genotoxic stress. Many cancer-associated mutations have been discovered in p53-TAD, but their structural and functional consequences are poorly understood. In this work, we combine extensive atomistic simulations, NMR spectroscopy, and protein-protein interaction assays and demonstrate that cancer-associated mutations can significantly perturb the balance of p53 interactions with key activation and degradation regulators. Importantly, some mutations do not directly disrupt the known interaction interfaces, but likely modulate the disordered state of p53-TAD to affect the interaction. Our work suggests that the disordered conformational ensemble of p53-TAD can serve as a central conduit in regulating the response to various cellular stimuli at the protein-protein interaction level. Understanding how the disordered state of IDPs may be modulated by regulatory signals and/or

disease associated perturbations will be essential in the studies on the role of IDPs in biology and diseases.

2.2 Introduction

Intrinsically disordered proteins (IDPs), which lack stable 3D structures under physiological conditions, have challenged the traditional protein structure-function paradigm [49]–[53]. IDPs tend to contain low-complexity amino-acid sequences with poor conservation among orthologs, are enriched with charged residues, and display low overall hydrophobicity [54], [55]. Importantly, IDPs do not usually exhibit fully random conformations in the absence of binding partners [56]. Frequently, they display a heterogeneous distribution of partially folded but highly mobile conformations [57]. The inherent thermodynamic instability of IDP conformations arguably provides a robust mechanism for allosteric regulation [58]–[62], by allowing the disordered ensemble to efficiently respond to various cellular signals. Combinations of post-translational modifications, ligand or protein binding, changes in environmental conditions (pH, ionic strength, etc.), and mechanical stress are common regulators of IDP conformational ensembles and interaction interfaces [30], [63]–[67]. Cooperative responses of IDPs, such as through binding-induced folding [68], could naturally integrate multiple signal pathways, thus allowing them to function as hubs within cellular signaling networks [54]. Indeed, IDPs have been estimated to account for nearly half of signaling-associated proteins in eukaryotic cells [69], and nearly all eukaryotic transcription factors contain one or more disordered regions [70]. It is therefore unsurprising that many human diseases, including neurodegeneration, cancer, and diabetes have been linked to mis-function of IDPs [71]–[74]. There is thus an important need and great interest in understanding how intrinsic disorder mediates versatile biological functions and how such mechanisms may fail in human diseases.

A prominent example of biologically important IDPs is the transactivation domain (TAD) of the tumor suppressor and transcription factor p53. As a critical integrator of cellular responses against genotoxic stress, p53 is the most frequently mutated protein associated with human cancers [75]–[77]. p53 is a multidomain protein comprised of the TAD, proline-rich domain (P-rich), DNA-binding domain (DBD), tetramerization domain (TD), and negative regulatory domain (NRD) (Figure 2.1(A)). Only the DBD is stably folded in the unbound state of p53. In healthy cells, p53 is negatively regulated via targeting for degradation by E3 ubiquitin ligase human double minute 2 (HDM2), which binds p53-TAD and promotes poly-ubiquitination of the p53 NRD [78], [79] (Figure 2.1(B)). Upon prolonged exposure to genotoxic stressors, p53 accumulates post-translational modifications at an array of positions throughout the molecule [20], [80]–[86]. Key phosphorylation events within TAD including those at Ser15, Thr18, and Ser20 destabilize the p53-TAD:HDM2 complex, inhibit the HDM2-mediated poly-ubiquitination of p53, and prevent the degradation of p53 [87]–[90]. Moreover, the phosphorylation of p53-TAD recruits the histone acetyl transferases and the transcriptional co-activators, cyclic-AMP response element-binding (CREB)-binding protein (CBP) and its paralog p300, to p53-TAD [16], [18], [81], [91]–[94]. These interactions enable site-specific acetylation of p53 NRD, which promotes p53 nuclear localization and further stabilizes the complex with CBP/p300 [95], [96] (Figure 2.1(B)). These processes subsequently lead to activation of p53-regulated apoptotic genes [97]–[99].

The molecular basis of how multi-site phosphorylation modulates the balance of p53 interactions with the degradation and activation pathways is not fully understood. Many phosphorylation sites do not locate at and near the known interaction sites of p53-TAD and thus do not directly disrupt protein-protein interactions [32]. Alternatively, phosphorylation may

modulate the unbound conformational properties of p53-TAD to affect its interactions. Specifically, an overall shape, transient secondary structures and/or long-range contacts in p53-TAD may be perturbed, which could alter the availability of key binding sites and/or change the entropic cost of binding to various regulators. Such conformational modulation mechanisms have been demonstrated in previous studies of CREB/CBP [100], Connexin43 and 4E-BP2 [101], [102], as well as interaction of IDPs with several drug molecules [103]–[107].

Importantly, although cancer-associated mutations in p53 are predominantly located within its DBD, a number of clinically relevant residue substitutions occur within the TAD [31]. Many of these cancer-associated mutations locate outside of the known interaction sites of p53-TAD. It is plausible that these mutations could also modulate the disordered ensemble of p53-TAD to perturb the balance of its interactions and further interfere with phosphorylation regulation of the p53 homeostasis in cellular response to genotoxic stresses. Testing these hypotheses is challenging and requires a detailed analysis of the disordered ensemble as well as dynamics of p53-TAD. p53-TAD has been characterized extensively using various experimental methods, including NMR spectroscopy [108]–[112], small-angle X-ray scattering (SAXS) [108], and single molecule fluorescence resonance energy transfer (smFRET) [113]. These studies show that unbound p53-TAD contains partial helical structures in the regions AD1 and AD2 (Figure 2.1) [108], [109], with the AD1 region estimated to be up to ~30% helical. The overall TAD shape appears to be extended with a Stokes radius of ~2.38 nm (residues 1-73) according to size exclusion chromatography [108], [112], [113]. Transient long-range contacts have also been observed within the peptide [110], [114]. Importantly, these experimental approaches measure ensemble-averaged properties and do not directly reveal the full range of potentially heterogeneous conformational states sampled by TAD. For this, molecular dynamics (MD)

simulations employing physics-based force fields could provide the much-needed details of the disordered ensemble, which is necessary for understanding the molecular basis of their functions [62]. This challenging task has become increasingly feasible in recent years, thanks to a continuous development of more accurate protein force fields [115]–[119], highly efficient GPU-enabled MD algorithms [120]–[126], and advanced conformational sampling algorithms [127]–[138]. The simulated ensembles could then be properly validated by direct comparison with experimental observables. Combining information from atomistic simulation and biophysical characterization can thus provide reliable molecular details for the mechanistic links between the disordered conformational ensemble and protein function, and for revealing how such mechanisms may fail in human diseases.

In this study, we investigated four variants of p53-TAD, each with one or two amino-acid substitution(s) found in human cancers [31], [139]–[141]: K24N, N29K/N30D, D49Y, and W53G. These mutants have been associated with female genital, brain, bladder or breast cancers [31], and were previously predicted to substantially modulate the level of disorder in unbound p53-TAD [32]. By combining enhanced sampling with total aggregated atomistic simulation time of over 240 μs, NMR spectroscopy, and protein-protein interaction assays, we assessed functional consequences and conformational ensemble changes in p53-TAD resulting from the above missense mutations. We discovered that while the TAD structure and dynamics were only marginally affected by the cancer-associated mutations, these mutations had profound effects upon the TAD interactions with the partner domains of its key regulators, HDM2 and CBP. Our results demonstrate that subtle changes in the conformational ensemble of an IDP may produce distinct molecular responses and highlight that the conformational ensemble of unstructured domains can serve as a central conduit in regulating the response to various cellular stimuli.

2.3 Materials and methods

2.3.1 Expression, mutagenesis and purification of p53-TAD (1-73)

A pET28a vector with sub-cloned cDNA of *Hs*p53-(1-73) (72R) was procured from Addgene (Addgene plasmid #62082; http://n2t.net/addgene:62082); deposition courtesy of Dr. Gary Daughdrill, University of South Florida. The p53-(1-73) plasmid was then transformed into DH10β cells (New England Biolabs) to generate a plasmid stock in preparation of p53-TAD variants. p53/pET28a plasmid was then templated to create the desired mutations in p53-TAD through primer mismatch using Agilent Technologies Quickchange Lightning Site-Directed Mutagenesis kit:

K24N: (FP: 5'-TCAGAVCTATGGAATCTACTTCCT-3'; RP: 5'-AGGAAGTAGATTCCATAGGTCTGA-3'),

N29K/N30D: (FP: 5'-GGAAACTACTTCCTGAAAAAGACGTTCTGTCCCCCTTGCC-3'; RP: 5'-GGCAAGGGGGACAGAACGTCTTTTTCAGGAAGTAGTTTCC-3'),

D49Y: (FP: 5'-CTGTCCCCGGACTATATTGAA-3'; RP: 5'-TTCAATATCGTCCGGGGACAG-3'), and

W53G: (FP: 5'-CGGACGATATTGAACAAGGCTTCACTGAAGACCCAGGTCC-3'; RP: 5'-GGACCTGGGTCTTCAGTGAAGCCTTGTTCAATATCGTCCG-3')

The p53-TAD protein was purified using previously described protocol [142], with some modifications. p53-TAD variants were transformed into BL21(DE3) low background strain (LOBSTR) chemically competent cells. The transformed cells were selected using lysogeny broth (LB) agar plates with 50 µg/mL Kanamycin and 1% (w/v) glucose. Expression of p53-TAD for Biolayer Interferometry (BLI) was accomplished using LB media. p53-TAD wild type

(WT) and variants were expressed using LB media were grown at 30°C in 4 L cultures of LB with Kanamycin and 1% glucose. 1 mM Isopropyl-β-D-thiogalactoside (IPTG) was added when cultures reached an OD_{600}= 0.5-0.7. After IPTG induction, the cultures were grown at 37°C for an additional 6hrs. Purification of p53-TAD from the clarified cell lysate was accomplished using a Nickel Nitrilotriacetic Acid (Ni-NTA) column using 50 mM Tris-HCl with 300 mM NaCl at pH=8.0. Further, Tris-HCl buffers containing 10 mM Imidazole for wash or 250 mM Imidazole for elution were used. Elution fractions were subjected to G75 Superdex gel filtration column chromatography at 4°C using 50 mM NaH_2PO_4 buffer with 300 mM NaCl and 1 mM EDTA at pH=7.0. The p53-TAD fractions collected from gel filtration chromatography were subjected to second round of Ni-NTA purification. Elution fractions of p53-TAD were then concentrated using Amicon Ultra-15 (3.5KDa MWCO) centrifugal filters. Final concentrates of p53-TAD were then dialyzed overnight at 4°C in a Slide-A-Lyzer dialysis cassette (3.5KDa MWCO) into a 20 mM Tris-HCl buffer with 50 mM NaCl and 10% glycerol at pH=7.5 prior to storage at -20°C.

2.3.2 Expression and purification of HDM2 (17-125)

A pGEX-6p-2 vector with subcloned cDNA of HDM2 (17-125) was procured from Addgene (Addgene # 62063; http://n2t.net/addgene:62063); deposition courtesy of Dr. Gary Daughdrill, University of South Florida. The pGEX-6p-2-HDM2(17-125) plasmid was transformed into DH10-β chemically competent cells (New England Biolabs) for plasmid stock preparation. The plasmid was then transformed into Rosetta BL21(DE3) chemically competent cells. The HDM2 (17-125) protein was purified according to previously published protocol [143], with some modifications. Transformed cells were selected for expression on LB agar plates in the presence of 25 µg/mL Chloramphenicol and 100 µg/mL of Ampicillin. HDM2 was

expressed using LB media and grown at 37°C with Chloramphenicol and Ampicillin to an OD_{600} = 0.8. 1 mM Isopropyl-β-D-thiogalactoside (IPTG) was added to the culture and cells were grown at 25°C for an additional 5hrs. The culture was centrifuged at 5000 rpm for 30 mins and pellets were resuspended in GST binding buffer (50 mM Tris-HCl, 300 mM NaCl, 2.5 mM EDTA, 0.02% NaN_3, 2 mM DTT, pH=7.4). The cells were lysed using French press and centrifuged at 13,000 rpm for 40 mins. HDM2 containing supernatant was loaded onto a Glutathione Sepharose 4B resin (GE Healthcare) column that had been already equilibrated with GST binding buffer. The HDM2 protein fractions were eluted with GST binding buffer with 10 mM reduced glutathione. The GST tag was removed using HRV 3C protease (Thermo Scientific) with an enzyme to substrate ratio of 1:100 at 4°C for 16 hrs. After the removal of the GST tag the HDM2 protein was concentrated using Amicon Ultra-15 (3.5 KDa MWCO) centrifugal filters. Finally, the protein was buffer exchanged into 20 mM Tris-HCl buffer with 150 mM NaCl, 1mM DTT, and 10% glycerol at pH 8.0 prior to storage at -20°C.

2.3.3 Expression and purification of TAZ2 domain of CBP (performed by Dr. Lynn Schrag)

A pET22b recombinant plasmid coding a bi-cistronic mRNA containing 6His-Gb1-STAT1 (710-750) and CBP (1764-1850) TAZ2 was provided as a gift from Dr. Peter Wright (Addgene plasmid #99342; http://n2t.net/addgene:99342). The pET22b/STAT1/TAZ2 encodes a second ribosomal binding site at the start of the TAZ2 sequence. This enables concurrent expression of TAZ2 and binding partner STAT1. pET22b/STAT1/TAZ2 was transformed into BL21(DE3) Rosetta chemically competent cells. Transformed cells were selected for expression on LB agar plates with 1% glucose in the presence of 25 µg/mL Chloramphenicol and 100 µg/mL of Ampicillin.

STAT1/TAZ2 expression was accomplished using 2 L volume of LB with 1% glucose, 25 µg/mL Chloramphenicol and 100 µg/mL Ampicillin via shaker at 37°C. Once the culture OD_{600}=0.1-0.3, approximately 250 µM of $ZnCl_2$ was introduced to facilitate stability of expressed TAZ2. Induction of culture expression with 1 mM IPTG was accomplished when OD_{600}=0.5-0.7. Cultures were then allowed to express for an additional 3 hrs at 37°C. Following pelleting and cell lysis STAT1/TAZ2 was resuspended in 20 mM Tris-HCl buffer with 300 mM NaCl and 10 mM Imidazole at pH=8.0. The clarified cell lysate was then passed over Ni-NTA column and elution fractions containing STAT1/TAZ2 were collected. The STAT1/TAZ2 fractions were applied to a sulfopropyl (SP) sepharose cation exchange column with a flow rate of 0.3 mL/min. TAZ2 was independently eluted using a linear salt gradient from 50 mM to 1M NaCl in a 20 mM Tris-HCl buffer with 10 mM DTT at pH=6.5. Purified TAZ2 fractions were quickly concentrated and dialyzed into a 25 mM Na-Acetate buffer with 25 mM NaCl, 0.1 mM DTT, and 0.1 mM $ZnCl_2$ at pH=5.8. Appropriate folding of TAZ2 was confirmed through Circular Dichroism (CD) spectroscopy prior to storage at 4°C.

2.3.4 Expression and purification of TAZ1 domain of CBP (performed by Dr. Lynn Schrag)

pET21a plasmid containing CBP-(340-439) TAZ1 was given as a gift by Dr. Peter Wright's lab. The transformation, selection, and expression of pET21a/TAZ1 was carried as outlined for STAT1/TAZ2. TAZ1 containing lysate was suspended in 20 mM Tris-HCl buffer with 50 mM NaCl and 10 mM DTT at pH=6.9. Further isolation of TAZ1 was carried using SP-sepharose linear gradient procedure as outline for TAZ2. TAZ1 containing eluent was then concentrated and injected onto a G75 superdex column using a 20 mM Tris-HCl buffer with 300 mM NaCl and 10 mM DTT at pH=6.8. Fractions containing TAZ1 were concentrated and

dialyzed into 20 mM Tris-HCl buffer with 50 mM NaCl and 1 mM DTT at pH=6.8. TAZ1 folding, as with TAZ2, was assessed using CD spectroscopy prior to storage at 4°C.

2.3.5 Biolayer interferometry (BLI) assays of p53-TAD variants to HDM2 (17-125), CBP-TAZ1 and CBP-TAZ2

Aliquots of HDM2 (17-125), TAZ1, and TAZ2 from storage buffer were diluted into freshly generated 20 mM Tris-HCl buffer with 50 mM NaCl and 1 mM DTT at pH=7.5 (pH=6.8 for TAZ1) to appropriate target concentrations (HDM2 (17-125) : 2 µM, 3µM, and 4 µM; TAZ1: 20 µM, 30 µM, and 40 µM; TAZ2: 300 nM, 600 nM and 900 nM) for analysis of binding kinetics. The buffer for binding kinetics studies was prepared with a composition of 20 mM Tris-HCl, 50 mM NaCl, 0.01% (W/V) BSA, and 0.002% (V/V) Tween-20. p53-TAD WT and its variants were prepared via dilution into binding kinetics buffer solution to generate a final concentration of 100 ng/µL of p53-TAD. All samples were stored on ice until used. 200 µL aliquots of the p53-TAD samples were transferred into 500 µL black eppendorf tubes. These samples were equilibrated to room temperature 10 minutes prior to their use. Ni-NTA conjugated BLItz biosensors were hydrated for ~10 minutes using binding kinetics buffer prior to data acquisition. Each experiment was conducted using BLItz Pro 1.2 software and the instrument operated at an oscillation frequency of 2200 Hz for the duration of each acquisition. The scheme for data acquisition was conducted as follows: Step 1 (baseline 1), the tip is equilibrated an additional 30 s with kinetics buffer and instrument self-calibrates; Step 2 (analyte binding), p53-TAD variant or control sample was applied to the tip for 180 s; Step 3 (baseline 2), tip is exposed to empty kinetics buffer for 120 s; Step 4 (association), HDM2 (17-125), CBP-TAZ1, CBP-TAZ2 dilutions in 20 mM Tris-HCl, 50 mM NaCl, and 1 mM DTT at pH=7.5 is applied for 300 s (120 s for TAZ1 acquisitions); And step 5 (dissociation), tip is exposed to kinetics buffer for

300 s (120 s for TAZ1 acquisitions). After analyzing the data, the figure was generated using GraphPad Prism 8.1.0 software to approximate effective on-rate (k_{on}), off-rate (k_{off}), and K_D values.

2.4 Results

We examined the functional impact of four cancer-associated mutations, K24N, N29K/N30D, D49Y, and W53G by studying their interactions with the partner domains of the key regulators, HDM2 and CBP. For this study, p53-TAD protein was purified using previously described protocol [142], with some modifications. A plasmid construct with p53-TAD (residues 1-61) was used initially to produce the protein. Unfortunately, we were not able to purify the protein and we found that p53-TAD (residues 1-61) was degraded at the final stage of purification after the thrombin cleavage (results not shown). Next, we obtained the p53-TAD construct which contained part of the proline-rich domain (residues 62-73) together with the TAD domain (residues 1-61). The pET28a vector with a sub-cloned cDNA of *Hs*p53-(1-73) (72R) was procured from Addgene. In addition to the p53 residues 1-73, the construct contained an N-terminal 6-histidine tag followed by a thrombin cleavage site.

The p53-TAD protein was initially expressed in BL21 (DE3) *E. coli* cells. The protein purity was not high enough because some of the endogenous histidine-rich *E. coli* proteins co-purified along with the p53-TAD protein. This was due to a relatively low level of p53-TAD expression in *E. coli*. To overcome this issue, we used BL21(DE3) low background strain (LOBSTR) cells to express the p53-TAD protein. LOBSTR is an engineered *E. coli* expression strain, which eliminates the most abundant histidine-rich contaminants [144]. Figure 2.2 shows the p53-TAD WT expression in LOBSTR cells and the following purification steps as described above in Materials and methods (section 2.3.1). Coomassie-stained gels detected expression of a

protein with an apparent molecular weight of ~18 kDa, which was higher than the predicted molecular weight of p53-TAD (1-73) (Figure 2.2(A)). The identity of the p53-TAD protein was confirmed by mass spectrometry. The N-terminal 6His-extension was not removed in order to preserve the protein's affinity towards nickel (see BLI studies, below). As shown in Figure 2.2, we purified P53-TAD WT to near homogeneity and mutant proteins with high purity. The same procedure was used to produce 4 variants of p53-TAD with cancer-linked mutations, K24N, N29K/N30D, D49Y, and W53G.

Next, we purified the recombinant N-terminal region of the E3 ubiquitin ligase HDM2 (residues 17-125), which is sufficient for binding to p53-TAD [143], using a construct, containing an N-terminal GST tag. A pGEX-6p-2 vector with subcloned cDNA of HDM2 (17-125) was procured from Addgene and the protein was purified according to previously published protocols [143], with some modifications. The expression and purification steps of HDM2 (17-125) are shown in Figure 2.3. After cleaving off the GST tag, we successfully produced a high amount (~ 0.5 g) of highly pure HDM2 (17-125) protein (see Figure 2.3(D)).

Next, we purified the p53-TAD interacting partner domains of CBP. CBP contains four domains, TAZ1, KIX, TAZ2, and NCBD that all interact with p53-TAD, but binding affinities for TAZ1 and TAZ2 are significantly higher than those for KIX and NCBD [16]. Therefore, we focused on the interactions of the p53-TAD variants with TAZ1 and TAZ2 domains because they could best reflect how CBP regulation of p53 may be impacted by the cancer-linked mutations. After successful expression of CBP-TAZ1 and CBP-TAZ2 in *E. coli* and purification, the folding and structural response to Zn^{2+} of CBP-TAZ1 and CBP-TAZ2 was confirmed via circular dichroism (CD) analysis [145], [146]. CBP-TAZ1 (340-439) and CBP-TAZ2 (1764-1850)

proteins were purified by Dr. Lynn Schrag as described in the Materials and methods (sections 2.3.3 – 2.3.4).

To investigate protein-protein interactions mediated by WT p53-TAD and its variants, we conducted a series of *in vitro* protein binding assays utilizing biolayer interferometry (BLI) with the p53-TAD variants non-covalently immobilized on Ni-NTA biosensor tips. To establish an appropriate concentration range for collection of binding isotherms, the p53 immobilization was performed at a concentration ~10-fold higher than the previously published K_D values for interactions of wild-type (WT) p53-TAD with the selected partner domains. The immobilized p53-TAD variants were then exposed, with 2 replicates, to 3 unique concentrations for each interaction partner and the BLI responses were recorded (see Materials and methods, section 2.3.5). Using Blitz Pro1.2 software, a global non-linear least-squares fitting of all 6 measurements for each interaction was used to calculate the association (k_{on}) and dissociation (k_{off}) rate constants (Figure 2.4 and Table 2.1). The k_{on} and k_{off} values for p53-TAD WT indicate fast association/dissociation kinetics for binding of all three partners, in agreement with the previous NMR observations [16]. Importantly, the calculated K_D values from BLI (Table 2.1) reproduce the thermodynamics for p53-TAD WT interactions with HDM2, TAZ1, and TAZ2, as previously determined using isothermal titration calorimetry (ITC) and NMR [16], [92], [94]. For example, the K_D value of p53-TAD WT-HDM2 interaction determined from BLI (260 ± 61 nM) is in agreement with the values of 230 ± 20 nM [16] or 200 ± 20 nM [94] from ITC experiments, despite slight differences in the length of the p53-TAD constructs and solution conditions used in different studies. Among the selected mutants, only the p53-TAD K24N-HDM2 interaction has been studied before, which showed that the K24N substitution did not significantly affect the binding affinity [111], as confirmed by the BLI results (Table 2.1).

As summarized in Table 2.1 and Figure 2.4, the four selected cancer-associated mutations in p53-TAD induce significant and diverse changes in the binding affinities towards the regulatory partners. Specifically, the K24N substitution in p53-TAD leads to a significant reduction of binding affinity for TAZ2, but does not affect interactions with TAZ1 or HDM2. N29K/N30D abolishes p53-TAD binding to HDM2 and TAZ2, but does not affect interactions with TAZ1. Both D49Y and W53G lose affinity towards TAZ2, but not towards HDM2. In addition, W53G, but not D49Y, reduces the affinity towards TAZ1. Taken together, these results suggest that mutations in the disordered p53-TAD could significantly perturb the regulatory network of p53, thus affecting its tumor suppressor function.

2.5 Discussion

The recombinant expression of specific domains of human proteins in bacteria is challenging. Protein expression protocols typically need to be optimized to obtain sufficient protein levels with high purity. We noticed that p53-TAD mutants have different level of expression in bacteria. This may be due to different levels of cytotoxicity of those proteins, which could arise from harmful interactions of human p53-TAD with bacterial proteins. We used the BL21(DE3) low background strain (LOBSTR) cells for p53-TAD expression, which eliminated the major histidine-rich bacterial protein contaminants [144]. We used different protein purification methods to obtain the p53-TAD (1-73; 72R), HDM2 (17-125), CBP-TAZ1 (340-439), and CBP-TAZ2 (1764-1850) high purity preparations for protein binding studies using Biolayer interferometry (BLI).

The BLI studies were performed to investigate the impact of four cancer-associated mutations in the intrinsically disordered p53-TAD on interactions with the partner domains. Notably, all the selected cancer-linked mutations decreased the p53-TAD affinity towards TAZ2,

including three mutations, N29D/N30D, D49Y, and W53G, for which the TAD-TAZ2 interaction was below the limit of detection (see Figure 2.4 and Table 2.1). In contrast, only one mutated variant, N29D/N30D showed a defect in binding to HDM2 and only W53G showed a weaker binding to TAZ1, as compared to p53-TAD WT.

Biologically active p53 forms a tetramer whose four TADs can simultaneously interact with four CBP domains, TAZ1, TAZ2, KIX, and NCBD [147]. The K_D value for the p53-TAD interaction with TAZ1 is ~30-fold higher than for the interaction with TAZ2 (see Table 2.1) and the affinity of p53-TAD towards KIX and NCBD is even weaker than towards TAZ1 [16]. Thus, among the four CBP domains, the dominant interaction, which regulates the status of the p53 activation is provided by the TAD-TAZ2 interface. Our results indicate that all four selected cancer-linked variants of p53-TAD may be defective in their interactions with CBP through TAZ2 with N29D/N30D, D49Y, and W53G showing the strongest binding deficiencies. A loss of interaction with CBP implies inhibition of the mechanisms that stabilize and activate p53, which could lead to a loss of the pro-apoptotic tumor suppressor activity and, as a result, contribute to a tumor development.

Interestingly, the N29D/N30D variant of p53-TAD apparently lost affinity towards both HDM2 and CBP (see Table 2.1). On the one hand, the N29D/N30D targeting for proteasomal degradation could be suppressed due to a less efficient ubiquitination by HDM2. However, even if the cellular stability of the N29D/N30D variant is enhanced, as compared to WT p53, the simultaneous CBP interaction defect could produce a dominant tumor-suppressor loss-of-function phenotype. Altogether, our results help explain why the selected mutations in p53-TAD are associated with cancer occurrence by documenting a loss of affinity in those protein variants towards the activating factor CBP.

Importantly, not all of the selected mutations directly perturb the binding interface of p53-TAD with the affected partners. For example, whereas the K24N variant shows a ~3-fold reduction of binding affinity towards TAZ2 (see Table 2.1), Lys24 is fully solvent-exposed in the p53:TAZ2 complex [148], [149]. Similarly, the p53-TAD binding to TAZ2 is abolished in the D49Y variant, although Asp49 does not form contacts with the TAZ2 surface [148], [149]. In addition, the N29K/N30D variant shows a binding defect towards both CBP (through TAZ2 domain) and HDM2, although the Asn29/30 tandem is positioned outside the AD1 and AD2 regions that harbor the interaction sites for HDM2 and CBP [78], [148], [149]. Altogether, these results suggest that the selected mutations may indirectly affect the interaction affinities by modulating the unbound conformational ensemble of p53-TAD. To examine this hypothesis, in a parallel study we combined NMR spectroscopy with atomistic simulations to characterize the structural and dynamic properties of WT p53-TAD and its variants. The NMR studies briefly described below were performed by Dr. Lynn Schrag at Kansas State University and the atomistic simulation studies were performed by Dr. Xiaorong Liu at the University of Massachusetts.

The chemical shift perturbation (CSP) analysis of ^{1}H and ^{15}N chemical shift assignments of the p53-TAD mutants against their WT counterparts was analyzed. The results suggest that the lack of distal CSPs from the site of mutation for all p53 variants implies that the overall structure of the sampled conformational ensemble for unbound p53-TAD is only marginally impacted by the selected mutations. The results obtained from NMR relaxation parameters (spin-lattice relaxation rate (R_1), spin-spin relaxation rate (R_2), hetero-nuclear NOE (NHNOE/NONOE)) suggest that all mutants remain similarly disordered at the secondary structure level. Altogether, the NMR analyses suggest that cancer mutants may perturb the

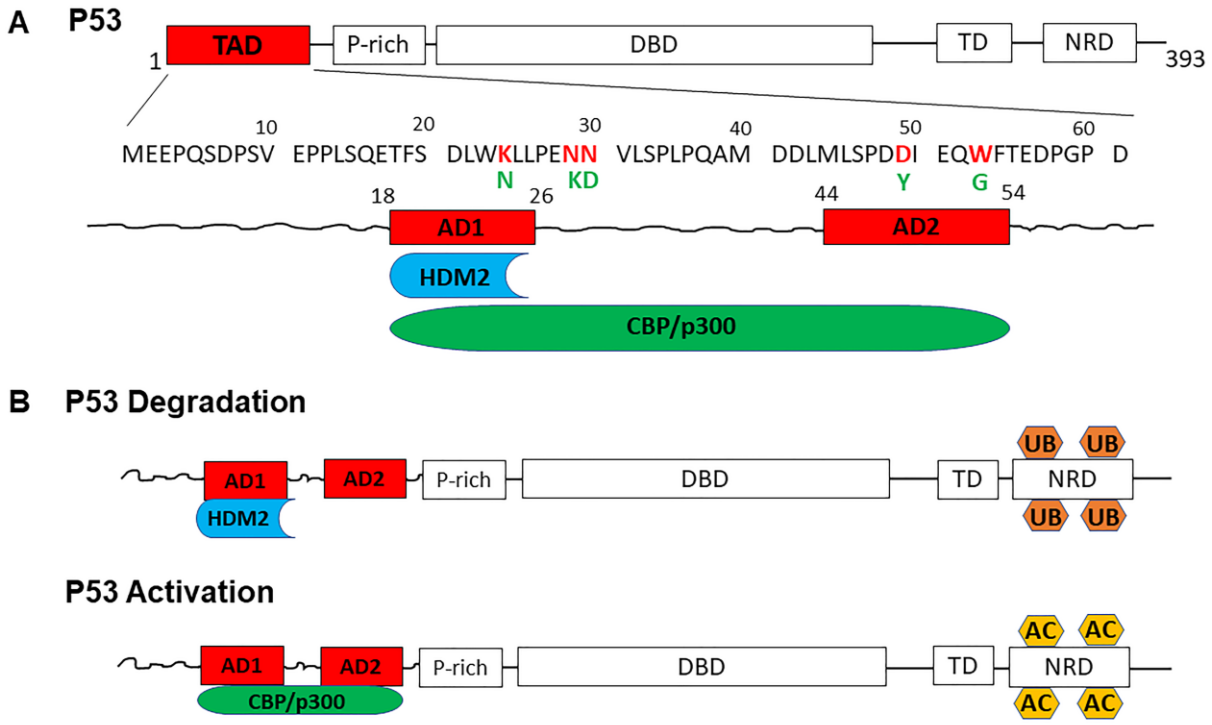
disordered ensemble of p53-TAD without significantly shifting the local secondary structure propensities.

Atomistic simulations were performed to further reveal how the disordered ensembles of p53-TAD may be modulated by mutations. Atomistic ensembles of four p53-TAD cancer-associated mutants were generated. These ensembles recapitulate the observation from NMR chemical shift analysis that these four cancer mutations do not substantially perturb local structures of p53-TAD on the ensemble level. The four cancer-associated mutations likely perturb the unbound ensemble of p53-TAD locally and distally, inducing subtle but non-negligible changes in both the overall dimensions and secondary structures of the protein.

Our observation that even subtle modulation of the disordered conformational ensemble could lead to dramatic changes in function such as specific binding with key regulators is intriguing. It may reflect a fundamental principle on how IDPs perform versatile functions in biology, that the disordered ensemble of IDPs is poised to respond sensitively and rapidly to various cellular stimuli, thus acting as a central conduit in cellular signaling and regulation. It further suggests a new conceptual framework to guide the design of novel therapeutic strategies targeting IDPs, where small molecules could modulate the structural ensembles of IDPs through dynamic interactions instead of specific ones to alter the interaction profiles and control their functions.

Figure 2.1:

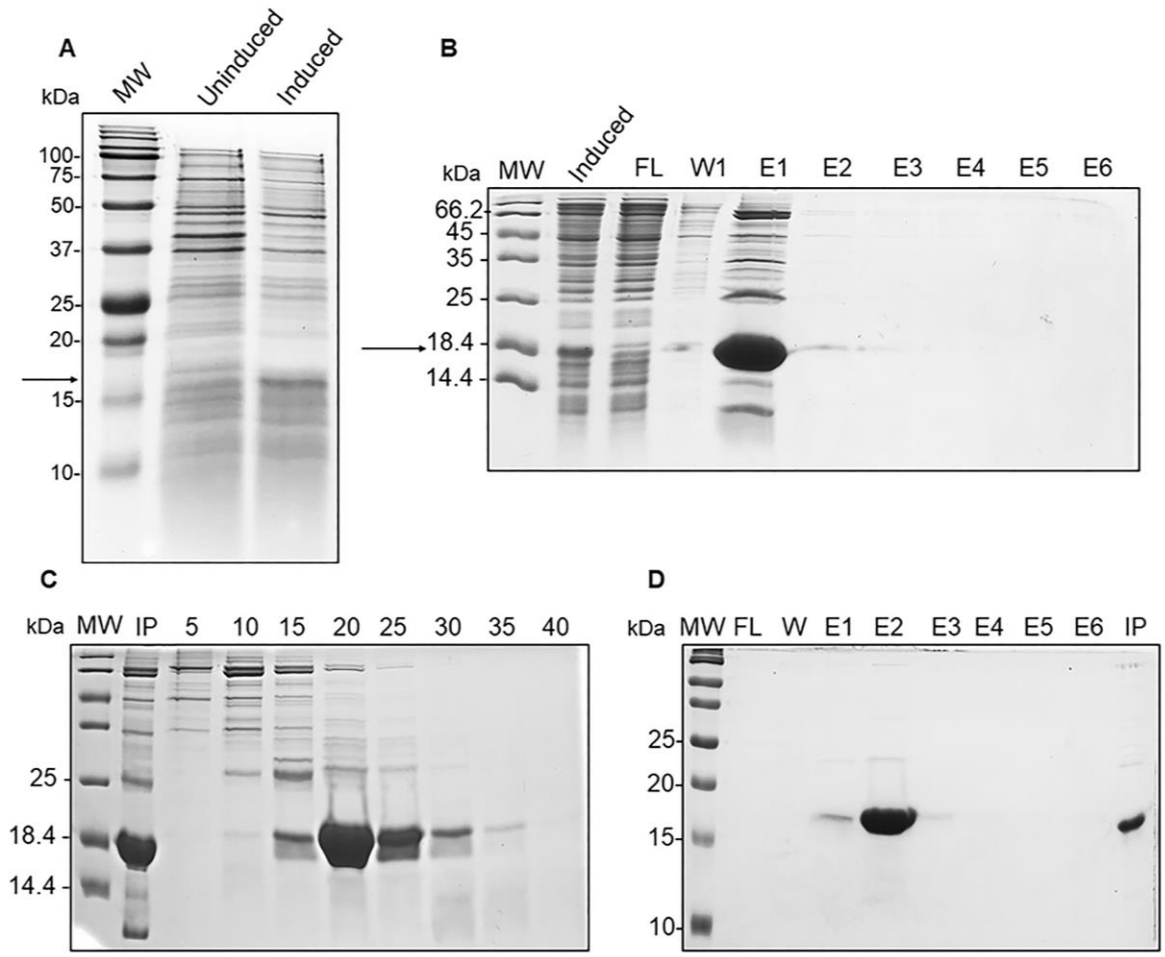
p53-TAD domain structure, interactions and regulation



A) Domain structure of p53: TAD (Transactivation domain), P-rich (Proline-rich domain), DBD (DNA-binding domain), TD (Tetramerization domain), NRD (Negative regulatory domain). The sequence p53-TAD is shown, with the selected cancer-linked mutation positions marked in red and the amino acid substitutions shown in green. The helical segments within TAD, AD1 and AD2, are shown below with the interacting partners, HDM2 and CBP/p300. B) p53 regulation is mediated by HDM2 and CBP/p300. Top: In unstressed cells, HDM2 binds strongly to AD1 and negatively regulates p53 stability by polyubiquitination (UB) of NRD. Bottom: Under genotoxic stress, phosphorylation of TAD inhibits HDM2-p53-TAD interaction and enhances the formation of a complex with CBP/p300, which promotes acetylation (AC) of NRD and, activates and stabilizes p53.

Figure 2.2:

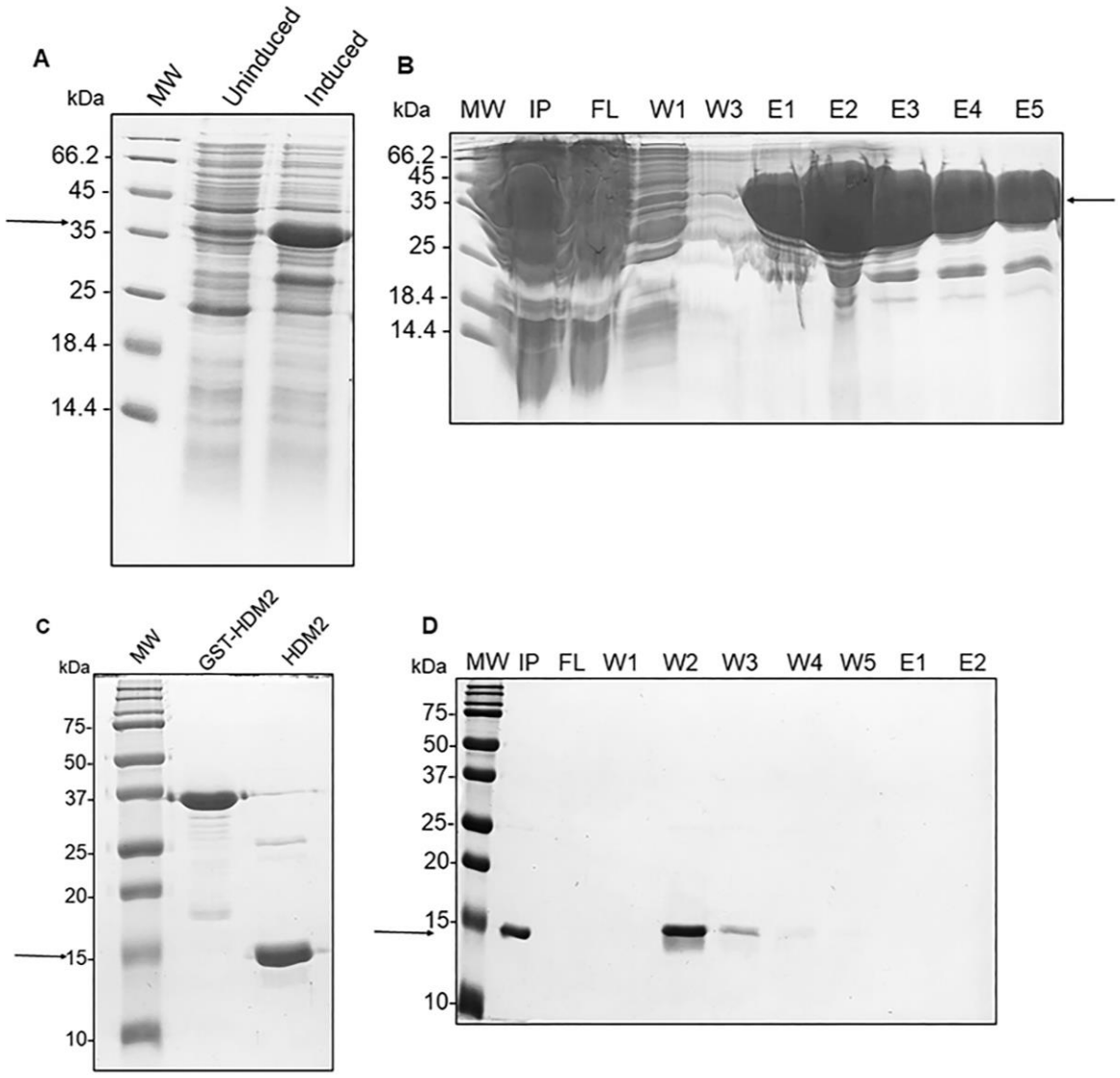
His-p53-TAD (1-73) wild type expression and purification



A) The p53-TAD wild type expression was checked without IPTG (Uninduced) and with IPTG (Induced). B) Ni-NTA purification; FL-Flow through, W-Wash fraction, E-Elution fraction. C) Gel-filtration chromatography and every 5th fraction was analyzed as shown in the figure; IP-Input. D) Second round of Ni-NTA purification; FL-Flow through, W-Wash fraction, E-Elution fraction.

Figure 2.3:

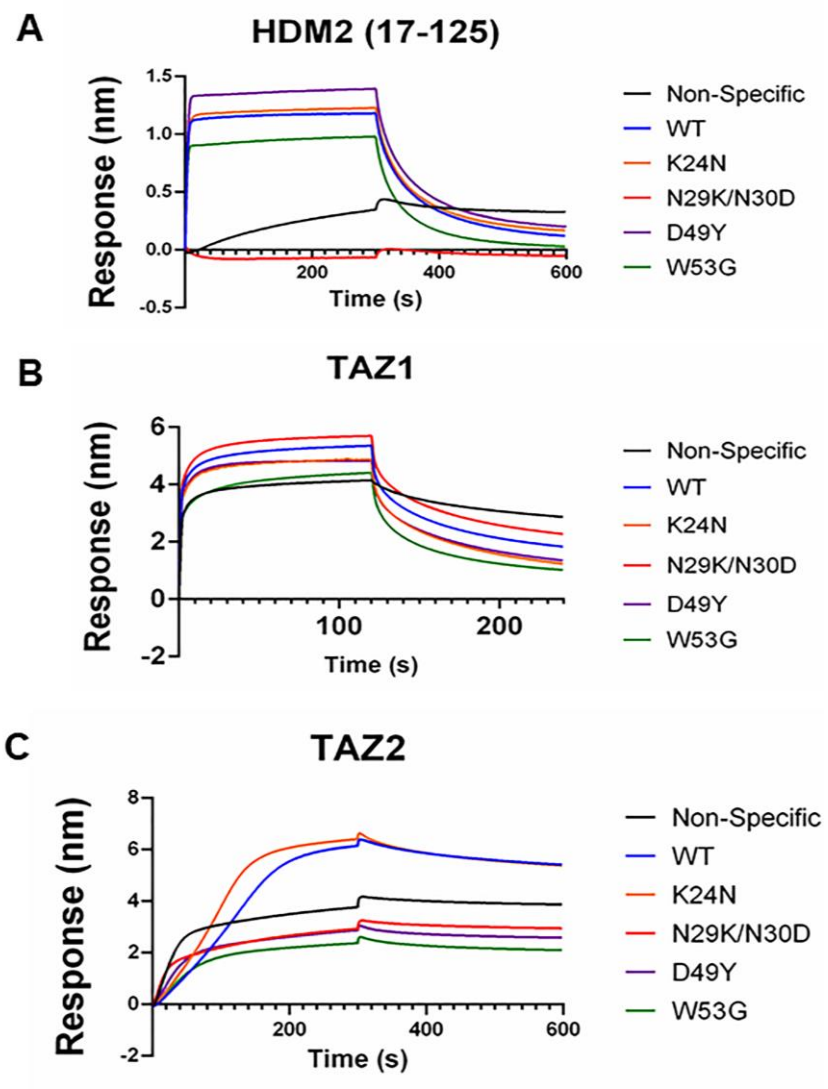
GST-HDM2 (17-125) expression and purification



A) GST-HDM2 protein expression was checked without IPTG (Uninduced) and with IPTG (Induced). B) Glutathione Sepharose 4B resin column purification; IP- Input, FL- Flow through, W- Wash fraction, E- Elution fraction. C) Cleavage of GST tag from HDM2 protein using HRV 3C protease. D) Glutathione Sepharose 4B resin column purification. The cleaved HDM2 protein was collected in the wash fractions; FL- Flow through, W- Wash fraction, E- Elution fraction

Figure 2.4:

Binding of p53-TAD variants to binding partners HDM2 (17-125), CBP-TAZ1 and CBP-TAZ2



100 ng/µL of p53-TAD WT and its variants were applied to Ni-NTA biosensor tip prior to the introduction of 3 µM HDM2 (A), 30 µM TAZ1 (B), or 300 nM TAZ2 (C). Binding response was calibrated to 0 at the start of association phase. All binding curves represent average response of two independent experiments. The Non-Specific curves were obtained with no p53 variants applied to the biosensor tip prior to the application of a binding partner.

Table 2.1:

Thermodynamic and kinetic parameters of interactions between p53-TAD variants and the binding partners

Binding partner	p53 variant	K_D (nM)	*k*on ($10^4 M^{-1}s^{-1}$)	*k*off ($10^{-3}s^{-1}$)
HDM2	WT	260±61	7.7±1.3	19.2±1.1
	K24N	301±39	7.1±0.9	21.0±1.0
	N29K/N30D	*N.D.*	*N.D.*	*N.D.*
	D49Y	275±52	8.0±1.1	21.6±2.6
	W53G	287±44	6.8±0.6	19.2±1.4
TAZ1	WT	1639±380	1.45±0.3	22.8±1.9
	K24N	2159±470	1.52±0.21	32.0±4.4
	N29K/N30D	1650±430	1.35±0.24	21.3±1.8
	D49Y	2188±630	1.45±0.25	30.2±2.8
	W53G	*4953±1500*	*0.67±0.27*	*30.0±5.2*
TAZ2	WT	61.10±0.97	8.36±0.13	5.26±0.5
	K24N	*210±14*	*4.53±0.31*	*9.5±1.6*
	N29K/N30D	*N.D.*	*N.D.*	*N.D.*
	D49Y	*N.D.*	*N.D.*	*N.D.*
	W53G	*N.D.*	*N.D.*	*N.D.*

Each interaction was measured with two replicates at three unique concentrations of each binding partner using BLI with an Ni-NTA biosensor tip. Concentrations used were 2 µM, 3 µM, and 4 µM for HDM2 (17-125), 20 µM, 30 µM, and 40 µM for TAZ1, and 300 nM, 600 nM, and 900 nM for TAZ2. k_{on} and k_{off} rates were then used to determine the effective K_D. N.D. indicates that the interaction was not detectable at this concentration range. The data indicating significant deviations from p53- TAD WT are shown in italics.

Chapter 3 - Human CLPB forms ATP-dependent complexes in the mitochondrial intermembrane space

This chapter has been published as the following journal article:

Thevarajan, I., Zolkiewski, M., & Zolkiewska, A. (2020). Human CLPB forms ATP-dependent complexes in the mitochondrial intermembrane space. *The International Journal of Biochemistry & Cell Biology*, 105841.

https://doi.org/10.1016/j.biocel.2020.105841

3.1 Abstract

Human caseinolytic peptidase B protein homolog (CLPB), also known as suppressor of potassium transport defect 3 (SKD3), is a broadly-expressed member of the family of ATPases associated with diverse cellular activities (AAA+). Mutations in the human *CLPB* gene cause 3-methylglutaconic aciduria type VII. CLPB is upregulated in acute myeloid leukemia (AML), where it contributes to anti-cancer drug resistance. The biological function of CLPB in human cells and mechanistic links to the clinical phenotypes are currently unknown. Herein, subcellular fractionation of human HEK-293 and BT-549 cells showed that a single 57-kDa form of CLPB was present in the mitochondria and not in the cytosolic fraction. Immunofluorescence staining of HEK-293 and BT-549 cells with anti-CLPB antibody co-localized with the mitochondrial staining using a MitoTracker dye. In purified intact mitochondria, CLPB was protected against externally added proteinase K, but it was susceptible to degradation after disruption of the outer membrane, indicating that CLPB resides in the mitochondrial intermembrane space. Overexpressed CLPB, while properly trafficked to the mitochondria, appeared to form large clusters/aggregates that were resistant to extraction with non-ionic detergents and were readily

visualized by immunofluorescence microscopy. Importantly, endogenous CLPB formed high molecular weight protein complexes in an ATP-dependent manner that were detected by blue native polyacrylamide gel electrophoresis. These results demonstrate that ATP induces a structural change in CLPB and controls its ability to self-associate or form complexes with other proteins in the intermembrane space of mitochondria.

Keywords: Mitochondria, ATPase, mutations, disease, chaperone, aggregation

3.2 Introduction

Human caseinolytic peptidase B protein homolog, CLPB, also known as a suppressor of potassium transport defect 3, SKD3, is a poorly characterized member of the AAA+ family of ATPases associated with diverse cellular activities [150]–[152]. The *CLPB* gene is located on the human chromosome 11q13 and is broadly expressed in various human organs and tissues [35]–[37], [153]. Homozygous or compound heterozygous mutations in the human *CLPB* gene are the cause of 3-methylglutaconic aciduria with cataracts, neurologic involvement, and neutropenia (MEGCANN; MIM entry # 616271), also referred to as 3-methylglutaconic aciduria (3-MGA-uria) type VII, an autosomal recessive genetic disease affecting young children [36]–[39], [41], [42], [154]. In its severe form, the disease causes death at a few months of age as a result of significant neonatal neurologic symptoms (absence of voluntary movements, hypotonia or hypertonia, swallowing problems, respiratory problems, and epilepsy) and severe neutropenia associated with life-threatening infections [42]. Upregulation of CLPB expression is observed in human acute myeloid leukemia (AML) upon acquisition of resistance to venetoclax, a selective BCL2 inhibitor and an FDA-approved drug for the treatment of AML [47]. Ablation of CLPB sensitizes AML cells to venetoclax [47]. A clearly documented link of

CLPB to severe human pathologies is intriguing, but the biological function of this protein is unknown.

Human CLPB is listed in MitoCarta 2.0 as a mitochondrial protein, with a Maestro score of 24 (www.broadinstitute.org/pubs/MitoCarta). MitoCarta 2.0 is an inventory of 1158 human proteins with strong indications of mitochondrial localization [155], [156]. Maestro score is an integrated score of mitochondrial localization, where the cut-off score of 5.22 indicates 5% false discovery rate, and the maximum score is 43 [155]. Consistent with the mitochondrial localization of human CLPB is the presence of a mitochondrial targeting presequence at the N-terminus of the protein and enrichment of CLPB in the mitochondrial proteome of several tissues [156]. Notably, CLPB was detected in the mitochondrial intermembrane space (IMS), but not in the mitochondrial matrix of HEK-293 cells, after mapping the IMS and matrix proteomes using an engineered ascorbate peroxidase (APEX) [157], [158]. The proteomic analyses suggest that CLPB resides in the IMS and if so, it may participate in IMS-specific functions, such as mitochondrial protein import and quality control [159]–[162]. Furthermore, CLPB was reported to be a substrate for presenilin-associated rhomboid-like (PARL) protease [163], [164]. PARL resides in the inner mitochondrial membrane and executes the intramembrane cleavage of substrate proteins during mitochondrial import, often leading to the liberation of a soluble protein into the IMS [165]. Importantly, some IMS proteins are dually localized and found not only in the IMS, but also in the cytosol [166], [167]. Whether or not CLPB localizes exclusively to the IMS in human cells is not known.

Human CLPB protein contains a P-loop nucleotide-binding AAA module characteristic of all AAA+ ATPases [150]. This domain converts the energy of ATP hydrolysis into mechanical force driving diverse cellular processes, including DNA replication, transcription,

recombination, cytosolic and mitochondrial protein quality control, ribosome biogenesis, cytoskeletal rearrangement, membrane fusion, viral genome packaging, or malarial infection [152], [168]–[173]. The AAA domain of human CLPB is closely related to the C-terminal AAA domain of the microbial ClpB/Hsp104/Hsp78 family of proteins. Bacterial ClpB, yeast cytosolic Hsp104, and yeast mitochondrial Hsp78 are well characterized molecular chaperones that cooperate with Hsp70 and Hsp40 during reactivation of aggregated proteins under severe cellular stresses [43], [174], [175]. Because of the similarity between the AAA domains of human CLPB and microbial ClpB/Hsp104/Hsp78, it has been postulated that human CLPB may also function as a molecular chaperone/disaggregase [44]. However, there are essential structural differences between human CLPB and microbial ClpB, which suggest that these two proteins may perform diverse functions. First, human CLPB contains a single AAA domain, whereas microbial ClpBs contain two AAA modules, separated by a unique coiled-coil middle domain that controls the disaggregase activity. This coiled-coil domain is not found in human CLPB. Second, the N-terminal region of human CLPB contains an ankyrin-repeat domain, but ankyrin repeats are not present in the microbial ClpBs. Whereas ankyrin repeats are known to mediate protein-protein interactions in a nucleotide-independent manner [176], [177], microbial ClpBs associate into hexameric rings in the presence of ATP or ADP [46]. Furthermore, in microbial ClpBs, chaperone activity strictly depends on ClpB oligomerization, which is supported by each of the two AAA domains [45], [178]. It is not clear whether human CLPB, with its single AAA domain and the non-canonical ankyrin repeats, could engage in stable nucleotide-dependent protein complexes in cells.

The goal of our studies was to establish the precise localization of endogenous CLPB in human cells, determine to what extent an overexpressed version of CLPB mirrors endogenous

CLPB, and examine whether CLPB forms ATP-dependent complexes in intact cells. We show that endogenous CLPB is found predominantly in the IMS, where it forms large molecular weight protein complexes in an ATP-dependent manner. Overexpressed CLPB, while properly trafficked to the mitochondria, appears to form large clusters/aggregates that are resistant to extraction with non-ionic detergents and are readily visualized by immunofluorescence microscopy. These findings should inform future studies on the biological function of CLPB in human cells and on mechanistic links between pathogenic CLPB mutations and clinical phenotypes.

3.3 Materials and methods

3.3.1 Reagents and antibodies

X-tremeGENE HP DNA transfection reagent, aprotinin, pepstatin, leupeptin, phenylmethanesulfonyl fluoride (PMSF), adenosine 5'-triphospahte (ATP) disodium salt (#A2383), adenosine 5'[γ-thio]triphosphate tetralithium salt (ATPγS) (#A1388), and anti-CLPB rabbit pAb (#HPA039006) were from MilliporeSigma (Burlington, MA). 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF) was from Fisher Scientific. MitoTracker Orange CMTMRos dye (#M7510), NativePAGE sample prep kit (#BN2008), NativePAGE running buffer (#BN2001), NativePAGE cathode buffer additive (#BN2002), NativePAGE 4-16%, Bis-Tris mini protein gels (#BN1002BOX), NuPAGE transfer buffer (#NP0006), and NativeMark unstained protein standards (#LC0725) were from ThermoFisher Scientific (Waltham, MA). ON-TARGETplus human CLPB (81570) siRNA SMARTpool (siRNAs J-017179-09, J-017179-10, J-017179-11, J-017179-12), ON-TARGETplus non-targeting siRNA pool (D-001810-10), and Dharmafect 1 transfection reagent were from Dharmacon (Lafayette, CO). Anti-Myc mouse mAb (clone 9B11, #2276), anti-FLAG tag rabbit

mAb (clone D6W5B, #14793), anti-Cox IV rabbit mAb (clone 3E11, #4850), anti-MEK1/2 rabbit mAb (clone D1A5, #8727), anti-HSP60 rabbit mAb (clone D6F1, #12165), and anti-GAPDH rabbit mAb (clone D16H11, #5174) were from Cell Signaling Technology (Danvers, MA). Anti-YME1L1 rabbit pAb (#11510-1-AP) and anti-TOMM20 rabbit pAb (#11802-1-AP) were from Proteintech (Rosemont, IL). Proteinase K (#K1037) was from APExBIO (Houston, TX). Cy5- or AlexaFluor 488-conjugated secondary antibodies were from Jackson ImmunoResearch Laboratories, Inc. (West Groove, PA).

3.3.2 Cell culture

HEK-293A (referred to as HEK-293), HEK-293T, and BT-549 cells were from the American Tissue Culture Collection (Manassas, VA). HEK- 293A and HEK-293T cells were cultured in Dulbecco's Modified Eagle medium (DMEM) with 4.5 g/L glucose and 10% FBS. BT-549 cells were grown in RMPI-1640 medium containing 10% FBS and 5 µg/ml insulin. Cells were maintained at 37°C under humidified atmosphere containing 5% CO_2.

3.3.3 Transient overexpression of human CLPB

Cells were transfected with pCMV6-Entry vector containing an insert encoding human CLPB (NM_030813) with a C-terminal Myc/FLAG (OriGene Technologies, Rockville, MD; #RC203013). Cells in 6-well plates were transfected at 60-70% confluency using 1 µg of plasmid DNA and 2 µl of X-tremeGENE HP DNA transfection reagent, according to the manufacturer's protocol. After 48 hours, cell lysates were collected for immunoblotting.

3.3.4 Stable overexpression of human CLPB

Lentiviral particles encoding human CLPB were generated in HEK-293T cells. Cells were transfected with 1 µg of pLenti-C-Myc-DDK-P2A-Puro vector encoding human CLPB

(NM_030813) with a C-terminal Myc/FLAG (OriGene; #RC203013L3) and 1.2 µg of lentiviral packaging plasmid mix (OriGene; #TR30037), using TurboFectin transfection reagent (OriGene; TF810002, 6.6 µl per well). After 18 hours, medium was replaced, and after additional 24 hours conditioned medium containing lentiviral particles was collected, centrifuged for 5 min at 600xg, passed through a 0.45-µm filter, and stored at -80°C. HEK-293A cells were plated in 6-well plates ($2.5x10^5$ cells/well). After 24 h, medium was removed and 2 ml of conditioned medium containing lentiviral particles was added and supplemented with polybrene (8 µg/ml). After 24 hours, transduction medium was replaced with fresh medium. Selection with puromycin (0.5 µg/ml) started 72 h after adding viruses and continued for 7-10 days.
pLenti-C-Myc-DDK-P2A-Puro is a bicistronic vector that encodes a Myc/FLAG-tagged insert, the P2A self-cleaving peptide, and puromycin N-acetyl transferase. The sequence of the P2A peptide is GSGATNFSLLKQAGDVEENPGP, with the self-cleavage occurring between P and G. Thus, the stably expressed CLPB-Myc/FLAG protein contained a 20-aa extension at the C-terminus that was not present in transiently expressed CLPB-Myc/FLAG.

3.3.5 siRNA knockdown of CLPB

Cells in 6-well plates were transfected at 60-70% confluency using 50 nM siRNAs targeting human CLPB or control siRNAs and 2 µl of Dharmafect 1 reagent, according to the manufacturer's protocol. After 24 h, the medium was replaced, and after additional 48 h cell lysates were collected and analyzed by immunoblotting.

3.3.6 Immunoblotting

Cells were washed twice with DPBS and incubated for 10 min with lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 5 mM EDTA, 1 mM AEBSF, 5 µg/mL pepstatin, 5 µg/mL leupeptin, 5 µg/mL aprotinin, and 10

mM 1,10-phenanthroline), with shaking. Lysates were centrifuged for 15 minutes at 15,000xg, supernatants were supplemented with 3xSDS sample buffer and resolved by SDS-PAGE, followed by transfer to a nitrocellulose membrane. Membranes were blocked for 1 h with 5% skim dry milk and 0.3% Tween-20 in DPBS, incubated overnight at 4°C with primary antibodies diluted in blocking buffer and then for 1 h with HRP-conjugated secondary antibodies, followed by signal detection using SuperSignal West Pico chemiluminescence detection kit (Pierce) and Azure c500 digital imaging system. Band intensities were quantified using the ImageJ analysis software.

3.3.7 Confocal microscopy

Cells were plated on glass cover slips placed in 6-well plates and allowed to attach overnight. Cells were then washed twice with DPBS and incubated for 30 minutes at 37°C with 200 nM MitoTracker Orange CMTMRos dye in DPBS. After washing, cells were treated for 20 min with 4% paraformaldehyde and permeabilized for 5 min with 0.1% Triton X-100 in DPBS. Cells were then blocked for 30 min with 1% bovine serum albumin (BSA) and 5% donkey serum, incubated for 1 h with primary antibodies in DPBS containing 1% BSA and then for 45 min with dye-conjugated secondary antibodies. After washing, mounting on glass slides, and drying at room temperature in the dark, cells were imaged using a Zeiss 880 confocal fluorescence microscope.

3.3.8 Isolation of mitochondria

Mitochondria were isolated from HEK-293 and BT-549 cells using previously described protocols [163], with some modifications. Cells were trypsinized and centrifuged at 600xg for 5 minutes at 4°C. Cell pellets were washed with DPBS, suspended in homogenization buffer (20 mM HEPES pH 7.4, 220 mM mannitol, 70 mM sucrose, 1 mM EDTA), and transferred to a pre-

chilled glass dounce homogenizer. Cells were broken with 20-30 strokes of the pestle, on ice. The homogenates were centrifuged at 600xg for 5 minutes at 4°C. The supernatants were transferred to fresh tubes and centrifuged at 15,000xg for 15 minutes at 4°C. The supernatants (cytosols) were removed and pellets (mitochondria) were washed once with isotonic buffer (20 mM HEPES pH 7.4, 250 mM sucrose) and then resuspended to the same volume as the supernatants.

3.3.9 Proteinase K treatment

Mitochondrial pellets were resuspended in isotonic buffer (20 mM HEPES pH 7.4, 250 mM sucrose) or hypotonic buffer (10 mM HEPES pH 7.4, 1 mM EDTA), as indicated, and treated for 10 min at room temperature with different concentrations of proteinase K. The proteinase K digestion was stopped by adding 4 mM PMSF and incubating the samples for 5 min at room temperature and then for 10 min at 95°C. In parallel, some mitochondrial pellets resuspended in isotonic buffer were further solubilized for 10 min with 0.5% (v/v) Trition X-100 prior to the proteinase K treatment. All samples were then analyzed by SDS-PAGE and immunoblotting.

3.3.10 Blue-Native PAGE

Mitochondrial pellets were incubated for 15 min on ice with 1X Native PAGE sample buffer with protease inhibitors and 1% digitonin or 1% n-Dodecyl-β-D-Maltoside (DDM), in the presence or absence of 5 mM ATP or ATPγS and 10 mM $MgCl_2$. Lysates were then centrifuged for 30 min at 15,000xg, at 4°C. Supernatants were collected, supplemented with 0.25% G-250 sample additive, and electrophoresed in NativePAGE 4-16% Bis-Tris gels. After electrophoresis, proteins were transferred to a PVDF membrane and immunoblotted using anti-CLPB Ab and HRP-conjugated secondary Ab. SuperSignal WestPico chemiluminescent substrate (Pierce) was

used to detect the signal and images were acquired using Azure c500 digital imaging system.

3.3.11 Statistics

All experiments were repeated 2-5 times. Representative immunoblots or microscopic images are shown.

3.4 Results

3.4.1 Endogenous human CLPB resides in the mitochondrial intermembrane space

To detect the endogenous CLPB protein in human cells, we used a commercially available polyclonal antibody generated against aa 589 - 693 from human CLPB. In human embryonic kidney HEK-293 cells and human breast cancer cell line BT-549, a single band of ~57-kDa was detected by immunoblotting. Intensity of this band was reduced by ~90% after transfection of cells with siRNAs targeting CLPB, but not with control siRNAs (Figure 3.1(A)), which confirmed the identity of the CLPB band. To analyze the relative abundance of CLPB in mitochondria and cytosol, cell lysates were subjected to differential centrifugation. The purity of subcellular fractions was assessed by immunoblotting using antibodies against GAPDH and COX IV, a cytosolic and a mitochondrial marker, respectively. CLPB was present in post-nuclear supernatants (PNS) and in the mitochondria, with less than 5% of the protein detected in the cytosolic fractions (Figure 3.1(B)). Immunofluorescence staining of HEK-293 and BT-549 cells with anti-CLPB antibody and a mitochondrial MitoTracker dye showed that endogenous CLPB was located in the mitochondria. The specificity of the CLPB signal was confirmed by using CLPB-targeting siRNA (Figure 3.1(C)).

To explore CLPB localization within the mitochondria, we compared the accessibility of

CLPB to externally added proteinase K (PK) under three conditions: in intact mitochondria suspended in isotonic buffer, upon disruption of the outer mitochondrial membrane after osmotic swelling in hypotonic buffer, and upon a complete membrane solubilization with Triton X-100. We observed that CLPB was protected against PK in intact mitochondria but was degraded, in a PK dose dependent manner, upon disruption of the outer membrane in both HEK-293 and BT-549 cells (Figure 3.2(A), (B)). The CLPB susceptibility to PK was similar to that of YME1L1, the i-AAA protease that is anchored in the inner mitochondrial membrane and exposed on the intermembrane side [179]. Unlike CLPB, TOMM20, a component of the receptor complex responsible for the import of mitochondrial proteins, which functions at the surface of the mitochondrial outer membrane [180], was susceptible to PK degradation even in the isotonic buffer. In contrast, HSP60, a molecular chaperone located in the mitochondrial matrix [181], became accessible to externally added PK only after solubilization of the inner membrane (Figure 3.2(A), (B)). These results indicate that CLPB resides in the mitochondrial intermembrane space (IMS).

3.4.2 Overexpressed human CLPB is prone to aggregation in the mitochondria

In contrast to a single form of the endogenous CLPB protein, two different forms of exogenously expressed CLPB were detected in HEK-293 cells (Figure 3.3). Transient transfection of HEK-293 cells with a construct encoding human CLPB bearing a C-terminal Myc/FLAG tandem tag produced a ~76-kDa and a ~62-kDa forms (Figure 3.3(A), (B)). Subcellular fractionation showed that the 76-kDa form represented the precursor CLPB-Myc/FLAG protein present in the cytosol and the 62-kDa form corresponded to the mature protein localized to the mitochondria (Figure 3.3(C)). The mitochondrial form of CLPB-

Myc/FLAG was larger than the endogenous CLPB protein, 62-kDa *versus* 57-kDa, due to the presence of the Myc/FLAG tag (the predicted MW of the tag is 3.9 kDa). The molecular weight difference between the cytosolic and mitochondrial forms of CLPB-Myc/FLAG, 76 kDa versus 62 kDa, is consistent with the removal of the N-terminal 127 amino acids of CLPB (12.9 kDa) during a sequential cleavage mediated by the mitochondrial processing peptidase MMP and the intramembrane protease PARL during the protein import into mitochondria [163].

Although the 62-kDa mitochondrial form of CLPB-Myc/FLAG was likely derived from the 76-kDa precursor, the time course of CLPB-Myc/FLAG expression suggested that the 76-kDa form persisted for a longer period of time after transient transfection than the 62-kDa form (Figure 3.3(B)). This result suggested that: (a) import of CLPB into mitochondria was rate-limiting for an overexpressed CLPB protein, and (b) processing of overexpressed CLPB by mitochondrial proteases was very efficient, resulting in a single mitochondrial form of CLPB. Analogous results were also obtained in CLPB-Myc/FLAG-transfected SUM-159 breast cancer cells (results not shown).

Two forms of CLPB were also observed in HEK-293 cells with stable overexpression of CLPB (Figure 3.3(D)). The molecular weights of these two forms, ~78 kDa and ~64 kDa, were slightly higher than those detected during transient transfection, because the lentiviral CLPB construct used for stable expression contained a 20-aa (~2-kDa) P2A extension at the C-terminus (see Materials and methods). The 78-kDa form, partitioned to the cytosol, whereas the 64-kDa band was present in the mitochondria (Figure 3.3(E)). Immunofluorescence microscopy of HEK-293-CLPB-Myc/FLAG cells using anti-Myc antibody detected a signal localized in discrete puncta that was distinct from the more diffuse signal observed in parental HEK-293 cells probed with anti-CLPB antibody (compare Figures 3.3(G) and 3.1(C)). Furthermore, anti-Myc antibody

staining did not fully overlap with anti-COX IV or anti-YME1L1 staining (Figure 3.3(G)). To better quantify the relationship between anti-Myc, anti-COX IV, and anti-YME1L1 signals, we used the intensity profile feature of the confocal analysis software, which quantifies the staining intensity along a line drawn in the image (Figure 3.3(G), right panels). Clearly, some of the intense anti-Myc puncta were only weakly stained with anti-COX IV or anti-YME1L1 antibodies, and some mitochondrial regions positive for COX IV or YME1L1 were only weakly stained with anti-Myc antibody. The strong puncta-like anti-Myc signals might represent CLPB clusters formed prior to the import of CLPB into the mitochondria and corresponding to the 78-kDa form in the Western blots (Figure 3.3(D)) or conglomerates of the mitochondrial 64-kDa CLPB form. Collectively, these results suggested that overexpressed CLPB, while properly targeted to mitochondria, may be prone to aggregation in cells.

To determine whether overexpressed CLPB might form aggregates inside the mitochondria, mitochondrial membranes were solubilized with a neutral non-denaturing detergents digitonin or n-dodecyl-beta-D-maltoside, and the lysates were fractionated by centrifugation. Supernatant fractions contained water-soluble proteins, multiprotein complexes, and solubilized membrane proteins, whereas pellets contained insoluble membrane protein complexes and protein aggregates [182], [183]. Remarkably, more than 90% of the overexpressed CLPB was found in the pellets, suggesting that it was a part of large protein aggregates located in the mitochondria (Figure 3.3(F)). In contrast, analogous fractionation of mitochondria from parental HEK-293 cells showed that more than 70% of endogenous CLPB was present in the supernatants (Figure 3.3(F)). Thus, overexpressed CLPB does not fully mimic the biochemical properties, and most likely the biological function, of endogenous CLPB due to its aggregation propensity, which is most likely caused by its high non-physiological

concentration in the IMS. This may be remedied in the future by tuning down the level of CLPB overexpression, for example by using an inducible expression system and applying suboptimal concentrations of an inducing agent. In any case, the apparent aggregation propensity of overexpressed CLPB should warrant caution when studying the role of pathogenic mutations in the *CLPB* gene and expressing individual CLPB mutants in mammalian cells.

3.4.3 Endogenous CLPB forms high molecular weight protein complexes in an ATP-dependent manner

Human CLPB contains a single AAA domain and recombinant CLPB expressed in *E. coli* is an active ATPase [37], [44]. Whether ATP binding induces any structural effects in CLPB that might control its ability to self-associate or form complexes with other proteins has not been investigated. Here, we analyzed endogenous CLPB in digitonin-solubilized HEK-293 and BT-549 mitochondria by blue-native polyacrylamide gel electrophoresis (BN-PAGE). In this approach, Coomassie G-250 dye binds to the surface of proteins and confers a net negative charge while maintaining the proteins in their native conformation, which is then followed by separation of the negatively charged complexes according to their molecular masses [184]–[186]. We observed that in the absence of nucleotides, CLPB migrated in BN-PAGE as a single band of ~120 kDa (Figure 3.4). This apparent molecular mass was larger than expected for a CLPB monomer (~57 kDa), suggesting that CLPB could form dimers or indicating an abnormal mobility of CLPB monomers in BN-PAGE, possibly due to their non-spherical shape. Importantly, when mitochondrial extracts were pre-incubated with 5 mM ATP prior to the electrophoresis, the migration pattern of CLPB changed significantly, indicating the presence of complexes with a higher molecular weight containing CLPB (Figure 3.4). The population of high-molecular weight complexes further increased when ATP was replaced with ATPγS, a non-

hydrolyzable ATP analog (Figure 3.4). Thus, assembly of protein complexes that contain CLPB in the mitochondrial IMS critically depends on the presence of ATP.

3.5 Discussion

3.5.1 CLPB is the second protein in the mitochondrial IMS linked to 3-methylglutaconic aciduria

We demonstrate that CLPB localizes to the mitochondrial IMS in human cells, where it engages in ATP-dependent protein-protein interactions. Our results agree with a previous large scale proteomic study, in which proteins located in the mitochondrial IMS in live human cells were tagged using APEX-based proximity labeling and analyzed by mass spectrometry [157]. Our results are also consistent with the finding that CLPB is cleaved by PARL during the transport through the inner mitochondrial membrane [163], which should then lead to the release of soluble CLPB into the IMS. Thus, the *CLPB* gene is one of the two genes encoding water-soluble proteins in the IMS, next to *HTRA2*, that are involved in 3-MGA-urias (OMIM phenotypic series PS250950; https://www.omim.org/phenotypicSeries/PS250950, downloaded on 06/04/2020).

For the nine currently catalogued types of 3-MGA-urias, genes associated with the disease have been identified for eight types, I-III and V-IX [187], [188]. Among these eight genes, seven - *AUH*, *TAZ*, *OPA3*, *DNAJC19*, *SERAC1*, *HTRA2*, and *TIMM50* - encode proteins for which biological functions are known. In contrast, the function of CLPB in human mitochondria is poorly understood. Since bacterial ClpB and its two yeast orthologs, Hsp104 in the cytosol and Hsp78 in the mitochondrial matrix, are molecular chaperones that reactivate protein aggregates, it has been postulated that human CLPB acts as a protein disaggregase as

well. Indeed, it has been recently shown that bacterially-expressed human CLPB can disaggregate two model substrates, firefly luciferase aggregates and α-synuclein fibrils, *in vitro* [44]. However, another report showed that recombinant human CLPB was much less potent than microbial disaggregases and, in contrast to microbial disaggregases, had no activity towards thermally unfolded GFP [189]. It is not clear whether human CLPB acts as a disaggregase *in vivo* and the identities of its putative substrates in the IMS are unknown.

3.5.2 Key differences between animal CLPB and microbial ClpB/HSP104/Hsp78

Remarkably, comparative and integrative genomic approaches showed that during evolution the *HSP104* and *HSP78* genes were lost in animals [190]. The gene designated as *CLPB* in animals arose as a fusion between an N-terminal ankyrin domain-encoding gene and a C-terminal fragment of the bacterial *clpB* gene [190]. However, because of the key features distinguishing the protein product of this novel animal *CLPB* gene from the microbial ClpB/Hsp104/Hsp78 proteins, namely a single AAA module, the presence of ankyrin repeats, and the localization in the mitochondrial IMS, it is likely that the biological function of animal CLPB differs from the function of its apparent ancestors in microorganisms. For example, recent studies suggested that CLPB is important for maintaining the mitochondrial cristae and contributes to the survival of acute myeloid leukemia cells [47] and to the mitochondrially-mediated antiviral innate immunity [48].

3.5.3 CLPB undergoes nucleotide-dependent structural changes

Recombinant human CLPB produced in *E. coli* is an active ATPase [37], [44], [189], as might be expected of a protein containing an AAA domain. Our results show that ATP, in

addition to serving as a substrate for CLPB ATPase, controls CLPB's ability to form large protein complexes in cells. While in the absence of added nucleotides CLPB migrated as a single low-molecular weight species in BN-PAGE, this form was almost undetectable in the presence of added ATP or ATPγS and was converted into higher molecular-weight species. While it is not known at this time whether these species represented CLPB homo-oligomers or hetero-complexes with other mitochondrial proteins, our results suggest that CLPB undegoes ATP- (or, more generally, nucleotide-) dependent structural changes that control its homo- or hetero-interactions and thus may support its biological function.

Notably, recent reports indicate that human CLPB might interact with the inner membrane-localized prohibitin (PHB) complex [48] and dynamin-related GTPase optic atrophy type 1 (OPA1) [47], outer membrane-localized mitochondrial antiviral signaling protein (MAVS) [48], HAX1 protein localized on the inner and outer mitochondrial membranes and exposed to the intermembrane space [37], [47]. Whether these interactions with CLPB in the IMS are ATP- (or nucleotide-) dependent is a subject of further investigation.

3.6 Conclusions

Our results demonstrate a key difference between human CLPB and the microbial ClpB/Hsp104/Hsp78 family of molecular chaperones/disaggregases. While bacterial ClpB is located in the cytoplasm, and yeast Hsp104 and Hsp78 exist in the cytosol and the mitochondrial matrix, respectively, human CLPB resides principally within the mitochondrial intermembrane space. This localization of human CLPB, together with the presence of a single AAA domain and a unique ankyrin-repeat domain, suggest that the function of human CLPB may differ from the microbial ClpB/Hsp104/Hsp78 proteins. Furthermore, our finding that CLPB forms high molecular weight protein complexes in the intermembrane space in an ATP-dependent manner

provides an insight into the structural aspects of CLPB that may be key for understanding the mechanism of its action.

Declaration of competing interests

The authors declare that they have no competing interests.

Author contributions

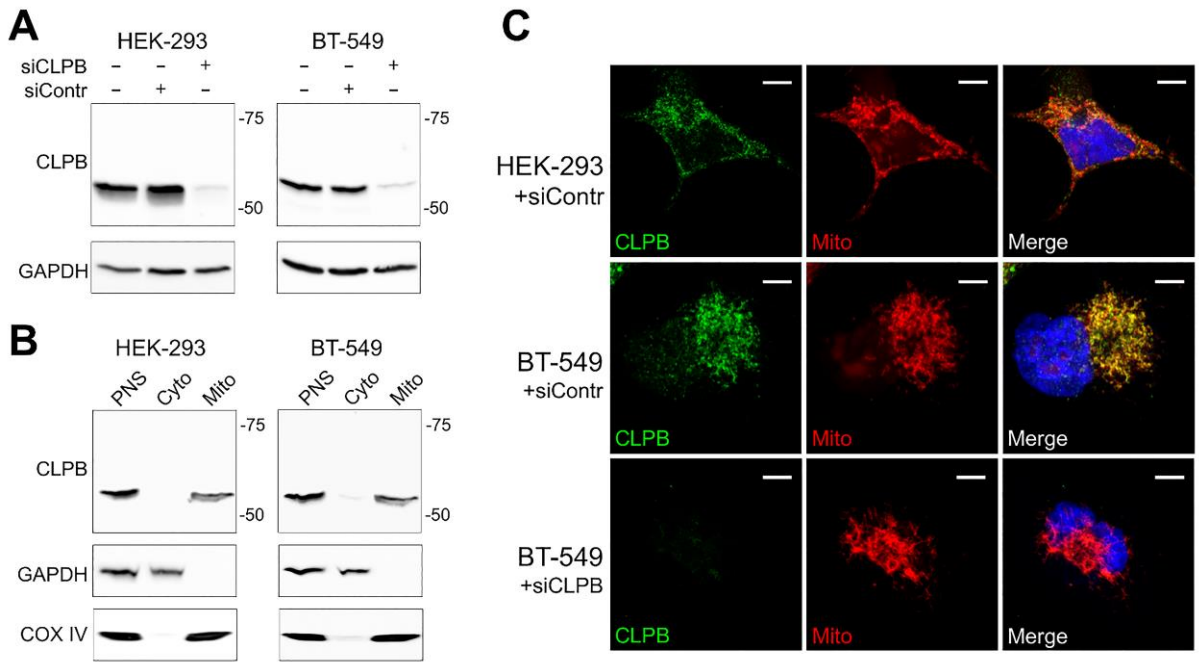
AZ and MZ conceived, designed, and supervised the study. IT performed the experiments. IT, AZ, and MZ performed data analyses. AZ and IT drafted the manuscript. All authors provided comments on the manuscript and approved the submitted version.

Acknowledgements

We would like to acknowledge the help of Joel Sanneman and the College of Veterinary Medicine Confocal Core Facility in acquiring confocal images. This is contribution 20-335-J from Kansas Agricultural Experiment Station.

Figure 3.1:

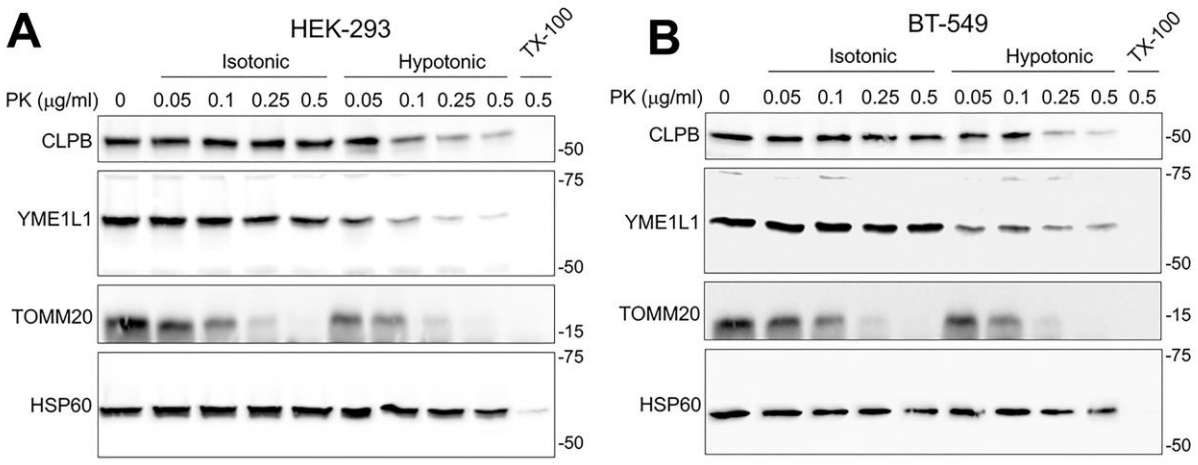
Subcellular localization of endogenous human CLPB



(A) Validation of anti-human CLPB antibody. HEK-293 or BT-549 cells were transfected with siRNAs targeting human CLPB (siCLPB) or with non-targeting siRNA (siContr). Total cell lysates were analyzed by SDS-PAGE and Western blotting 72 h after transfection. GAPDH is a gel-loading control. (B) Subcellular fractionation of HEK-293 and BT-549 cells. Post-nuclear supernatants (PNS), cytosolic (Cyto) and mitochondrial (Mito) fractions were subjected to SDS-PAGE and immunoblotting with anti-CLPB, anti-GAPDH (a cytosolic marker), or anti-COX IV (a mitochondrial marker) antibodies. (C) Confocal images of HEK-293 and BT-549 cells stained with anti-CLPB Ab (green), MitoTracker Orange CMTMRos dye (red), and DAPI (blue). Cells were transfected with siContr or siCLPB 72 h prior to staining. Bar, 5 μm.

Figure 3.2:

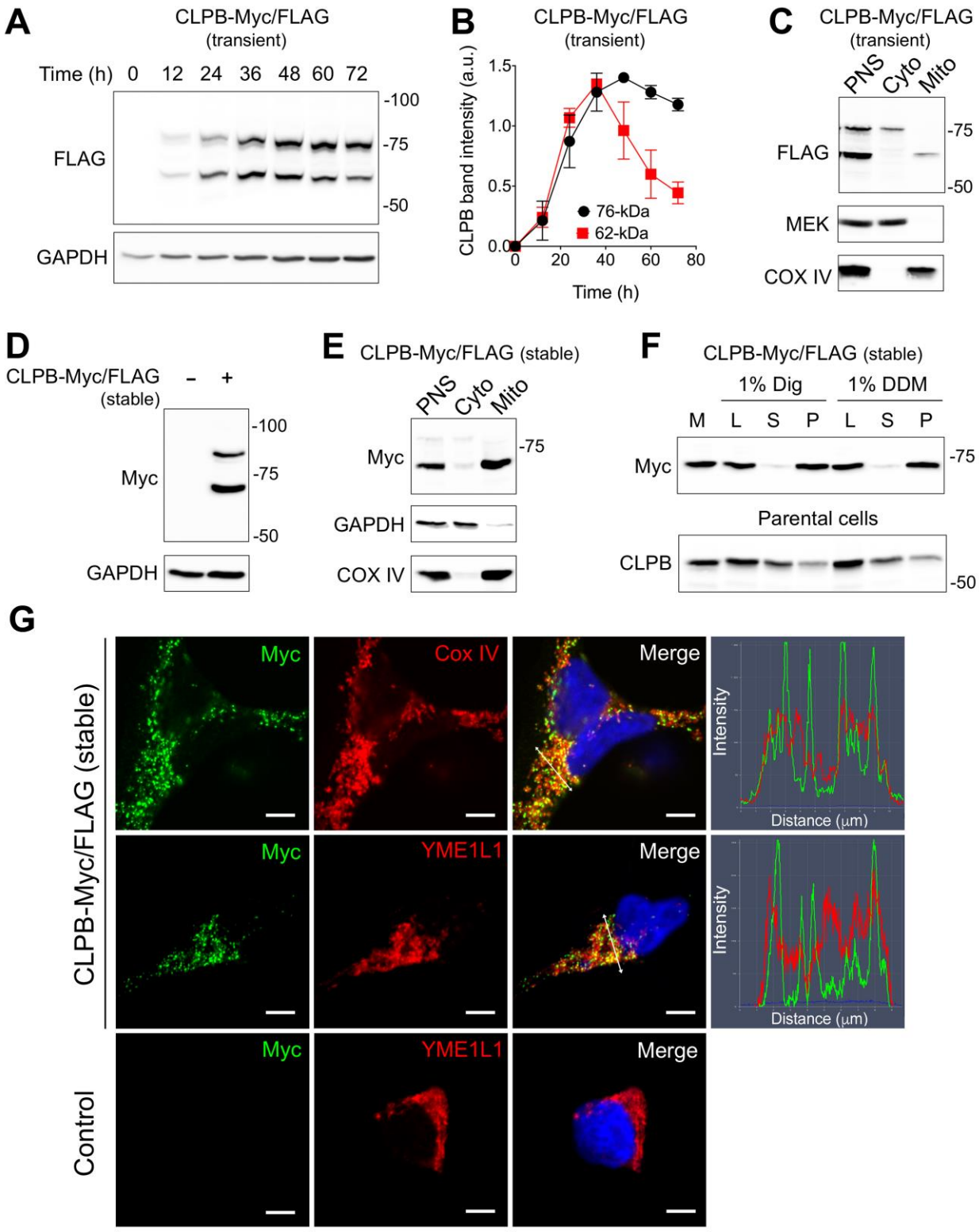
Accessibility of endogenous human CLPB to proteinase K



Intact mitochondria isolated from HEK-293 (A) or BT-549 cells (B), suspended in isotonic or hypotonic buffer, or mitochondria solubilized with 0.5% Triton X-100 (TX-100), were incubated for 15 min in the presence of the indicated concentrations of proteinase K (PK).
Equivalent amounts of each sample were resolved by SDS-PAGE and analyzed by immunoblotting using antibodies against CLPB, YME1L1 (an intermembrane space protein), anti-TOMM20 (an outer membrane protein) or HSP60 (a mitochondrial matrix protein).

Figure 3.3:

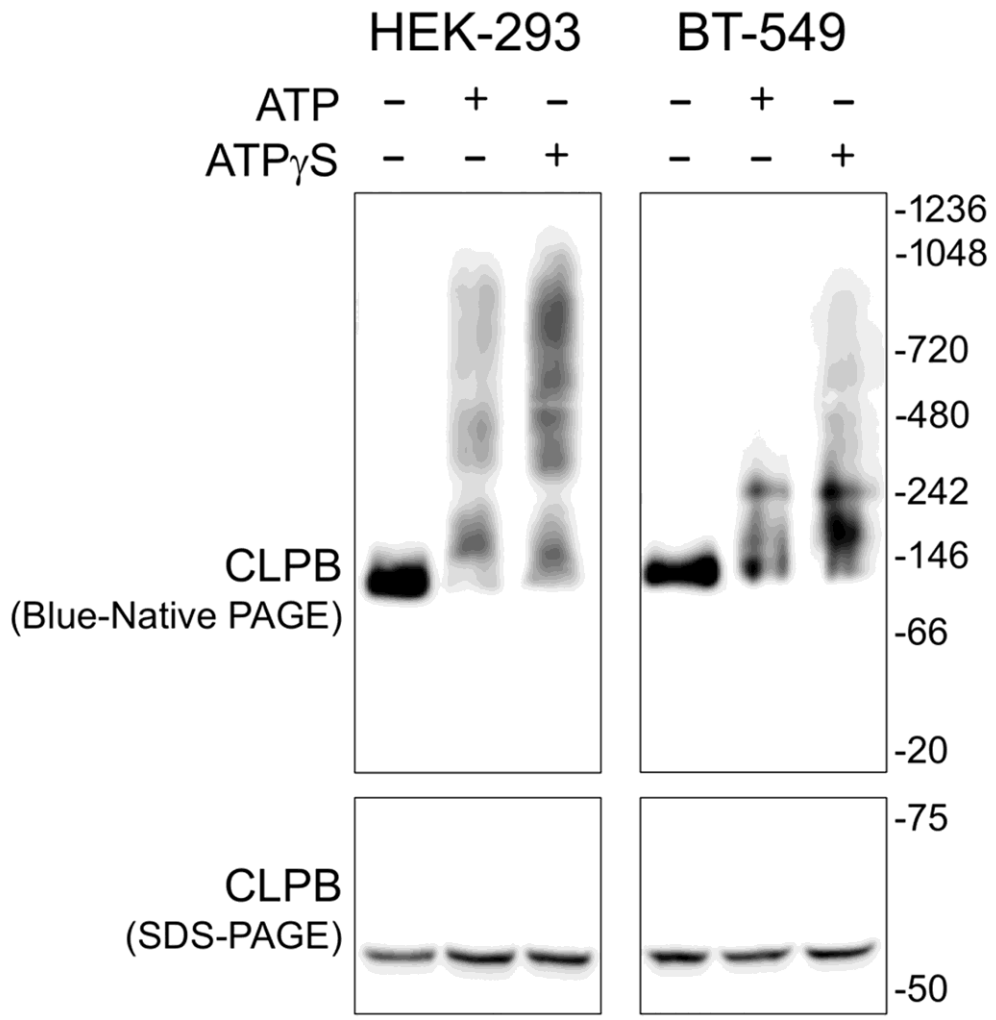
Characterization of human CLPB exogenously expressed in HEK-293 cells



(A) Time course of CLPB-Myc/FLAG expression in transiently transfected cells. Total cell lysates were analyzed at indicated times after transfection by immunoblotting with anti-FLAG antibody; GAPDH is a gel-loading control. (B) Intensities of the 76-kDa and 62-kDa CLPB-Myc/FLAG bands were quantified by densitometry and normalized to GAPDH levels. (C) Subcellular fractionation of HEK-293 cells with transient expression of CLPB-Myc/FLAG. Post-nuclear supernatant (PNS), cytosolic fraction (Cyto), and mitochondrial fraction (Mito) were subjected to SDS-PAGE and immunoblotting with anti-FLAG, anti-MEK (a cytosolic marker), or anti-COX IV (a mitochondrial marker) antibodies. (D) Stable overexpression of CLPB-Myc/FLAG in HEK-293 cells. Total cell lysates were analyzed by immunoblotting with anti-Myc antibody; GAPDH is a gel-loading control. (E) Subcellular fractionation of stably transduced HEK-293 cells was performed as in panel **C**; GAPDH was used as a cytosolic marker. (F) Detergent extraction of CLPB complexes from mitochondria. Mitochondria (M) were isolated from HEK-293 cells with stable expression of CLPB-Myc/FLAG (top) or from parental HEK-293 cells. Mitochondria were then treated with 1% digitonin (Dig) or 1% n-Dodecyl-β-D-Maltoside (DDM) and centrifuged for 30 min at 15,000xg. Lysate (L), supernatant (S), or pellet (P, resuspended in the volume equivalent to the supernatant) were analyzed by anti-Myc (top) or anti CLPB (bottom) antibodies. (G) Confocal images of HEK-293 cells stably overexpressing CLPB-Myc/FLAG (top two rows) or control HEK-293 cells (bottom row), stained with anti Myc Ab (green), anti-COX IV or anti-YME1L1 Ab (red), and DAPI (blue). Fluorescence signal intensities along white double-arrowed lines in two merged images were quantified using the line profile measurement tool in the Zeiss Zen software. Bar, 5 μm and plotted in right panels.

Figure 3.4:

Formation of ATP-dependent complexes by endogenous human CLPB



Mitochondria were isolated from HEK-293 or BT-549 cells, incubated with 1% digitonin in the presence or absence of 5 mM ATP or ATPγS, and centrifuged for 30 min at 15,000xg. The supernatants were analyzed by Blue-Native PAGE (top) or SDS-PAGE (bottom), followed by immunoblotting with anti-CLPB Ab.

Chapter 4 - Disease-linked mutations in Human CLPB

4.1 Introduction

As described in detail in Chapter 1, Section 1.7 (Human CLPB/SKD3 mutations and the associated diseases), mutations in the human *CLPB* gene are the cause of 3-methylglutaconic aciduria with cataracts, neurologic involvement, and neutropenia (MEGCANN; MIM entry # 616271), also referred to as 3-methylglutaconic aciduria (3-MGA-uria) type VII. The goal of the studies described below was to perform the first characterization of the selected disease-linked variants of CLPB during their expression in cultured mammalian cells.

4.2 Materials and methods

4.2.1 Reagents and antibodies

X-tremeGENE HP DNA transfection reagent, aprotinin, pepstatin, leupeptin, phenylmethanesulfonyl fluoride (PMSF), adenosine 5'-triphospahte (ATP) disodium salt (#A2383) were from MilliporeSigma (Burlington, MA). 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF) was from Fisher Scientific. MitoTracker Orange CMTMRos dye (#M7510), NativePAGE sample prep kit (#BN2008) were from ThermoFisher Scientific (Waltham, MA). Anti-Myc mouse mAb (clone 9B11, #2276), anti-FLAG tag rabbit mAb (clone D6W5B, #14793), anti-Cox IV rabbit mAb (clone 3E11, #4850), anti-MEK1/2 rabbit mAb (clone D1A5, #8727), and anti-GAPDH rabbit mAb (clone D16H11, #5174) were from Cell Signaling Technology (Danvers, MA). Cy5- or AlexaFluor 488-conjugated secondary antibodies were from Jackson ImmunoResearch Laboratories, Inc. (West Groove, PA).

4.2.2 Generation of human CLPB disease-linked mutations

Disease-linked mutations were introduced to CLPB-Myc/FLAG (OriGene; #RC203013) using Quickchange Lightning Site-Directed mutagenesis kit (Agilent Technologies; #210518). The mutagenesis primers were as follows:

T268M (FP: 5' GCCAGTTTCAAGGGCTGCATGGCCTTGCACTATGCTG 3';
RP: 5' CAGCATAGTGCAAGGCCATGCAGCCCTTGAAACTGGC 3')

K387T (FP: 5' GGATCATCTGGAATAGGAACAACAGAGCTGGCCAAGC 3'; RP: 5' GCTTGGCCAGCTCTGTTGTTCCTATTCCAGATGATCC 3')

R408G (FP: 5' GCTAAAAAGGGCTTCATCGGGCTGGACATGTCCGAGTTCC 3'; RP: 5' GGAACTCGGACATGTCCAGCCCGATGAAGCCCTTTTTAGC 3')

E455Q (FP: 5'GCTGTGGTGCTCTTTGATCAAGTAGACAAGGCCCATCC 3'; RP: 5'GGATGGGCCTTGTCTACTTGATCAAAGAGCACCACAGC 3')

Y567C (FP: 5'CGGATCAATGAGATCGTCTGCTTCCTCCCCTTCTGCCACTCG 3'; RP: 5'CGAGTGGCAGAAGGGGAGGAAGCAGACGATCTCATTGATCCG 3')

A591V (FP: 5'CTAAACTTCTGGGCCAAGAGAGTCAAGCAAAGGCACAAC 3'; RP: 5'GTTGTGCCTTTGCTTGACTCTCTTGGCCCAGAAGTTTAG 3')

After PCR, 2µl of *Dpn 1* enzyme was added to each PCR product and incubated at 37°C for 2 hours. Then 2 ul of PCR product was used for transformation of NEB 5-α competent *E. coli* cells (New England Biolabs; #C2987H). Transformants were selected with lysogeny broth (LB) agar plates containing 50 µg/mL Kanamycin. Single colonies were picked from plates and grown in LB media supplemented with 50 µg/mL Kanamycin at 37°C overnight. Plasmids were purified using QIAprep Spin Miniprep Kit (QIAGEN; #27104) and sent for sequencing. For

transfection of mammalian cells, plasmids were purified using EndoFree Plasmid Maxi Kit (10) (QIAGEN; #12362)

The K387T and E455Q mutants of CLPB in pLenti-C-Myc-DDK-P2A-Puro vector were created using a similar protocol, but using 34 µg/ml Chloramphenicol instead of Kanamycin as the bacterial selection marker.

4.2.3 Cell culture, transient overexpression of human CLPB, stable overexpression of human CLPB, immunoblotting, confocal microscopy, isolation of mitochondria, and digitonin and DDM treatment

These methods were performed as described in Chapter 3.3.

4.3 Results

Four CLPB disease-linked mutations were selected for this study: T268M, R408G, Y567C, and A591V. Clustal Omega multiple sequence alignment tool was used to check the amino acid conservation of CLPB across different vertebrate species for the selected mutants (NCBI RefSeq numbers are: *Homo_sapiens*, NP_110440.1; *Pan_troglodytes*, XP_001174710.1; *Mus_musculus*, NP_033217.1; *Bos_taurus*, NP_001015625.1; *Equus_caballus*, XP_001917508.1; *Xenopus_tropicalis*, XP_031753289.1). Remarkably, all the four amino acids, T268, R408, Y567, A591, are highly conserved among vertebrate species (Figure 4.1(A) and (C)). The X-ray crystal structure of human CLPB is not available yet, therefore we used the I-TASSER protein structure prediction tool to build 3D models of the ankyrin-repeat domain and the nucleotide binding domain (Figure 4.1(B) and (D)) [192], [193]. The individual ankyrin repeats display helix-loop-helix conformation and the whole ANK domain exhibits a concave

inner surface, which is considered to be the main site for mediating protein-protein interactions [176], [177]. Threonine 268 is located in the helical region of one of the ankyrin-repeat motifs (Figure 4.1(B)) and the threonine hydroxyl group can form a hydrogen bond, which is necessary for the stability of the ankyrin repeat domain [176]. Thus, the amino acid substitution of threonine with methionine might affect the protein stability. The other disease-linked mutations are present in the nucleotide binding domain and are located at different positions (Figure 4.1(D)). Arginine 408 is located adjacent to the Walker B motif, which supports the ATP hydrolysis activity. The substitution of the basic arginine residue with glycine is a significant change of the residue type and it might affect the function of the protein. Tyrosine 567 and alanine 591 are located at a distance from the ATP binding site, which is located between the Walker A, Walker B, and sensor-2 motifs (Figure 4.1(D)). One can speculate that the Y567C and A591V mutations could affect the structure and/or stability of the nucleotide binding domain of CLPB.

The effects of amino acid substitution on protein function were further evaluated using SIFT and PolyPhen scores obtained from the Ensembl Genome Browser (Table 4.1). SIFT is a sequence homology-based tool that sorts intolerant vs tolerant amino acid substitutions and predicts whether an amino acid substitution in a protein will have a phenotypic effect (sift.bii.a-star.edu.sg/www/SIFT_help.html, [194]). PolyPhen-2 is a tool which predicts possible impact of an amino acid substitution on the structure and function of a protein using physical and chemical parameters (genetics.bwh.harvard.edu/pph2, [195]). The SIFT scores <= 0.05 qualify the mutations as “Deleterious”, and the scores > 0.05 as “Tolerated”. The PolyPhen scores >= 0.908 qualify the mutations as “Probably Damaging” i.e., those that can affect protein function or structure with high confidence. According to the SIFT scores (Table 4.1), all selected mutations

except A591V have “deleterious” effects in CLPB, and the PolyPhen scores predicted that all four disease-linked mutations are “probably damaging” for CLPB structure/function.

To functionally characterize the effects of disease-linked mutations in human CLPB, we first transiently overexpressed the wild type CLPB (WT) and the selected mutants in human embryonic kidney HEK-293 cells and human breast cancer cell line SUM-159, using a construct encoding human CLPB bearing a C-terminal Myc/FLAG tag. We observed a ~76-kDa and a ~62-kDa form in both cell lines by immunoblotting (Figure 4.2). Surprisingly, the WT CLPB expression efficiency was low when compared with the mutants (with the exception of T268M). To rule out additional mutations in the CLPB WT construct that might be inadvertently introduced during PCR, the coding sequence in the expression plasmid was fully sequenced and confirmed to contain no extraneous mutations in the construct. The WT plasmid purification was repeated several times and the protein expression was tested with different WT plasmid preparations to make sure the lower expression level of WT CLPB was not due to poor quality of the plasmid preparation. Ultimately, we concluded that the expression of WT CLPB in transiently transfected cells occurred at lower levels than the expression of the mutants.

The K387T and E455Q mutations in CLPB have been included in this study because they produce significant defects in the ATPase activity of AAA+ proteins. The K387T substitution is present in the Walker A motif and it inhibits binding of ATP. The E455Q substitution is in the Walker B motif, it does not inhibit ATP binding, but it suppresses the ATP hydrolysis activity. Since ATP-bound conformations of AAA+ proteins often show high affinity towards their substrates, the Walker B mutant is often called a “trap” variant. These two variants with substitutions within the Walker A and B motifs have been biochemically characterized and previously used in the studies on other AAA+ ATPases [150].

In HEK-293 and SUM-159 cells, the expression levels of K387T and E455Q mutants were much higher than the expression level of WT CLPB. The four disease-linked mutants showed different levels of expression (Figure 4.2). The T268M mutant, where the mutation is present in the ankyrin-repeat domain, was expressed at the lowest levels among all the disease-linked mutants. The 76-kDa form of T268M was noticeable, but the 62-kDa form was almost undetectable. For the R408G mutant, we observed both the 76-kDa and the 62-kDa forms, but the 62-kDa form less pronounced than the 76-kDa form. This was different from the remaining CLPB variants, where both protein forms were equally abundant or the 62-kDa form was even a predominant one. Overall, the disease-linked CLPB mutants that have been selected for this study have variable levels of expression when compared with the WT CLPB.

Next, we investigated the localization of the 76-kDa and the 62-kDa forms of different CLPB mutants inside the cells. Subcellular fractionation showed that the 76-kDa form of WT CLPB in transiently transfected cells resides in the cytosol, while the 62-kDa protein is localized to the mitochondria (Figure 4.3). Thus, the 76-kDa protein likely represented the precursor form of WT CLPB and the 62-kDa protein corresponded to the mature form after the cleavage of the N-terminal mitochondrial targeting presequence by mitochondrial processing peptidase (MPP) and the intramembrane protease PARL. Interestingly, unlike for WT CLPB, low amounts of the 76-kDa mature form were detected in the mitochondrial fractions for all the mutants. This could be due to somewhat inefficient processing of the mature form of all the mutants in the mitochondria. Also, surprisingly, unlike for WT CLPB, low amounts of the 62-kDa mature form was observed in the cytosol for all the mutants. This could indicate that mutations caused some defects in the folding of CLPB during import into mitochondria, leading to its retro-translocation into the cytosol. The retro-translocation of proteins from the mitochondrial intermembrane space

to cytosol has been reported as a quality control mechanism [196]. Alternatively, the presence of the 76-kDa form in the mitochondria and the 62-kDa form in the cytosol may be a result of suboptimal fractionation of mutant-expressing cells.

To examine the production of the mature form of CLPB, we performed a time course experiment for the transiently overexpressed WT CLPB and disease-linked mutants. The 76-kDa precursor mutant form was present for a longer period of time than the 62-kDa mature form of WT CLPB (Figure 4.4). Importantly, the mature form of the T268M mutant declined at 36 hours after transfection and the mature form of the R408G mutant was present in lower amounts than the precursor form till 72 hours post-transfection. These data are consistent with our previous transient overexpression experiment (Figure 4.2), where the samples were collected after 48 hours after transfection. Intriguingly, the opposite effect was observed for the Y567C and A591V mutants. For both mutants, the 62-kDa form was more abundant than the 76-kDa precursor form. This suggested that: a) the precursor form is less stable than the mature form, and/or b) processing by mitochondrial proteases is very efficient for the Y567C and A591V mutants. The protein import machinery into mitochondria is regulated at different levels under physiological and pathophysiological conditions [197]. Thus, it is important to understand how the processing of WT CLPB and CLPB mutants is regulated inside the mitochondria.

To determine the intracellular localization of CLPB, we used immunofluorescence microscopy of transiently transfected HEK-293 and SUM-159 cells. Anti-Myc antibody was used to detect the CLPB protein, MitoTracker orange dye was used to stain the mitochondria, and anti-COX IV antibody was used to detect COX IV protein, which is a known marker of the mitochondrial inner membrane. In transiently transfected HEK-293-CLPB-Myc/FLAG cells, the WT CLPB staining produced a punctuated pattern that did not completely overlap with

mitochondria (Figure 4.5). Furthermore, anti-Myc antibody staining did not fully overlap with anti-COX IV staining. In SUM-159-CLPB-Myc/FLAG cells, the pattern of anti-Myc staining was similar to HEK 293-CLPB-Myc/FLAG cells (Figure 4.6). Interestingly, the K387T and E455Q mutants expressed in SUM-159 cells co-localized better with mitochondria and COX IV staining, but some of the intense anti-Myc puncta were only weakly stained with MitoTracker dye or anti-COX IV antibodies (Figure 4.7 and Figure 4.8). The discrete punctuated pattern of overexpressed CLPB in both cell lines is different from the endogenous CLPB immunofluorescence staining as shown before (Figure 3.1 (C)). The strong puncta-like anti-Myc signals might represent CLPB clusters of the precursor form or conglomerates of the mature form. These results suggested that overexpressed CLPB and the disease-linked mutants might be prone to aggregation in cells.

Next, we established stable overexpression of WT CLPB, K387T or E455Q mutants in HEK-293 cells using lentiviral delivery. Immunoblotting results showed a similar pattern of bands for stably overexpressed CLPB proteins as for transiently expressed CLPB, except for the K387T and E455Q mutants, for which the mature 62-kDa form was present in higher amounts than the precursor 76-kDa form (Figure 4.9(A)). This result suggested that the protein turnover in cells may be different for these mutants than for WT CLPB.

To determine whether overexpressed CLPB might form aggregates inside the mitochondria, mitochondrial membranes were solubilized with neutral non-denaturing detergents digitonin and n-dodecyl-beta-D-maltoside (DDM), and the lysates were fractionated by centrifugation. Supernatant fractions contained water-soluble proteins, multiprotein complexes, and solubilized membrane proteins, whereas pellets contained insoluble membrane protein complexes and protein aggregates [182], [183]. Overexpressed CLPB was found mostly in the

pellets of WT CLPB and the K387T and E455Q mutants (Figure 4.9(B)). The E455Q mutant showed a slightly higher amount of soluble CLPB in the supernatants than the WT and K387T mutant. This result was in contrast to the results obtained for endogenous CLPB, which was described in Chapter 3 (Figure 3.3(F)), suggesting that overexpressed CLPB does not fully mimic the biochemical properties of endogenous CLPB.

4.4 Discussion

We demonstrate that the selected disease-linked variants of human CLPB show different levels of expression than WT CLPB in transiently transfected cells. Specifically, the T268M and R408G mutations result in low amounts of the mature 62-kDa form of CLPB. In contrast, the Y567C and A591V mutations are linked to high levels of the mature CLPB. A recent study reported that the recombinant human CLPB has a higher level of disaggregase activity *in vitro* after the cleavage with PARL, i.e. after a removal of the whole mitochondrial targeting sequence [44]. Thus, the mature 62-kDa form of CLPB is believed to be a physiologically functional form localized in the mitochondrial intermembrane space. The CLPB function has not been fully explored *in vivo* and it is not clear whether it could act as a disaggregase inside the cells. Our time course experiment (Figure 4.4) also shows that the expression and processing of CLPB is different for the disease-linked mutants and WT CLPB. Collectively, our results indicate that the disease-linked mutations may produce different levels of protein expression and processing, which might contribute to a loss of stability and/or activity of CLPB.

The subcellular fractionation clearly demonstrates that the 62-kDa mature form of CLPB is localized to the mitochondria. The differences in the localization of the precursor form and the mature form for the mutants are intriguing and show that the mutated CLPB may undergo regulation by a protein quality control system in the mitochondria [198], [199]. Thus, cells could

potentially avoid harmful effects of the incorrect localization of CLPB and/or the loss of CLPB activity. The WT and mutant CLPB localization to mitochondria is also demonstrated by immunofluorescence experiments (Figure 4.5-4.8). These results suggest that overexpressed WT and mutant CLPB are properly trafficked to mitochondria, but they may be prone to aggregation (see also Chapter 3).

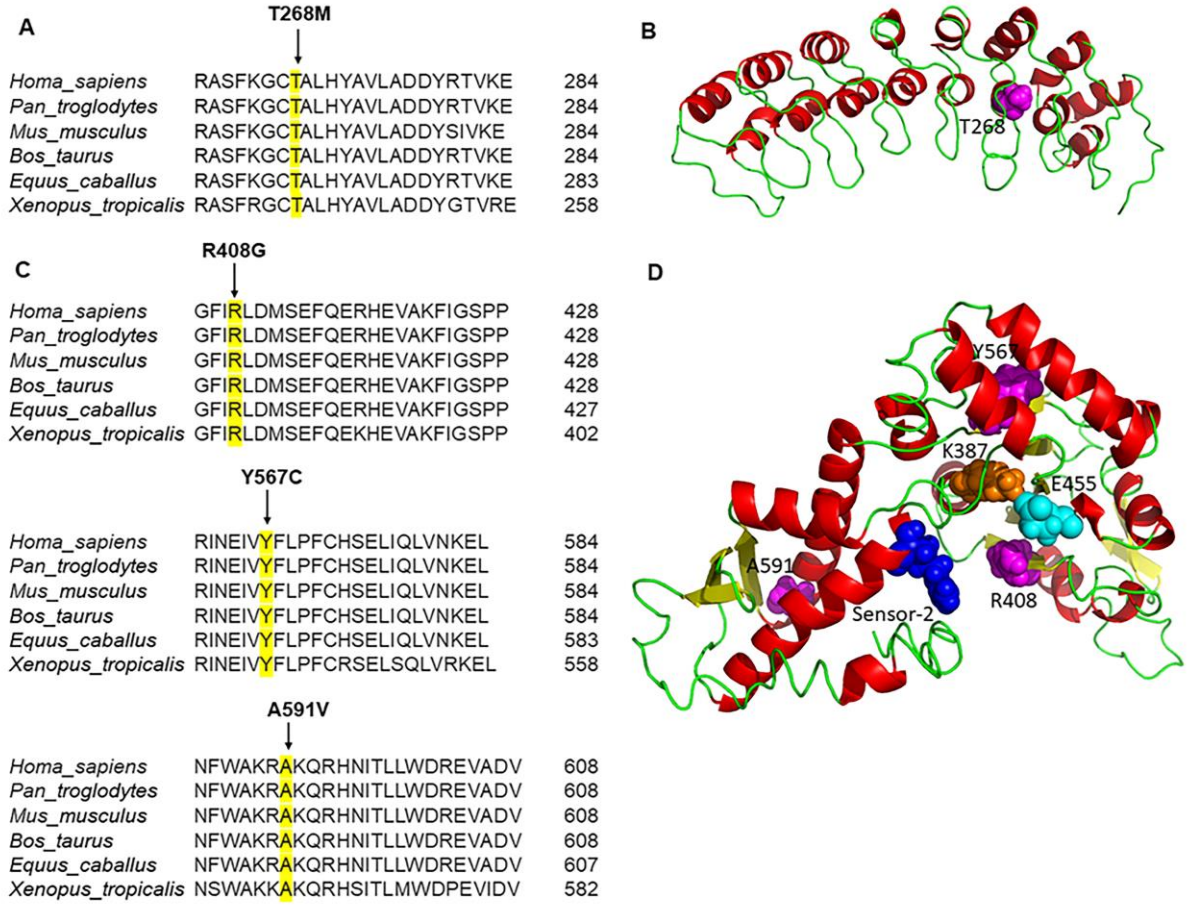
Stable overexpression of CLPB in HEK-293 cells confirmed that the WT CLPB level is low when compared to the K387T and E455Q mutants, which are fully inactive ATPases. One can speculate that cells limit overexpression of an active ATPase in order to preserve energy resources. Importantly, the digitonin and DDM treatment of cells with overexpression of CLPB clearly indicated that most of the overexpressed protein was in the pellet and thus could be a part of large protein aggregates/clusters located in the mitochondria. This behavior does not mimic the biochemical and biological properties of endogenous CLPB (see Chapter 3), which should be taken into consideration in future experiments. Thus, different levels of expression and differences in processing between disease-associated CLPB variants and WT CLPB during overexpression experiments should be treated with caution. The CLPB overexpression levels could be dialed down by using an inducible expression system in the future. Such an expression system may become useful for studying the role of pathogenic mutations in the *CLPB* gene and for expressing individual CLPB mutants in mammalian cells.

SIFT and PolyPhen scores for the T268M, R408G, Y567C and A591V mutations suggest that these mutations could produce a loss of function. A recently published study analyzed the ATPase activity and disaggregase activity of some disease-linked variants (T268M, R475Q, A591V, and R650P) *in vitro* [44]. Among these four disease-linked variants, T268M showed higher ATPase activity than WT CLPB, but had only ~27% disaggregase activity compared to

WT CLPB. The A591V mutant did not show any ATPase activity or disaggregase activity. These results indicate that the disease-linked mutations might have impact on the human CLPB function. It is not clear yet whether CLPB acts as a disaggregase *in vivo*. To further explore the functionality of human CLPB, it is necessary to identify its physiological substrates in the mitochondria. This will help close the gap in knowledge on the biological role of CLPB and the molecular basis of the diseases caused by CLPB deficiency.

Figure 4.1:

Selected CLPB disease-linked mutations



A) Amino acid sequence alignment of CLPB homologs from different vertebrate species in the region containing threonine 268. B) Model of the ankyrin-repeat domain of CLPB generated by the I-TASSER protein structure prediction tool. C) Amino acid sequence alignment of CLPB homologs from different vertebrate species in the regions containing arginine 408, tyrosine 567, and alanine 591. D) Model of the nucleotide binding domain of CLPB generated by the I-TASSER protein structure prediction tool.

Table 4.1:

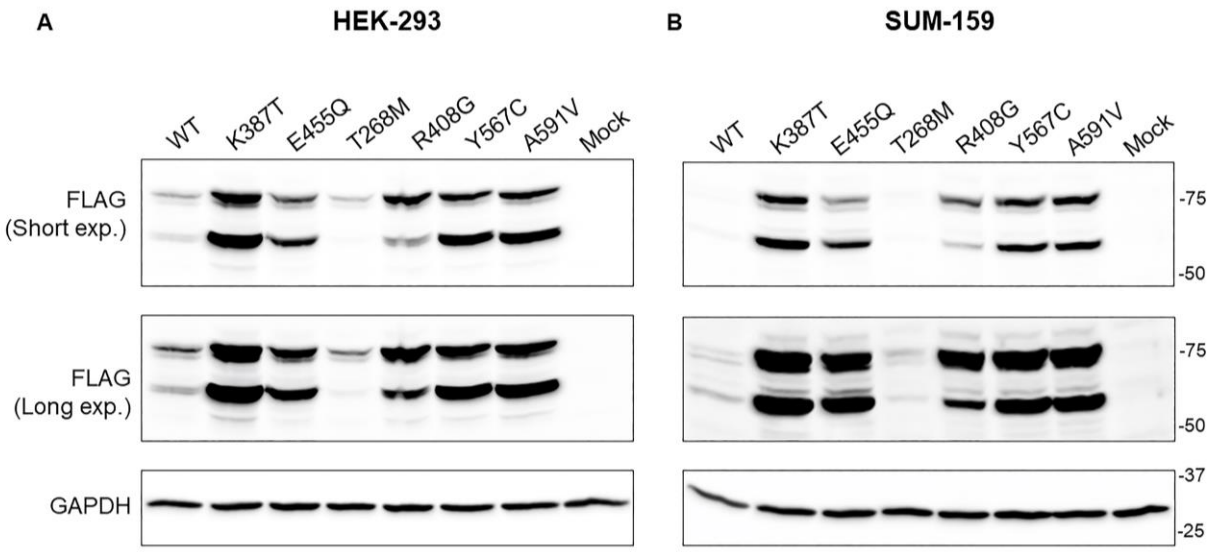
Predicted effects of missense mutations in CLPB

Mutation	Location of the mutation	SIFT score	PolyPhen score
T268M	ANK	0- Deleterious	0.998- Probably damaging
R408G	NBD	0- Deleterious	0.948- Probably damaging
Y567C	NBD	0- Deleterious	0.999- Probably damaging
A591V	NBD	0.09- Tolerated	0.93- Probably damaging

Predicted effects of selected disease-linked mutations on CLPB structure/function according to the SIFT algorithm and PolyPhen-2 algorithm obtained from the Ensembl Genome Browser (www.ensembl.org).

Figure 4.2:

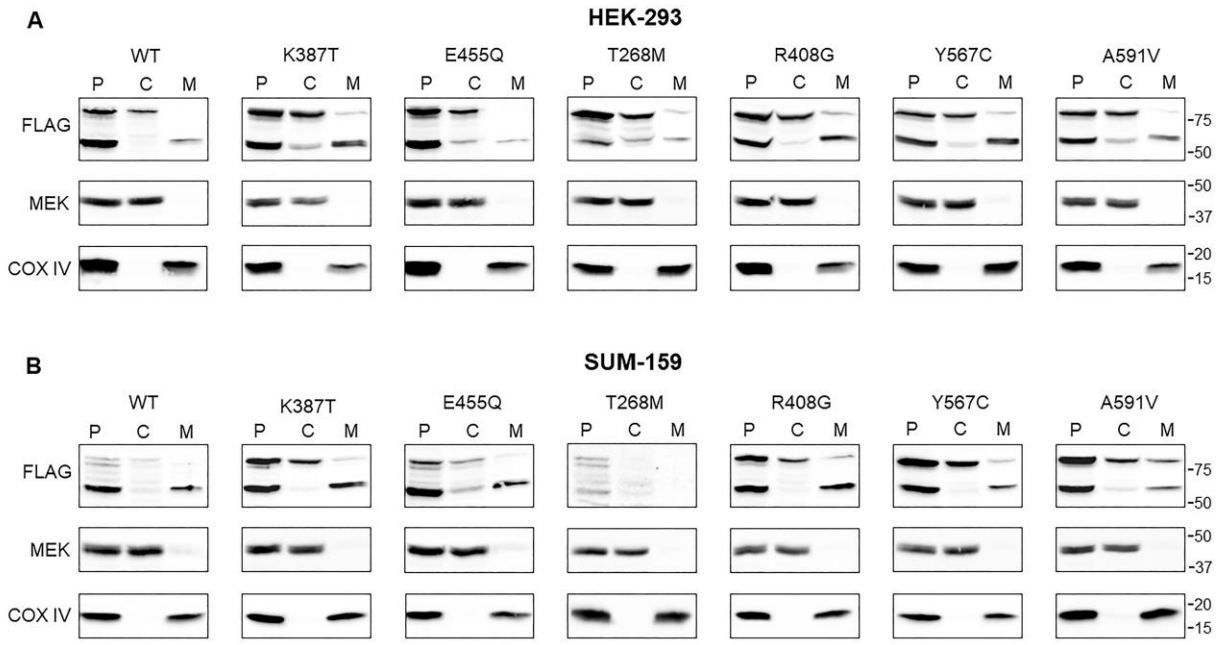
Transient overexpression of CLPB mutants in HEK-293 and SUM-159 cells



Transient overexpression of CLPB-Myc/FLAG in HEK-293 cells (A) or SUM-159 cells (B). Total cell lysates for both cell lines were analyzed by immunoblotting 48 hours after transfection using anti-FLAG antibody; GAPDH is a gel-loading control.

Figure 4.3:

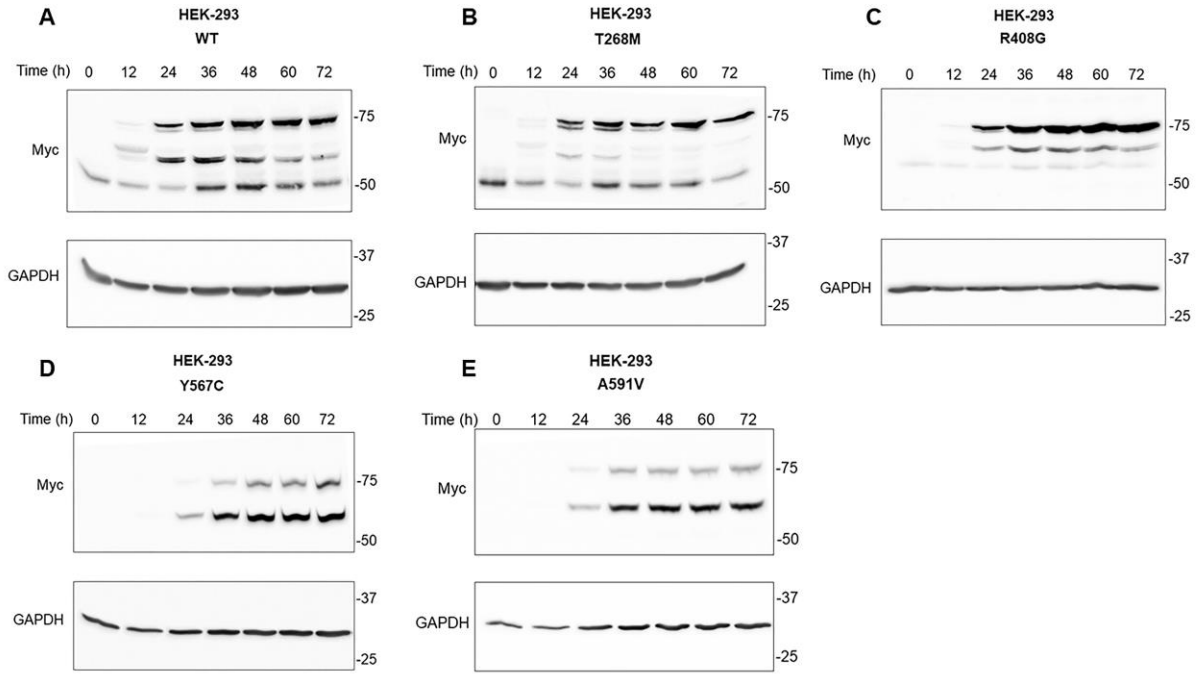
Subcellular fractionation of CLPB mutants in HEK-293 and SUM-159 cells



Subcellular fractionation of HEK-293 cells (A) or SUM-159 cells (B), with transient expression of CLPB-Myc/FLAG. Post-nuclear supernatant (P), cytosolic fraction (C), and mitochondrial fraction (M) were subjected to SDS-PAGE and immunoblotting with anti-FLAG, anti-MEK (a cytosolic marker), or anti-COX IV (a mitochondrial marker) antibodies.

Figure 4.4:

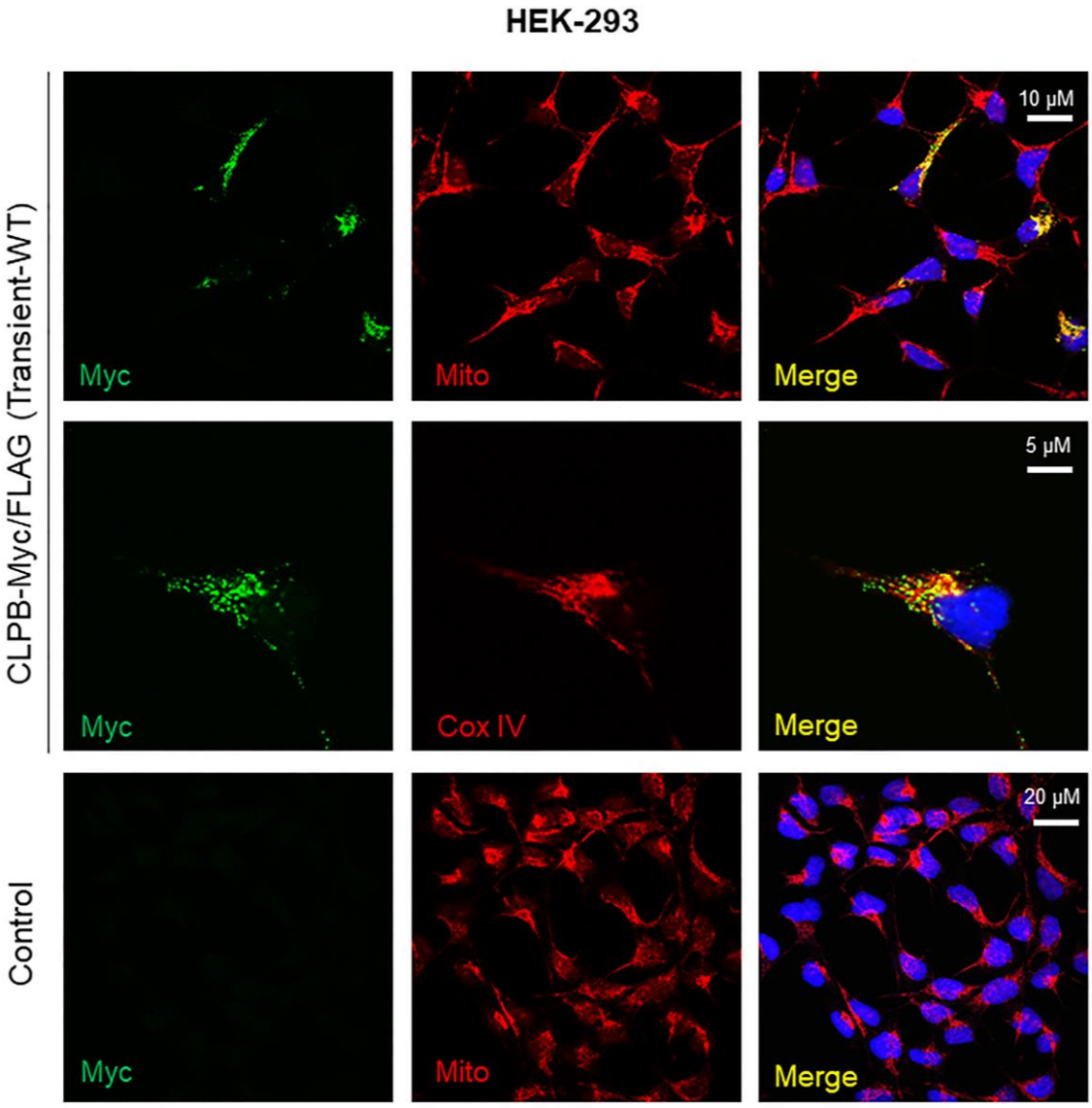
Time course of CLPB mutant expression in transiently transfected HEK-293 cells



Time course of CLPB-Myc/FLAG expression of WT CLPB (A), the T268M mutant (B), the R408G mutant (C), the Y567C mutant (D), and the A591V mutant (E). Total cell lysates were analyzed at indicated times after transfection by immunoblotting with anti-Myc antibody; GAPDH is a gel-loading control.

Figure 4.5:

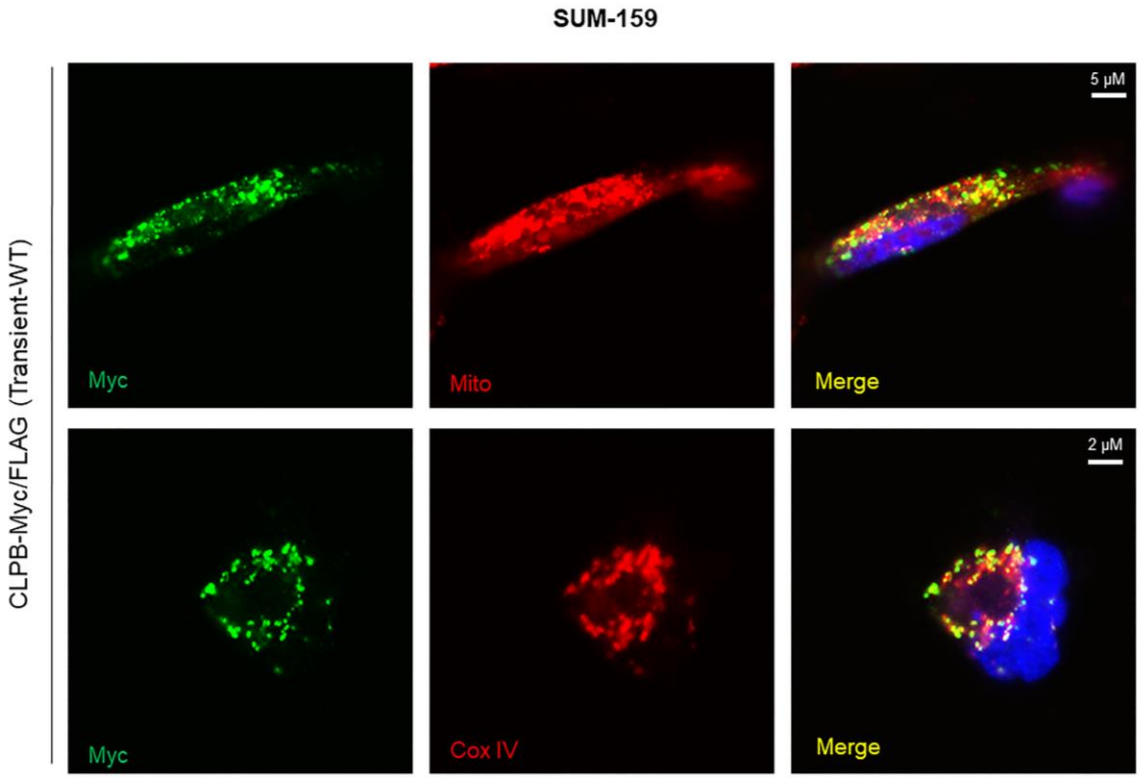
Confocal images of HEK-293 cells transiently overexpressing WT CLPB



Transient overexpression of CLPB-Myc/FLAG (top two rows) or control HEK-293 cells (bottom row), stained with anti-Myc Ab (green), MitoTracker Orange CMTMRos dye or anti-COX IV Ab (red), and DAPI (blue).

Figure 4.6:

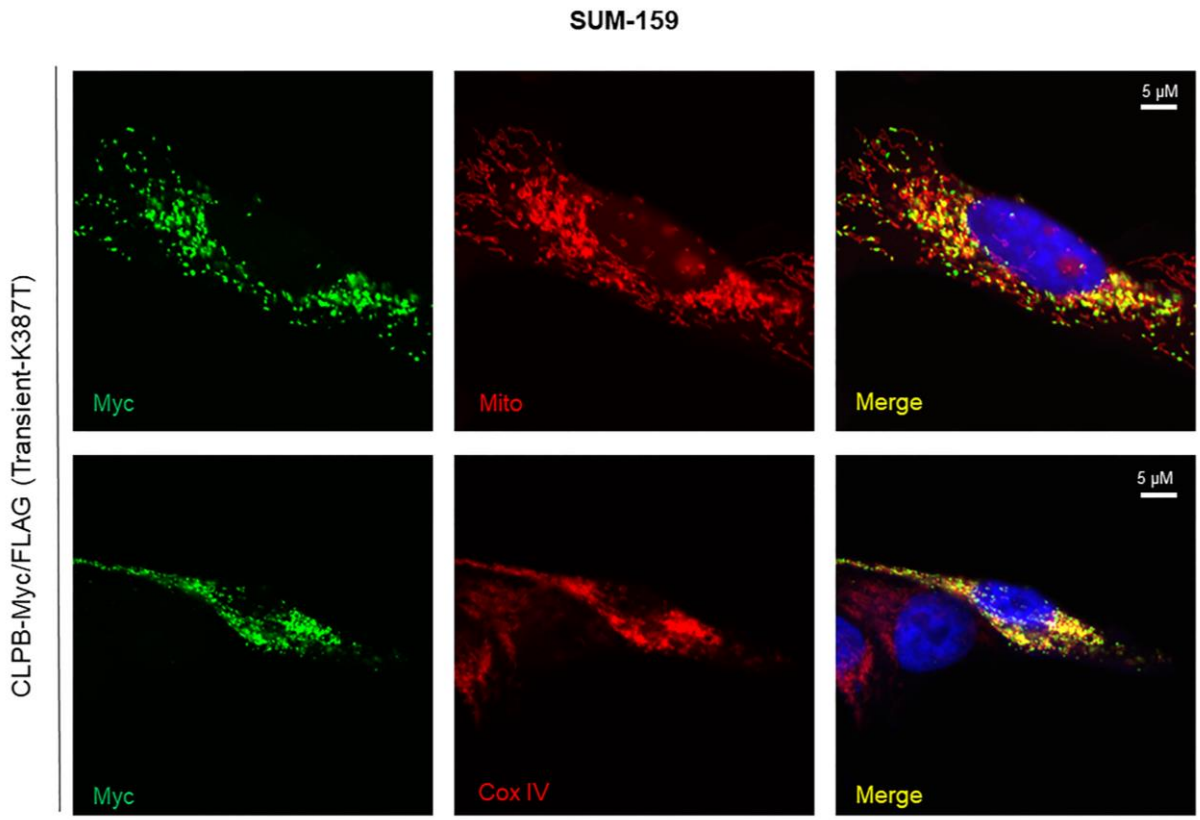
Confocal images of SUM-159 cells transiently overexpressing WT CLPB



Transient overexpression of CLPB-Myc/FLAG, stained with anti-Myc Ab (green), MitoTracker Orange CMTMRos dye or anti-COX IV Ab (red), and DAPI (blue).

Figure 4.7:

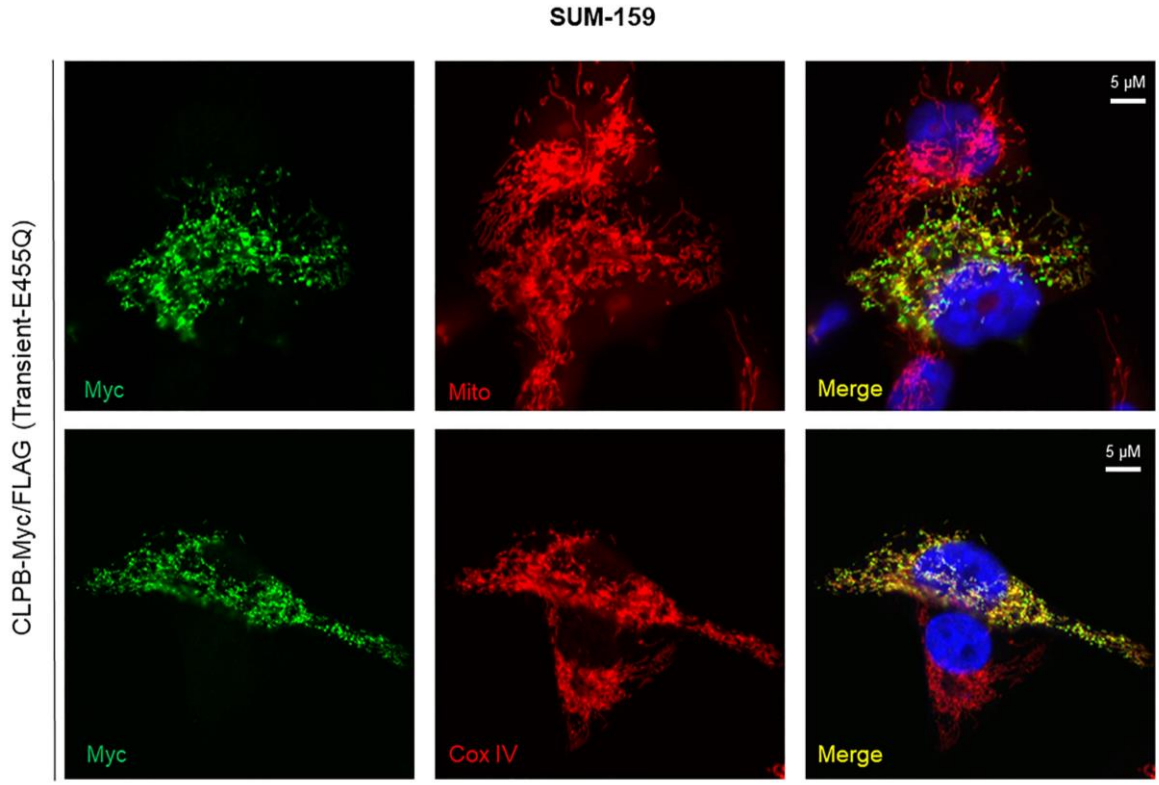
Confocal images of SUM-159 cells transiently overexpressing CLPB mutant K387T



Transient overexpression of CLPB-Myc/FLAG, stained with anti-Myc Ab (green), MitoTracker Orange CMTMRos dye or anti-COX IV Ab (red), and DAPI (blue).

Figure 4.8:

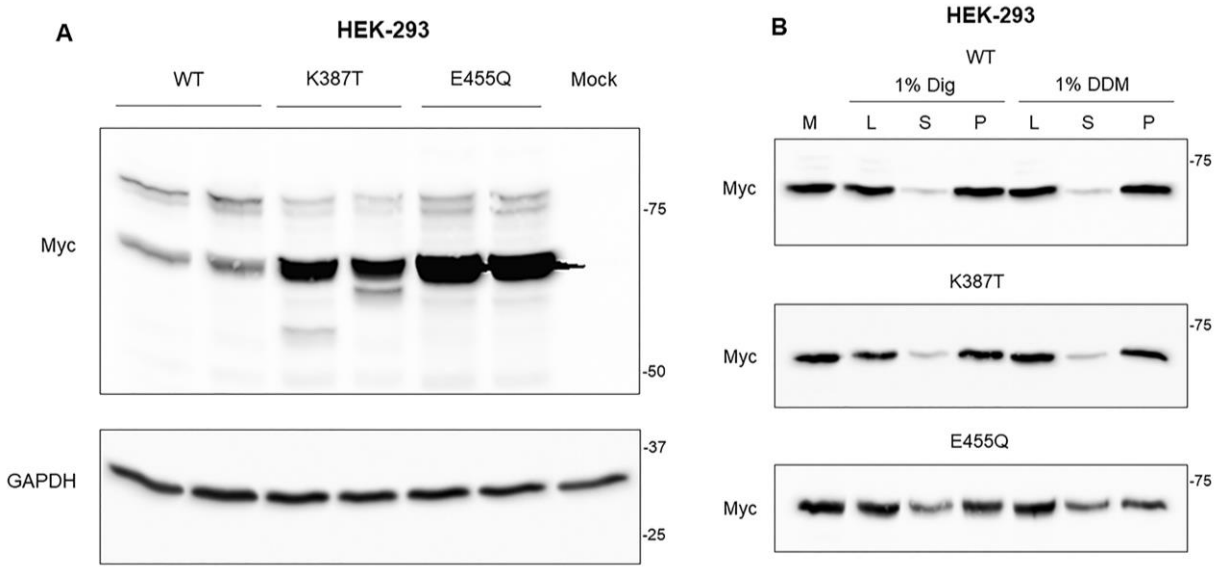
Confocal images of SUM-159 cells transiently overexpressing CLPB mutant E455Q



Transient overexpression of CLPB-Myc/FLAG, stained with anti-Myc Ab (green), MitoTracker Orange CMTMRos dye or anti-COX IV Ab (red), and DAPI (blue).

Figure 4.9:

Stable overexpression of WT CLPB and, the K387T and E455Q mutants in HEK-293 cells



A) Stable overexpression of CLPB-Myc/FLAG (WT, the K387T mutant, or the E455Q mutant) in HEK-293 cells. Total cell lysates were analyzed by immunoblotting with anti-Myc antibody; GAPDH is a gel-loading control. B) Detergent extraction of CLPB complexes from mitochondria. Mitochondria (M) were isolated from HEK-293 cells with stable expression of WT CLPB, the K387T mutant, or the E455Q mutant. Mitochondria were then treated with 1% digitonin (Dig) or 1% n-Dodecyl-β-D-Maltoside (DDM) and centrifuged for 30 min at 15,000xg. Lysate (L), supernatant (S), or pellet (P, resuspended in the volume equivalent to the supernatant) were analyzed by anti-Myc antibody.

Chapter 5 - Conclusions

In this dissertation, I have investigated the effects of missense mutations in two proteins linked to human disease: tumor suppressor p53 and the AAA+ ATPase human CLPB. I introduced selected cancer-linked mutations into the transactivation domain (TAD) of p53 and 3-MGA-uria type VII-associated mutations into human CLPB.

In Chapter 2, I showed that the cancer-linked mutations in the transactivation domain of p53 have impacts on binding with negative regulator HDM2, and general transcriptional co-activator CBP. The N-terminal TAD of p53 is essential for interactions with multiple signaling molecules that mediate the cellular response to genotoxic stresses. The p53 level and activity are regulated by interactions of TAD with HDM2 and CBP. Despite a few decades of p53 research, very little is known about the structural and functional consequence of numerous cancer-associated mutations in p53-TAD. It is thus crucial to understand how various mutations affect p53 at the molecular level.

In this study, I analyzed four p53-TAD cancer-associated mutations; K24N, N29K/N30D, D49Y, and W53G. I purified recombinant human p53-TAD protein of wild type (WT) and the four cancer-associated mutants, which are mentioned above. The p53-TAD interacting partners, such as HDM2 (17-125), CBP-TAZ1 (340-439), and CBP-TAZ2 (1764-1850) were also purified to perform protein binding studies using Biolayer Interferometry (BLI). I observed that N29K/N30D mutant lost its binding with HDM2 and the other mutants were not affected. The interaction with CBP-TAZ2 was affected for N29K/N30D, D49Y, and W53G mutants. The K24N mutant showed ~3 fold weaker binding with CBP-TAZ2 than WT. The CBP-TAZ1 interaction was not affected for all the mutants except W53G. The CBP-TAZ2 binds ~30-fold stronger to p53-TAD than does CBP-TAZ1. These results indicate that the attenuation

of binding of cancer-linked mutants with CBP-TAZ2 may lead to a loss of the p53 tumor suppressor function that can cause cancer. This is the first time that the p53-TAD cancer mutants have been reported at the functional level to understand the impacts of the missense mutations in the transactivation domain.

In the future, the modulation of p53-TAD interaction with HDM2 and CBP in the cancer-linked TAD variants could be further explored by including the effects of multi-site phosphorylation of TAD. The p53-TAD phosphorylation is an important post translational modification. The impact of phosphorylation on the cancer-linked variants could be tested using phosphomimetic approach, where the specific amino acid residues are substituted to mimic a phosphorylated protein. The understanding of the biological consequences by phosphorylation would be helpful to assess the potential cancer intervention strategies.

In chapter 3, I investigated endogenous human caseinolytic peptidase B protein homolog (CLPB) in HEK-293 and BT-549 cells. CLPB belongs to the AAA+ family of ATPases associated with different activities. The function of CLPB is unknown and its relation to the well-studied disaggregase ClpB/Hsp104 has been controversial. Unlike the microbial ClpB/Hsp104 protein family, human CLPB contains a N-terminal mitochondrial targeting presequence, followed by ankyrin-repeat domain, and a single ATP-binding AAA+ module.

I demonstrated that CLPB localizes to the mitochondrial intermembrane space (IMS). The accessibility of CLPB to externally added proteinase K (PK) was examined and the results showed that CLPB was protected against PK in intact mitochondria but was degraded upon disruption of the outer membrane. Immunofluorescence experiment also showed that the endogenous CLPB was predominantly located in the mitochondria. The exogenously expressed CLPB were examined by transient and stable overexpression in HEK 293 cells. The results

showed that CLPB produces two forms; the precursor form and the mature form. I found that the precursor form was present in the cytosol and the mature form localized to the mitochondria. The precursor form likely represented the full-length protein and the mature form corresponded to the protein after the cleavage of the N-terminal mitochondrial targeting presequence by mitochondrial processing peptidase (MPP) and the intramembrane protease PARL. Unexpectedly, we observed that the overexpressed CLPB formed large aggregates/protein complexes that were poorly extractable with non-ionic detergents and were readily visualized by immunofluorescence as a punctuate pattern, which did not completely overlap with mitochondria. Importantly, another discovery was that the endogenous CLPB formed high molecular weight protein complexes in an ATP-dependent manner. It is not known at this time whether these protein complexes represented CLPB homo-oligomers or hetero-complexes with other mitochondrial proteins.

In Chapter 4, I investigated the selected disease-linked CLPB variants and found that they showed different levels of expression than WT CLPB. The T268M and R408G mutants produce low amounts of the mature CLPB, while the Y567C and A591V mutants produce high levels of the mature protein form. The mature form of CLPB is believed to be the physiologically functional form localized in the mitochondrial intermembrane space. Overall, the selected disease-linked variants are characterized by different levels of protein expression and processing, which might contribute to a loss of stability and/or activity of CLPB. This also suggests that the mutated CLPB may be regulated by protein quality control system in the mitochondria to avoid any harmful effects from the incorrect localization of CLPB. The digitonin and DDM treatment of overexpressed CLPB clearly showed that it is prone to aggregation and it does not mimic the biochemical and biological properties of endogenous CLPB. This phenomenon should be

considered in future experiments when expressing individual CLPB mutants in mammalian cells to study the role of pathogenic mutations in the CLPB gene.

In the future, the CLPB disease-linked variants should be investigated with lower levels of expression using an inducible expression system to mimic the expression level of endogenous CLPB. The protein turnover could be examined for all the disease-linked variants in a stable cell line to understand how the stabilities of the mutated proteins differ from WT CLPB. The apparent degradation of mature form of CLPB for some disease-linked variants is intriguing and it could be further tested with specific mitochondrial protease inhibitors to understand the degradation of CLPB by proteases in the mitochondria. The interacting partners of CLPB also could be investigated by pull down assay and mass spectrometry analysis. Collectively, these findings should inform the effect of disease-causing mutations on CLPB function and on mechanism of CLPB action in human cells.

References

[1] S. Clancy, “A single base change can create a devastating genetic disorder or a beneficial adaptation, or it might have no effect. How do mutations happen, and how do they influence the future of a species?,” Nat. Educ., p. 12, 2008.

[2] B. A. Pierce, Genetics: a conceptual approach, 4th ed. New York: W.H. Freeman, 2012.

[3] H. Seidl et al., “Ultraviolet Exposure as the Main Initiator of p53 Mutations in Basal Cell Carcinomas from Psoralen and Ultraviolet A-Treated Patients with Psoriasis,” J. Invest. Dermatol., vol. 117, no. 2, pp. 365–370, Aug. 2001, doi: 10.1046/j.0022-202x.2001.01413.x.

[4] L. E. Rosenberg and D. D. Rosenberg, “Mutation,” in Human Genes and Genomes, Elsevier, 2012, pp. 117–124.

[5] M. R. Lieber, “The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway,” Annu. Rev. Biochem., vol. 79, no. 1, pp. 181–211, Jun. 2010, doi: 10.1146/annurev.biochem.052308.093131.

[6] I. Martincorena and P. J. Campbell, “Somatic mutation in cancer and normal cells,” Science, vol. 349, no. 6255, pp. 1483–1489, Sep. 2015, doi: 10.1126/science.aab4082.

[7] J. M. Lancaster, C. B. Powell, L. Chen, and D. L. Richardson, “Society of Gynecologic Oncology statement on risk assessment for inherited gynecologic cancer predispositions,” Gynecol. Oncol., vol. 136, no. 1, pp. 3–7, Jan. 2015, doi: 10.1016/j.ygyno.2014.09.009.

[8] B. Friedenson, “The BRCA1/2 pathway prevents hematologic cancers in addition to breast and ovarian cancers,” BMC Cancer, vol. 7, p. 152, Aug. 2007, doi: 10.1186/1471-2407-7-152.

[9] J. R. Pon and M. A. Marra, “Driver and Passenger Mutations in Cancer,” Annu. Rev. Pathol. Mech. Dis., vol. 10, no. 1, pp. 25–50, Jan. 2015, doi: 10.1146/annurev-pathol-012414-040312.

[10] N. Rivlin, R. Brosh, M. Oren, and V. Rotter, “Mutations in the p53 Tumor Suppressor Gene: Important Milestones at the Various Steps of Tumorigenesis,” Genes Cancer, vol. 2, no. 4, pp. 466–474, Apr. 2011, doi: 10.1177/1947601911408889.

[11] S. Kato et al., “Understanding the function–structure and function–mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis,” Proc. Natl. Acad. Sci., vol. 100, no. 14, pp. 8424–8429, Jul. 2003, doi: 10.1073/pnas.1431692100.

[12] W. Mcbride, D. Merry, and D. GIVOLt, “The gene for human p53 cellular tumor antigen is located on chromosome 17 short arm (17pl3),” Proc Natl Acad Sci USA, p. 5, 1986.

[13] T. Saha, R. K. Kar, and G. Sa, “Structural and sequential context of p53: A review of experimental and theoretical evidence,” Prog. Biophys. Mol. Biol., vol. 117, no. 2, pp. 250–263, Mar. 2015, doi: 10.1016/j.pbiomolbio.2014.12.002.

[14] A. M. Bode and Z. Dong, “Post-translational modification of p53 in tumorigenesis,” Nat. Rev. Cancer, vol. 4, no. 10, Art. no. 10, Oct. 2004, doi: 10.1038/nrc1455.

[15] A. W. Braithwaite, G. Del Sal, and X. Lu, “Some p53-binding proteins that can function as arbiters of life and death,” Cell Death Differ., vol. 13, no. 6, Art. no. 6, Jun. 2006, doi: 10.1038/sj.cdd.4401924.

[16] J. C. Ferreon, C. W. Lee, M. Arai, M. A. Martinez-Yamout, H. J. Dyson, and P. E. Wright, “Cooperative regulation of p53 by modulation of ternary complex formation with CBP/p300 and HDM2,” Proc. Natl. Acad. Sci., vol. 106, no. 16, pp. 6591–6596, Apr. 2009, doi: 10.1073/pnas.0811023106.

[17] R. Dawson, L. Müller, A. Dehner, C. Klein, H. Kessler, and J. Buchner, “The N-terminal Domain of p53 is Natively Unfolded,” J. Mol. Biol., vol. 332, no. 5, pp. 1131–1141, Oct. 2003, doi: 10.1016/j.jmb.2003.08.008.

[18] A. S. Krois, J. C. Ferreon, M. A. Martinez-Yamout, H. J. Dyson, and P. E. Wright, “Recognition of the disordered p53 transactivation domain by the transcriptional adapter zinc finger domains of CREB-binding protein,” Proc. Natl. Acad. Sci., vol. 113, no. 13, pp. E1853–E1862, Mar. 2016, doi: 10.1073/pnas.1602487113.

[19] N. W. Fischer, A. Prodeus, D. Malkin, and J. Gariépy, “p53 oligomerization status modulates cell fate decisions between growth, arrest and apoptosis,” Cell Cycle, vol. 15, no. 23, pp. 3210–3219, Dec. 2016, doi: 10.1080/15384101.2016.1241917.

[20] A. Hafner, M. L. Bulyk, A. Jambhekar, and G. Lahav, “The multiple mechanisms that regulate p53 activity and cell fate,” Nat. Rev. Mol. Cell Biol., vol. 20, no. 4, Art. no. 4, Apr. 2019, doi: 10.1038/s41580-019-0110-x.

[21] M. Hollstein, D. Sidransky, B. Vogelstein, and C. Harris, “p53 mutations in human cancers,” Science, vol. 253, no. 5015, pp. 49–53, Jul. 1991, doi: 10.1126/science.1905840.

[22] M. Olivier, R. Eeles, M. Hollstein, M. A. Khan, C. C. Harris, and P. Hainaut, “The IARC TP53 database: New online mutation analysis and recommendations to users,” Hum. Mutat., vol. 19, no. 6, pp. 607–614, Jun. 2002, doi: 10.1002/humu.10081.

[23] T. Soussi, K. Dehouche, and C. Béroud, “p53 Website and Analysis of p53 Gene Mutations in Human Cancer: Forging a Link Between Epidemiology and Carcinogenesis,” p. 9.

[24] M. Olivier et al., “Li-Fraumeni and Related Syndromes: Correlation between Tumor Type, Family Structure, and TP53 Genotype,” Cancer Res., vol. 63, no. 20, pp. 6643–6650, Oct. 2003.

[25] M. Olivier, M. Hollstein, and P. Hainaut, “TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use,” Cold Spring Harb. Perspect. Biol., vol. 2, no. 1, pp. a001008–a001008, Jan. 2010, doi: 10.1101/cshperspect.a001008.

[26] M. A. Resnick and A. Inga, “Functional mutants of the sequence-specific transcription factor p53 and implications for master genes of diversity,” Proc. Natl. Acad. Sci., vol. 100, no. 17, pp. 9934–9939, Aug. 2003, doi: 10.1073/pnas.1633803100.

[27] L. R. Dearth et al., “Inactive full-length p53 mutants lacking dominant wild-type p53 inhibition highlight loss of heterozygosity as an important aspect of p53 status in human cancers,” Carcinogenesis, vol. 28, no. 2, pp. 289–298, Feb. 2007, doi: 10.1093/carcin/bgl132.

[28] A. C. Joerger and A. R. Fersht, “Structure–function–rescue: the diverse nature of common p53 cancer mutants,” Oncogene, vol. 26, no. 15, Art. no. 15, Apr. 2007, doi: 10.1038/sj.onc.1210291.

[29] Y. Cho, S. Gorina, P. D. Jeffrey, and N. P. Pavletich, “Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations,” Science, vol. 265, no. 5170, pp. 346–355, Jul. 1994, doi: 10.1126/science.8023157.

[30] P. E. Wright and H. J. Dyson, “Intrinsically disordered proteins in cellular signalling and regulation,” Nat. Rev. Mol. Cell Biol., vol. 16, no. 1, Art. no. 1, Jan. 2015, doi: 10.1038/nrm3920.

[31] A. Petitjean et al., “Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database,” Hum. Mutat., vol. 28, no. 6, pp. 622–629, 2007, doi: 10.1002/humu.20495.

[32] D. Ganguly and J. Chen, “Modulation of the Disordered Conformational Ensembles of the p53 Transactivation Domain by Cancer-Associated Mutations,” PLOS Comput. Biol., vol. 11, no. 4, p. e1004247, Apr. 2015, doi: 10.1371/journal.pcbi.1004247.

[33] A. Petitjean, M. I. W. Achatz, A. L. Borresen-Dale, P. Hainaut, and M. Olivier, “TP53 mutations in human cancers: functional selection and impact on cancer prognosis and outcomes,” Oncogene, vol. 26, no. 15, Art. no. 15, Apr. 2007, doi: 10.1038/sj.onc.1210302.

[34] M. Olivier et al., “The clinical value of somatic TP53 gene mutations in 1,794 patients with breast cancer.,” Clin. Cancer Res., vol. 12, no. 4, pp. 1157–1167, Feb. 2006, doi: 10.1158/1078-0432.CCR-05-1029.

[35] F. Périer, C. M. Radeke, K. F. Raab-Graham, and C. A. Vandenberg, “Expression of a putative ATPase suppresses the growth defect of a yeast potassium transport mutant: identification of a mammalian member of the Clp/HSP 104 family,” Gene, vol. 152, no. 2, pp. 157–163, Jan. 1995, doi: 10.1016/0378-1119(94)00697-Q.

[36] C. Saunders et al., “CLPB Variants Associated with Autosomal-Recessive Mitochondrial Disorder with Cataract, Neutropenia, Epilepsy, and Methylglutaconic Aciduria,” Am. J. Hum. Genet., vol. 96, no. 2, pp. 258–265, Feb. 2015, doi: 10.1016/j.ajhg.2014.12.020.

[37] S. B. Wortmann et al., “CLPB mutations cause 3-methylglutaconic aciduria, progressive brain atrophy, intellectual disability, congenital neutropenia, cataracts, movement disorder,” Am. J. Hum. Genet., vol. 96, no. 2, pp. 245–257, Feb. 2015, doi: 10.1016/j.ajhg.2014.12.013.

[38] J.-M. Capo-Chichi et al., “Disruption of CLPB is associated with congenital microcephaly, severe encephalopathy and 3-methylglutaconic aciduria,” J. Med. Genet., vol. 52, no. 5, pp. 303–311, May 2015, doi: 10.1136/jmedgenet-2014-102952.

[39] M. Kanabus et al., “Bi-allelic CLPB mutations cause cataract, renal cysts, nephrocalcinosis and 3-methylglutaconic aciduria, a novel disorder of mitochondrial protein disaggregation,” J. Inherit. Metab. Dis., vol. 38, no. 2, pp. 211–219, Mar. 2015, doi: 10.1007/s10545-015-9813-0.

[40] A. Kiykim et al., “Novel CLPB mutation in a patient with 3-methylglutaconic aciduria causing severe neurological involvement and congenital neutropenia,” Clin. Immunol., vol. 165, pp. 1–3, Apr. 2016, doi: 10.1016/j.clim.2016.02.008.

[41] E. Pronicka et al., “A scoring system predicting the clinical course of CLPB defect based on the foetal and neonatal presentation of 31 patients,” J. Inherit. Metab. Dis., vol. 40, no. 6, pp. 853–860, 2017, doi: 10.1007/s10545-017-0057-z.

[42] S. B. Wortmann, R. A. Wevers, and A. P. de Brouwer, “CLPB Deficiency,” in GeneReviews®, M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. Bean, K. Stephens, and A. Amemiya, Eds. Seattle (WA): University of Washington, Seattle, 1993.

[43] S. M. Doyle and S. Wickner, “Hsp104 and ClpB: protein disaggregating machines,” Trends Biochem. Sci., vol. 34, no. 1, pp. 40–48, Jan. 2009, doi: 10.1016/j.tibs.2008.09.010.

[44] R. R. Cupo and J. Shorter, “Skd3 (human ClpB) is a potent mitochondrial protein disaggregase that is inactivated by 3-methylglutaconic aciduria-linked mutations,” eLife, vol. 9, p. e55279, Jun. 2020, doi: 10.7554/eLife.55279.

[45] A. Mogk, C. Schlieker, C. Strub, W. Rist, J. Weibezahn, and B. Bukau, “Roles of Individual Domains and Conserved Motifs of the AAA+ Chaperone ClpB in Oligomerization, ATP Hydrolysis, and Chaperone Activity,” J. Biol. Chem., vol. 278, no. 20, pp. 17615–17624, May 2003, doi: 10.1074/jbc.M209686200.

[46] V. Akoev, E. P. Gogol, M. E. Barnett, and M. Zolkiewski, “Nucleotide-induced switch in oligomerization of the AAA+ ATPase ClpB,” Protein Sci., vol. 13, no. 3, pp. 567–574, 2004, doi: 10.1110/ps.03422604.

[47] X. Chen et al., “Targeting Mitochondrial Structure Sensitizes Acute Myeloid Leukemia to Venetoclax Treatment,” Cancer Discov., vol. 9, no. 7, pp. 890–909, Jul. 2019, doi: 10.1158/2159-8290.CD-19-0117.

[48] T. Yoshinaka et al., “Structural Basis of Mitochondrial Scaffolds by Prohibitin Complexes: Insight into a Role of the Coiled-Coil Region,” iScience, vol. 19, pp. 1065–1078, Sep. 2019, doi: 10.1016/j.isci.2019.08.056.

[49] P. E. Wright and H. J. Dyson, “Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm,” J. Mol. Biol., vol. 293, no. 2, pp. 321–331, Oct. 1999, doi: 10.1006/jmbi.1999.3110.

[50] P. Tompa, “Intrinsically unstructured proteins,” Trends Biochem. Sci., vol. 27, no. 10, pp. 527–533, Oct. 2002, doi: 10.1016/S0968-0004(02)02169-2.

[51] A. K. Dunker, M. S. Cortese, P. Romero, L. M. Iakoucheva, and V. N. Uversky, “Flexible nets,” FEBS J., vol. 272, no. 20, pp. 5129–5148, 2005, doi: 10.1111/j.1742-4658.2005.04948.x.

[52] H. J. Dyson and P. E. Wright, “Intrinsically unstructured proteins and their functions,” Nat. Rev. Mol. Cell Biol., vol. 6, no. 3, Art. no. 3, Mar. 2005, doi: 10.1038/nrm1589.

[53] J. Habchi, P. Tompa, S. Longhi, and V. N. Uversky, “Introducing Protein Intrinsic Disorder,” Chem. Rev., vol. 114, no. 13, pp. 6561–6588, Jul. 2014, doi: 10.1021/cr400514h.

[54] V. N. Uversky, J. R. Gillespie, and A. L. Fink, “Why are ‘natively unfolded’ proteins unstructured under physiologic conditions?,” Proteins Struct. Funct. Bioinforma., vol. 41, no. 3, pp. 415–427, 2000, doi: 10.1002/1097-0134(20001115)41:3<415::AID-PROT130>3.0.CO;2-7.

[55] P. Romero, Z. Obradovic, X. Li, E. C. Garner, C. J. Brown, and A. K. Dunker, “Sequence complexity of disordered protein,” Proteins Struct. Funct. Bioinforma., vol. 42, no. 1, pp. 38–48, 2001, doi: 10.1002/1097-0134(20010101)42:1<38::AID-PROT50>3.0.CO;2-3.

[56] R. K. Das, K. M. Ruff, and R. V. Pappu, “Relating sequence encoded information to form and function of intrinsically disordered proteins,” Curr. Opin. Struct. Biol., vol. 32, pp. 102–112, Jun. 2015, doi: 10.1016/j.sbi.2015.03.008.

[57] M. Varadi et al., “pE-DB: a database of structural ensembles of intrinsically disordered and of unfolded proteins,” Nucleic Acids Res., vol. 42, no. D1, pp. D326–D335, Jan. 2014, doi: 10.1093/nar/gkt960.

[58] A. C. M. Ferreon, J. C. Ferreon, P. E. Wright, and A. A. Deniz, “Modulation of allostery by protein intrinsic disorder,” Nature, vol. 498, no. 7454, Art. no. 7454, Jun. 2013, doi: 10.1038/nature12294.

[59] A. Garcia-Pino et al., “Allostery and Intrinsic Disorder Mediate Transcription Regulation by Conditional Cooperativity,” Cell, vol. 142, no. 1, pp. 101–111, Jul. 2010, doi: 10.1016/j.cell.2010.05.039.

[60] R. B. Berlow, H. J. Dyson, and P. E. Wright, “Expanding the Paradigm: Intrinsically Disordered Proteins and Allosteric Regulation,” J. Mol. Biol., vol. 430, no. 16, pp. 2309–2320, Aug. 2018, doi: 10.1016/j.jmb.2018.04.003.

[61] V. J. Hilser and E. B. Thompson, “Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins,” Proc. Natl. Acad. Sci., vol. 104, no. 20, pp. 8311–8315, May 2007, doi: 10.1073/pnas.0700329104.

[62] J. Chen, “Towards the physical basis of how intrinsic disorder mediates protein function,” Arch. Biochem. Biophys., vol. 524, no. 2, pp. 123–131, Aug. 2012, doi: 10.1016/j.abb.2012.04.024.

[63] V. N. Uversky, C. J. Oldfield, and A. K. Dunker, “Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling,” J. Mol. Recognit., vol. 18, no. 5, pp. 343–384, 2005, doi: 10.1002/jmr.747.

[64] V. Csizmok, A. V. Follis, R. W. Kriwacki, and J. D. Forman-Kay, “Dynamic Protein Interaction Networks and New Structural Paradigms in Signaling,” Chem. Rev., vol. 116, no. 11, pp. 6424–6462, Jun. 2016, doi: 10.1021/acs.chemrev.5b00548.

[65] R. G. Smock and L. M. Gierasch, “Sending Signals Dynamically,” Science, vol. 324, no. 5924, pp. 198–203, Apr. 2009, doi: 10.1126/science.1169377.

[66] M. Fuxreiter, “Fuzziness in Protein Interactions—A Historical Perspective,” J. Mol. Biol., vol. 430, no. 16, pp. 2278–2287, Aug. 2018, doi: 10.1016/j.jmb.2018.02.015.

[67] D. M. Mitrea and R. W. Kriwacki, “Regulated unfolding of proteins in signaling,” FEBS Lett., vol. 587, no. 8, pp. 1081–1088, Apr. 2013, doi: 10.1016/j.febslet.2013.02.024.

[68] P. E. Wright and H. J. Dyson, “Linking folding and binding,” Curr. Opin. Struct. Biol., vol. 19, no. 1, pp. 31–38, Feb. 2009, doi: 10.1016/j.sbi.2008.12.003.

[69] A. K. Dunker et al., “Intrinsically disordered protein,” J. Mol. Graph. Model., vol. 19, no. 1, pp. 26–59, Feb. 2001, doi: 10.1016/S1093-3263(00)00138-8.

[70] J. Liu, N. B. Perumal, C. J. Oldfield, E. W. Su, V. N. Uversky, and A. K. Dunker, “Intrinsic disorder in transcription factors,” Biochemistry, vol. 45, no. 22, pp. 6873–6888, Jun. 2006, doi: 10.1021/bi0602718.

[71] V. N. Uversky et al., “Pathological Unfoldomics of Uncontrolled Chaos: Intrinsically Disordered Proteins and Human Diseases,” Chem. Rev., vol. 114, no. 13, pp. 6844–6879, Jul. 2014, doi: 10.1021/cr400713r.

[72] V. N. Uversky, C. J. Oldfield, and A. K. Dunker, “Intrinsically Disordered Proteins in Human Diseases: Introducing the D2 Concept,” Annu. Rev. Biophys., vol. 37, no. 1, pp. 215–246, 2008, doi: 10.1146/annurev.biophys.37.032807.125924.

[73] M. M. Babu, R. van der Lee, N. S. de Groot, and J. Gsponer, “Intrinsically disordered proteins: regulation and disease,” Curr. Opin. Struct. Biol., vol. 21, no. 3, pp. 432–440, Jun. 2011, doi: 10.1016/j.sbi.2011.03.011.

[74] J. H. Bushweller, “Targeting transcription factors in cancer — from undruggable to reality,” Nat. Rev. Cancer, vol. 19, no. 11, Art. no. 11, Nov. 2019, doi: 10.1038/s41568-019-0196-7.

[75] P. A. J. Muller and K. H. Vousden, “Mutant p53 in Cancer: New Functions and Therapeutic Opportunities,” Cancer Cell, vol. 25, no. 3, pp. 304–317, Mar. 2014, doi: 10.1016/j.ccr.2014.01.021.

[76] C. Kandoth et al., “Mutational landscape and significance across 12 major cancer types,” Nature, vol. 502, no. 7471, Art. no. 7471, Oct. 2013, doi: 10.1038/nature12634.

[77] M. S. Lawrence et al., “Discovery and saturation analysis of cancer genes across 21 tumour types,” Nature, vol. 505, no. 7484, Art. no. 7484, Jan. 2014, doi: 10.1038/nature12912.

[78] P. H. Kussie et al., “Structure of the MDM2 Oncoprotein Bound to the p53 Tumor Suppressor Transactivation Domain,” Science, vol. 274, no. 5289, pp. 948–953, Nov. 1996, doi: 10.1126/science.274.5289.948.

[79] S.-Y. Shieh, M. Ikeda, Y. Taya, and C. Prives, “DNA Damage-Induced Phosphorylation of p53 Alleviates Inhibition by MDM2,” Cell, vol. 91, no. 3, pp. 325–334, Oct. 1997, doi: 10.1016/S0092-8674(00)80416-X.

[80] A. L. Craig, L. Burch, B. Vojtesek, J. Mikutowska, A. Thompson, and T. R. Hupp, “Novel phosphorylation sites of human tumour suppressor protein p53 at Ser20 and Thr18 that disrupt the binding of mdm2 (mouse double minute 2) protein are modified in human cancers,” Biochem. J., vol. 342, no. 1, pp. 133–141, Aug. 1999, doi: 10.1042/bj3420133.

[81] O. Schon, A. Friedler, M. Bycroft, S. M. V. Freund, and A. R. Fersht, “Molecular Mechanism of the Interaction between MDM2 and p53,” J. Mol. Biol., vol. 323, no. 3, pp. 491–501, Oct. 2002, doi: 10.1016/S0022-2836(02)00852-5.

[82] J. N. Mavinahalli, A. Madhumalar, R. W. Beuerman, D. P. Lane, and C. Verma, “Differences in the transactivation domains of p53 family members: a computational study,” BMC Genomics, vol. 11, no. 1, p. S5, Feb. 2010, doi: 10.1186/1471-2164-11-S1-S5.

[83] N. Dumaz, D. M. Milne, and D. W. Meek, “Protein kinase CK1 is a p53-threonine 18 kinase which requires prior phosphorylation of serine 15,” FEBS Lett., vol. 463, no. 3, pp. 312–316, 1999, doi: 10.1016/S0014-5793(99)01647-6.

[84] R. S. Tibbetts et al., “A role for ATR in the DNA damage-induced phosphorylation of p53,” Genes Dev., vol. 13, no. 2, pp. 152–157, Jan. 1999.

[85] O. Kalid and N. Ben-Tal, “Study of MDM2 Binding to p53-Analogues: Affinity, Helicity, and Applicability to Drug Design,” J. Chem. Inf. Model., vol. 49, no. 4, pp. 865–876, Apr. 2009, doi: 10.1021/ci800352c.

[86] M. Matsumoto, M. Furihata, and Y. Ohtsuki, “Posttranslational phosphorylation of mutant p53 protein in tumor development,” Med. Mol. Morphol., vol. 39, no. 2, pp. 79–87, Jul. 2006, doi: 10.1007/s00795-006-0320-0.

[87] J. D. Oliner, J. A. Pietenpol, S. Thiagalingam, J. Gyuris, K. W. Kinzler, and B. Vogelstein, “Oncoprotein MDM2 conceals the activation domain of tumour suppressor p53,” Nature, vol. 362, no. 6423, Art. no. 6423, Apr. 1993, doi: 10.1038/362857a0.

[88] J. Chen, V. Marechal, and A. J. Levine, “Mapping of the p53 and mdm-2 interaction domains.,” Mol. Cell. Biol., vol. 13, no. 7, pp. 4107–4114, Jul. 1993, doi: 10.1128/MCB.13.7.4107.

[89] S. R. Grossman et al., “Polyubiquitination of p53 by a Ubiquitin Ligase Activity of p300,” Science, vol. 300, no. 5617, pp. 342–344, Apr. 2003, doi: 10.1126/science.1080386.

[90] C. L. Brooks and W. Gu, “p53 Ubiquitination: Mdm2 and Beyond,” Mol. Cell, vol. 21, no. 3, pp. 307–315, Feb. 2006, doi: 10.1016/j.molcel.2006.01.020.

[91] K. Sakaguchi, S. Saito, Y. Higashimoto, S. Roy, C. W. Anderson, and E. Appella, “Damage-mediated Phosphorylation of Human p53 Threonine 18 through a Cascade Mediated by a Casein 1-like Kinase EFFECT ON Mdm2 BINDING,” J. Biol. Chem., vol. 275, no. 13, pp. 9278–9283, Mar. 2000, doi: 10.1074/jbc.275.13.9278.

[92] C. W. Lee, J. C. Ferreon, A. C. M. Ferreon, M. Arai, and P. E. Wright, “Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation,” Proc. Natl. Acad. Sci., vol. 107, no. 45, pp. 19290–19295, Nov. 2010, doi: 10.1073/pnas.1013078107.

[93] P. F. Lambert, F. Kashanchi, M. F. Radonovich, R. Shiekhattar, and J. N. Brady, “Phosphorylation of p53 Serine 15 Increases Interaction with CBP,” J. Biol. Chem., vol. 273, no. 49, pp. 33048–33053, Dec. 1998, doi: 10.1074/jbc.273.49.33048.

[94] D. P. Teufel, M. Bycroft, and A. R. Fersht, “Regulation by phosphorylation of the relative affinities of the N-terminal transactivation domains of p53 for p300 domains and Mdm2,” Oncogene, vol. 28, no. 20, Art. no. 20, May 2009, doi: 10.1038/onc.2009.71.

[95] K. Sakaguchi et al., “DNA damage activates p53 through a phosphorylation–acetylation cascade,” Genes Dev., vol. 12, no. 18, pp. 2831–2841, Sep. 1998, doi: 10.1101/gad.12.18.2831.

[96] Y. Tang, W. Zhao, Y. Chen, Y. Zhao, and W. Gu, “Acetylation Is Indispensable for p53 Activation,” Cell, vol. 133, no. 4, pp. 612–626, May 2008, doi: 10.1016/j.cell.2008.03.025.

[97] S. Min, K. Kim, S.-G. Kim, H. Cho, and Y. Lee, “Chromatin-remodeling factor, RSF1, controls p53-mediated transcription in apoptosis upon DNA strand breaks,” Cell Death Dis., vol. 9, no. 11, Art. no. 11, Oct. 2018, doi: 10.1038/s41419-018-1128-2.

[98] D. Su et al., “Interactions of Chromatin Context, Binding Site Sequence Content, and Sequence Evolution in Stress-Induced p53 Occupancy and Transactivation,” PLOS Genet., vol. 11, no. 1, p. e1004885, Jan. 2015, doi: 10.1371/journal.pgen.1004885.

[99] A. Hafner, G. Lahav, and J. Stewart-Ornstein, “Stereotyped p53 binding tuned by chromatin accessibility,” bioRxiv, p. 177667, Aug. 2017, doi: 10.1101/177667.

[100] D. Ganguly and J. Chen, “Atomistic Details of the Disordered States of KID and pKID. Implications in Coupled Binding and Folding,” J. Am. Chem. Soc., vol. 131, no. 14, pp. 5214–5223, Apr. 2009, doi: 10.1021/ja808999m.

[101] A. Bah et al., “Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch,” Nature, vol. 519, no. 7541, Art. no. 7541, Mar. 2015, doi: 10.1038/nature13999.

[102] R. Grosely et al., “Effects of Phosphorylation on the Structure and Backbone Dynamics of the Intrinsically Disordered Connexin43 C-terminal Domain,” J. Biol. Chem., vol. 288, no. 34, pp. 24857–24870, Aug. 2013, doi: 10.1074/jbc.M113.454389.

[103] C. Liang, S. N. Savinov, J. Fejzo, S. J. Eyles, and J. Chen, “Modulation of Amyloid-β42 Conformation by Small Molecules Through Nonspecific Binding,” J. Chem. Theory Comput., vol. 15, no. 10, pp. 5169–5174, Oct. 2019, doi: 10.1021/acs.jctc.9b00599.

[104] M. J. Daniels et al., “Cyclized NDGA modifies dynamic α-synuclein monomers preventing aggregation and toxicity,” Sci. Rep., vol. 9, no. 1, Art. no. 1, Feb. 2019, doi: 10.1038/s41598-019-39480-z.

[105] D. I. Hammoudeh, A. V. Follis, E. V. Prochownik, and S. J. Metallo, “Multiple Independent Binding Sites for Small-Molecule Inhibitors on the Oncoprotein c-Myc,” J. Am. Chem. Soc., vol. 131, no. 21, pp. 7390–7401, Jun. 2009, doi: 10.1021/ja900616b.

[106] N. Krishnan et al., “Targeting the disordered C terminus of PTP1B with an allosteric inhibitor,” Nat. Chem. Biol., vol. 10, no. 7, Art. no. 7, Jul. 2014, doi: 10.1038/nchembio.1528.

[107] J. Chen, X. Liu, and J. Chen, "Targeting Intrinsically Disordered Proteins through Dynamic Interactions," Biomolecules, vol. 10, no. 5, Art. no. 5, May 2020, doi: 10.3390/biom10050743.

[108] M. Wells et al., "Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain," Proc. Natl. Acad. Sci., vol. 105, no. 15, pp. 5762–5767, Apr. 2008, doi: 10.1073/pnas.0801353105.

[109] H. Lee et al., "Local Structural Elements in the Mostly Unstructured Transcriptional Activation Domain of Human p53," J. Biol. Chem., vol. 275, no. 38, pp. 29426–29432, Sep. 2000, doi: 10.1074/jbc.M003107200.

[110] P. Vise, B. Baral, A. Stancik, D. F. Lowry, and G. W. Daughdrill, "Identifying long-range structure in the intrinsically unstructured transactivation domain of p53," Proteins Struct. Funct. Bioinforma., vol. 67, no. 3, pp. 526–530, 2007, doi: 10.1002/prot.21364.

[111] Y. A. Zhan, H. Wu, A. T. Powell, G. W. Daughdrill, and F. M. Ytreberg, "Impact of the K24N mutation on the transactivation domain of p53 and its binding to murine double-minute clone 2," Proteins Struct. Funct. Bioinforma., vol. 81, no. 10, pp. 1738–1747, 2013, doi: 10.1002/prot.24310.

[112] D. F. Lowry, A. Stancik, R. M. Shrestha, and G. W. Daughdrill, "Modeling the accessible conformations of the intrinsically unstructured transactivation domain of p53," Proteins Struct. Funct. Bioinforma., vol. 71, no. 2, pp. 587–598, 2008, doi: 10.1002/prot.21721.

[113] F. Huang et al., "Multiple conformations of full-length p53 detected with single-molecule fluorescence resonance energy transfer," Proc. Natl. Acad. Sci., vol. 106, no. 49, pp. 20758–20763, Dec. 2009, doi: 10.1073/pnas.0909644106.

[114] J. K. Lum, H. Neuweiler, and A. R. Fersht, "Long-Range Modulation of Chain Motions within the Intrinsically Disordered Transactivation Domain of Tumor Suppressor p53," J. Am. Chem. Soc., vol. 134, no. 3, pp. 1617–1622, Jan. 2012, doi: 10.1021/ja2078619.

[115] J. Huang et al., "CHARMM36m: an improved force field for folded and intrinsically disordered proteins," Nat. Methods, vol. 14, no. 1, Art. no. 1, Jan. 2017, doi: 10.1038/nmeth.4067.

[116] P. Robustelli, S. Piana, and D. E. Shaw, "Developing a molecular dynamics force field for both folded and disordered protein states," Proc. Natl. Acad. Sci., vol. 115, no. 21, pp. E4758–E4766, May 2018, doi: 10.1073/pnas.1800690115.

[117] K. H. Lee and J. Chen, "Optimization of the GBMV2 implicit solvent force field for accurate simulation of protein conformational equilibria," J. Comput. Chem., vol. 38, no. 16, pp. 1332–1341, 2017, doi: 10.1002/jcc.24734.

[118] H.-N. Wu, F. Jiang, and Y.-D. Wu, “Significantly Improved Protein Folding Thermodynamics Using a Dispersion-Corrected Water Model and a New Residue-Specific Force Field,” J. Phys. Chem. Lett., vol. 8, no. 14, pp. 3199–3205, Jul. 2017, doi: 10.1021/acs.jpclett.7b01213.

[119] M. J. Robertson, J. Tirado-Rives, and W. L. Jorgensen, “Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field,” J. Chem. Theory Comput., vol. 11, no. 7, pp. 3499–3509, Jul. 2015, doi: 10.1021/acs.jctc.5b00356.

[120] B. R. Brooks et al., “CHARMM: The biomolecular simulation program,” J. Comput. Chem., vol. 30, no. 10, pp. 1545–1614, 2009, doi: 10.1002/jcc.21287.

[121] P. Eastman et al., “OpenMM 4: A Reusable, Extensible, Hardware Independent Library for High Performance Molecular Simulation,” J. Chem. Theory Comput., vol. 9, no. 1, pp. 461–469, Jan. 2013, doi: 10.1021/ct300857j.

[122] M. J. Abraham et al., “GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers,” SoftwareX, vol. 1–2, pp. 19–25, Sep. 2015, doi: 10.1016/j.softx.2015.06.001.

[123] J. C. Phillips et al., “Scalable molecular dynamics with NAMD,” J. Comput. Chem., vol. 26, no. 16, pp. 1781–1802, 2005, doi: 10.1002/jcc.20289.

[124] A. W. Götz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand, and R. C. Walker, “Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born,” J. Chem. Theory Comput., vol. 8, no. 5, pp. 1542–1555, May 2012, doi: 10.1021/ct200909j.

[125] R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, and R. C. Walker, “Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald,” J. Chem. Theory Comput., vol. 9, no. 9, pp. 3878–3888, Sep. 2013, doi: 10.1021/ct400314y.

[126] D. A. Case et al., “The Amber biomolecular simulation programs,” J. Comput. Chem., vol. 26, no. 16, pp. 1668–1688, 2005, doi: 10.1002/jcc.20290.

[127] Y. I. Yang, Q. Shao, J. Zhang, L. Yang, and Y. Q. Gao, “Enhanced sampling in molecular dynamics,” J. Chem. Phys., vol. 151, no. 7, p. 070902, Aug. 2019, doi: 10.1063/1.5109531.

[128] R. C. Bernardi, M. C. R. Melo, and K. Schulten, “Enhanced sampling techniques in molecular dynamics simulations of biological systems,” Biochim. Biophys. Acta BBA - Gen. Subj., vol. 1850, no. 5, pp. 872–877, May 2015, doi: 10.1016/j.bbagen.2014.10.019.

[129] C. Abrams and G. Bussi, “Enhanced Sampling in Molecular Dynamics Using Metadynamics, Replica-Exchange, and Temperature-Acceleration,” Entropy, vol. 16, no. 1, Art. no. 1, Jan. 2014, doi: 10.3390/e16010163.

[130] W. Zhang and J. Chen, “Accelerate Sampling in Atomistic Energy Landscapes Using Topology-Based Coarse-Grained Models,” J. Chem. Theory Comput., vol. 10, no. 3, pp. 918–923, Mar. 2014, doi: 10.1021/ct500031v.

[131] K. Moritsugu, T. Terada, and A. Kidera, “Scalable free energy calculation of proteins via multiscale essential sampling,” J. Chem. Phys., vol. 133, no. 22, p. 224105, Dec. 2010, doi: 10.1063/1.3510519.

[132] P. Liu, B. Kim, R. A. Friesner, and B. J. Berne, “Replica exchange with solute tempering: A method for sampling biological systems in explicit water,” Proc. Natl. Acad. Sci., vol. 102, no. 39, pp. 13749–13754, Sep. 2005, doi: 10.1073/pnas.0506346102.

[133] A. Mittal, N. Lyle, T. S. Harmon, and R. V. Pappu, “Hamiltonian Switch Metropolis Monte Carlo Simulations for Improved Conformational Sampling of Intrinsically Disordered Regions Tethered to Ordered Domains of Proteins,” J. Chem. Theory Comput., vol. 10, no. 8, pp. 3550–3562, Aug. 2014, doi: 10.1021/ct5002297.

[134] E. Karl Peter and J.-E. Shea, “A hybrid MD-kMC algorithm for folding proteins in explicit solvent,” Phys. Chem. Chem. Phys., vol. 16, no. 14, pp. 6430–6440, 2014, doi: 10.1039/C3CP55251A.

[135] C. Zhang and J. Ma, “Enhanced sampling and applications in protein folding in explicit solvent,” J. Chem. Phys., vol. 132, no. 24, p. 244101, Jun. 2010, doi: 10.1063/1.3435332.

[136] L. Zheng and W. Yang, “Practically Efficient and Robust Free Energy Calculations: Double-Integration Orthogonal Space Tempering,” J. Chem. Theory Comput., vol. 8, no. 3, pp. 810–823, Mar. 2012, doi: 10.1021/ct200726v.

[137] A. Laio and M. Parrinello, “Escaping free-energy minima,” Proc. Natl. Acad. Sci., vol. 99, no. 20, pp. 12562–12566, Oct. 2002, doi: 10.1073/pnas.202427399.

[138] L. Wang, R. A. Friesner, and B. J. Berne, “Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2),” J. Phys. Chem. B, vol. 115, no. 30, pp. 9431–9438, Aug. 2011, doi: 10.1021/jp204407d.

[139] E. L. Kim et al., “Comparative Assessment of the Functional p53 Status in Glioma Cells,” Anticancer Res., vol. 25, no. 1A, pp. 213–224, Jan. 2005.

[140] Uchida Toyoaki et al., “p53 Mutations and Prognosis in Bladder Tumors,” J. Urol., vol. 153, no. 4, pp. 1097–1104, Apr. 1995, doi: 10.1016/S0022-5347(01)67517-7.

[141] A. Thirion, P. Rouanet, S. Thezenas, D. Detournay, J. Grenier, and E. Lopez-Crapez, “Interest of investigating p53 status in breast cancer by four different methods,” Oncol. Rep., vol. 9, no. 6, pp. 1167–1172, Nov. 2002, doi: 10.3892/or.9.6.1167.

[142] P. D. Vise, B. Baral, A. J. Latos, and G. W. Daughdrill, “NMR chemical shift and relaxation measurements provide evidence for the coupled folding and binding of the p53 transactivation domain,” Nucleic Acids Res., vol. 33, no. 7, pp. 2061–2077, Apr. 2005, doi: 10.1093/nar/gki336.

[143] W. Borcherds et al., “Disorder and residual helicity alter p53-Mdm2 binding affinity and signaling in cells,” Nat. Chem. Biol., vol. 10, no. 12, Art. no. 12, Dec. 2014, doi: 10.1038/nchembio.1668.

[144] K. R. Andersen, N. C. Leksa, and T. U. Schwartz, “Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification,” Proteins Struct. Funct. Bioinforma., vol. 81, no. 11, pp. 1857–1861, 2013, doi: 10.1002/prot.24364.

[145] R. N. De Guzman, H. Y. Liu, M. Martinez-Yamout, H. J. Dyson, and P. E. Wright, “Solution structure of the TAZ2 (CH3) domain of the transcriptional adaptor protein CBP11Edited by M. F. Summers,” J. Mol. Biol., vol. 303, no. 2, pp. 243–253, Oct. 2000, doi: 10.1006/jmbi.2000.4141.

[146] R. N. De Guzman, M. A. Martinez-Yamout, H. J. Dyson, and P. E. Wright, “Structure and Function of the CBP/p300 TAZ Domains,” in Zinc Finger Proteins: From Atomic Contact to Cellular Function, S. Iuchi and N. Kuldell, Eds. Boston, MA: Springer US, 2005, pp. 114–120.

[147] D. P. Teufel, S. M. Freund, M. Bycroft, and A. R. Fersht, “Four domains of p300 each bind tightly to a sequence spanning both transactivation subdomains of p53,” Proc. Natl. Acad. Sci., vol. 104, no. 17, pp. 7009–7014, Apr. 2007, doi: 10.1073/pnas.0702010104.

[148] C. W. Lee, M. A. Martinez-Yamout, H. J. Dyson, and P. E. Wright, “Structure of the p53 Transactivation Domain in Complex with the Nuclear Receptor Coactivator Binding Domain of CREB Binding Protein,” Biochemistry, vol. 49, no. 46, pp. 9964–9971, Nov. 2010, doi: 10.1021/bi1012996.

[149] P. Di Lello et al., “Structure of the Tfb1/p53 Complex: Insights into the Interaction between the p62/Tfb1 Subunit of TFIIH and the Activation Domain of p53,” Mol. Cell, vol. 22, no. 6, pp. 731–740, Jun. 2006, doi: 10.1016/j.molcel.2006.05.007.

[150] P. I. Hanson and S. W. Whiteheart, “AAA+ proteins: have engine, will work,” Nat. Rev. Mol. Cell Biol., vol. 6, no. 7, Art. no. 7, Jul. 2005, doi: 10.1038/nrm1684.

[151] A. F. Neuwald, L. Aravind, J. L. Spouge, and E. V. Koonin, “AAA+: A Class of Chaperone-Like ATPases Associated with the Assembly, Operation, and Disassembly of Protein Complexes,” Genome Res., vol. 9, no. 1, pp. 27–43, Jan. 1999, doi: 10.1101/gr.9.1.27.

[152] C. Puchades, C. R. Sandate, and G. C. Lander, “The molecular principles governing the activity and functional diversity of AAA+ proteins,” Nat. Rev. Mol. Cell Biol., vol. 21, no. 1, Art. no. 1, Jan. 2020, doi: 10.1038/s41580-019-0183-6.

[153] D. G. Murdock, B. E. Boone, L. A. Esposito, and D. C. Wallace, “Up-regulation of nuclear and mitochondrial genes in the skeletal muscle of mice lacking the heart/muscle isoform of the adenine nucleotide translocator,” J. Biol. Chem., vol. 274, no. 20, pp. 14429–14433, May 1999, doi: 10.1074/jbc.274.20.14429.

[154] E. Pronicka et al., “New perspective in diagnostics of mitochondrial disorders: two years’ experience with whole-exome sequencing at a national paediatric centre,” J. Transl. Med., vol. 14, no. 1, p. 174, Jun. 2016, doi: 10.1186/s12967-016-0930-9.

[155] S. E. Calvo, K. R. Clauser, and V. K. Mootha, “MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins,” Nucleic Acids Res., vol. 44, no. D1, pp. D1251–D1257, Jan. 2016, doi: 10.1093/nar/gkv1003.

[156] D. J. Pagliarini et al., “A Mitochondrial Protein Compendium Elucidates Complex I Disease Biology,” Cell, vol. 134, no. 1, pp. 112–123, Jul. 2008, doi: 10.1016/j.cell.2008.06.016.

[157] V. Hung et al., “Proteomic Mapping of the Human Mitochondrial Intermembrane Space in Live Cells via Ratiometric APEX Tagging,” Mol. Cell, vol. 55, no. 2, pp. 332–341, Jul. 2014, doi: 10.1016/j.molcel.2014.06.003.

[158] H.-W. Rhee et al., “Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging,” Science, vol. 339, no. 6125, pp. 1328–1331, Mar. 2013, doi: 10.1126/science.1230593.

[159] M. Cardenas-Rodriguez and K. Tokatlidis, “Cytosolic redox components regulate protein homeostasis via additional localisation in the mitochondrial intermembrane space,” FEBS Lett., vol. 591, no. 17, pp. 2661–2670, 2017, doi: 10.1002/1873-3468.12766.

[160] M. Habich, S. L. Salscheider, and J. Riemer, “Cysteine residues in mitochondrial intermembrane space proteins: more than just import,” Br. J. Pharmacol., vol. 176, no. 4, pp. 514–531, 2019, doi: 10.1111/bph.14480.

[161] J. M. Herrmann and J. Riemer, “The Intermembrane Space of Mitochondria,” Antioxid. Redox Signal., vol. 13, no. 9, pp. 1341–1358, Apr. 2010, doi: 10.1089/ars.2009.3063.

[162] E. Nuebel, P. Manganas, and K. Tokatlidis, “Orphan proteins of unknown function in the mitochondrial intermembrane space proteome: New pathways and metabolic cross-talk,” Biochim. Biophys. Acta BBA - Mol. Cell Res., vol. 1863, no. 11, pp. 2613–2623, Nov. 2016, doi: 10.1016/j.bbamcr.2016.07.004.

[163] S. Saita et al., “PARL mediates Smac proteolytic maturation in mitochondria to promote apoptosis,” Nat. Cell Biol., vol. 19, no. 4, Art. no. 4, Apr. 2017, doi: 10.1038/ncb3488.

[164] M. Spinazzi et al., “PARL deficiency in mouse causes Complex III defects, coenzyme Q depletion, and Leigh-like syndrome,” Proc. Natl. Acad. Sci., vol. 116, no. 1, pp. 277–286, Jan. 2019, doi: 10.1073/pnas.1811938116.

[165] R. B. Hill and L. Pellegrini, “The PARL family of mitochondrial rhomboid proteases,” Semin. Cell Dev. Biol., vol. 21, no. 6, pp. 582–592, Aug. 2010, doi: 10.1016/j.semcdb.2009.12.011.

[166] C. Petrungaro and J. Riemer, “Mechanisms and physiological impact of the dual localization of mitochondrial intermembrane space proteins,” Biochem. Soc. Trans., vol. 42, no. 4, pp. 952–958, Aug. 2014, doi: 10.1042/BST20140104.

[167] S. Saita, T. Tatsuta, P. A. Lampe, T. König, Y. Ohba, and T. Langer, “PARL partitions the lipid transfer protein STARD7 between the cytosol and mitochondria,” EMBO J., vol. 37, no. 4, p. e97909, Feb. 2018, doi: 10.15252/embj.201797909.

[168] P. Hänzelmann, C. Galgenmüller, and H. Schindelin, “Structure and Function of the AAA+ ATPase p97, a Key Player in Protein Homeostasis,” in Macromolecular Protein Complexes II: Structure and Function, J. R. Harris and J. Marles-Wright, Eds. Cham: Springer International Publishing, 2019, pp. 221–272.

[169] C.-M. Ho et al., “Malaria parasite translocon structure and mechanism of effector export,” Nature, vol. 561, no. 7721, Art. no. 7721, Sep. 2018, doi: 10.1038/s41586-018-0469-4.

[170] M. Opalińska and H. Jańska, “AAA Proteases: Guardians of Mitochondrial Function and Homeostasis,” Cells, vol. 7, no. 10, Art. no. 10, Oct. 2018, doi: 10.3390/cells7100163.

[171] M. Prattes, Y.-H. Lo, H. Bergler, and R. E. Stanley, “Shaping the Nascent Ribosome: AAA-ATPases in Eukaryotic Ribosome Biogenesis,” Biomolecules, vol. 9, no. 11, Art. no. 11, Nov. 2019, doi: 10.3390/biom9110715.

[172] T. E. Steele and S. E. Glynn, “Mitochondrial AAA proteases: A stairway to degradation,” Mitochondrion, vol. 49, pp. 121–127, Nov. 2019, doi: 10.1016/j.mito.2019.07.012.

[173] S. Zhang and Y. Mao, “AAA+ ATPases in Protein Degradation: Structures, Functions and Mechanisms,” Biomolecules, vol. 10, no. 4, Art. no. 4, Apr. 2020, doi: 10.3390/biom10040629.

[174] T. Haslberger, B. Bukau, and A. Mogk, “Towards a unifying mechanism for ClpB/Hsp104-mediated protein disaggregation and prion propagationThis paper is one of a selection of papers published in this special issue entitled 8th International Conference on AAA Proteins and has undergone the Journal’s usual peer review process.,” Biochem. Cell Biol., vol. 88, no. 1, pp. 63–75, Jan. 2010, doi: 10.1139/O09-118.

[175] M. Zolkiewski, T. Zhang, and M. Nagy, “Aggregate reactivation mediated by the Hsp100 chaperones,” Arch. Biochem. Biophys., vol. 520, no. 1, pp. 1–6, Apr. 2012, doi: 10.1016/j.abb.2012.01.012.

[176] J. Li, A. Mahajan, and M.-D. Tsai, “Ankyrin Repeat: A Unique Motif Mediating Protein−Protein Interactions,” Biochemistry, vol. 45, no. 51, pp. 15168–15178, Dec. 2006, doi: 10.1021/bi062188q.

[177] S. G. Sedgwick and S. J. Smerdon, “The ankyrin repeat: a diversity of interactions on a common structural framework,” Trends Biochem. Sci., vol. 24, no. 8, pp. 311–316, Aug. 1999, doi: 10.1016/S0968-0004(99)01426-7.

[178] M. E. Barnett and M. Zolkiewski, “Site-Directed Mutagenesis of Conserved Charged Amino Acid Residues in ClpB from Escherichia coli,” Biochemistry, vol. 41, no. 37, pp. 11277–11283, Sep. 2002, doi: 10.1021/bi026161s.

[179] M. Koppen and T. Langer, “Protein Degradation within Mitochondria: Versatile Activities of AAA Proteases and Other Peptidases,” Crit. Rev. Biochem. Mol. Biol., vol. 42, no. 3, pp. 221–242, Jan. 2007, doi: 10.1080/10409230701380452.

[180] N. Wiedemann and N. Pfanner, “Mitochondrial Machineries for Protein Import and Assembly,” Annu. Rev. Biochem., vol. 86, no. 1, pp. 685–714, 2017, doi: 10.1146/annurev-biochem-060815-014352.

[181] P. Bross and P. Fernandez-Guerra, “Disease-Associated Mutations in the HSPD1 Gene Encoding the Large Subunit of the Mitochondrial HSP60/HSP10 Chaperonin Complex,” Front. Mol. Biosci., vol. 3, 2016, doi: 10.3389/fmolb.2016.00049.

[182] P. Banerjee, J. B. Joo, J. T. Buse, and G. Dawson, “Differential solubilization of lipids along with membrane proteins by different classes of detergents,” Chem. Phys. Lipids, vol. 77, no. 1, pp. 65–78, Aug. 1995, doi: 10.1016/0009-3084(95)02455-R.

[183] V. G. Nadeau, A. Rath, and C. M. Deber, “Sequence Hydropathy Dominates Membrane Protein Response to Detergent Solubilization,” Biochemistry, vol. 51, no. 31, pp. 6228–6237, Aug. 2012, doi: 10.1021/bi201853n.

[184] V. Reisinger and L. A. Eichacker, “Isolation of Membrane Protein Complexes by Blue Native Electrophoresis,” in 2D PAGE: Sample Preparation and Fractionation, A. Posch, Ed. Totowa, NJ: Humana Press, 2008, pp. 423–431.

[185] H. Schägger, “Chapter 13 Blue-native gels to isolate protein complexes from mitochondria,” in Methods in Cell Biology, vol. 65, Academic Press, 2001, pp. 231–244.

[186] H. SCHÄGER and G. VON JAGOW, “Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form,” Blue Native Electrophor. Isol. Membr. Protein Complexes Enzym. Act. Form, vol. 199, no. 2, pp. 223–231, 1991.

[187] E. Nicolas, R. Tricarico, M. Savage, E. A. Golemis, and M. J. Hall, "Disease-Associated Genetic Variation in Human Mitochondrial Protein Import," Am. J. Hum. Genet., vol. 104, no. 5, pp. 784–801, May 2019, doi: 10.1016/j.ajhg.2019.03.019.

[188] J. Vamecq, B. Papegay, V. Nuyens, J. Boogaerts, O. Leo, and V. Kruys, "Mitochondrial dysfunction, AMPK activation and peroxisomal metabolism: A coherent scenario for non-canonical 3-methylglutaconic acidurias," Biochimie, vol. 168, pp. 53–82, Jan. 2020, doi: 10.1016/j.biochi.2019.10.004.

[189] D. Mróz et al., "CLPB (caseinolytic peptidase B homolog), the first mitochondrial protein refoldase associated with human disease," Biochim. Biophys. Acta Gen. Subj., vol. 1864, no. 4, p. 129512, 2020, doi: 10.1016/j.bbagen.2020.129512.

[190] A. J. Erives and J. S. Fassler, "Metabolic and Chaperone Gene Loss Marks the Origin of Animals: Evidence for Hsp104 and Hsp78 Chaperones Sharing Mitochondrial Enzymes as Clients," PLOS ONE, vol. 10, no. 2, p. e0117192, Feb. 2015, doi: 10.1371/journal.pone.0117192.

[191] E. Fernández-Vizarra, G. Ferrín, A. Pérez-Martos, P. Fernández-Silva, M. Zeviani, and J. A. Enríquez, "Isolation of mitochondria for biogenetical studies: An update," Mitochondrion, vol. 10, no. 3, pp. 253–262, Apr. 2010, doi: 10.1016/j.mito.2009.12.148.

[192] A. Roy, A. Kucukural, and Y. Zhang, "I-TASSER: a unified platform for automated protein structure and function prediction," Nat. Protoc., vol. 5, no. 4, Art. no. 4, Apr. 2010, doi: 10.1038/nprot.2010.5.

[193] Y. Zhang, "I-TASSER server for protein 3D structure prediction," BMC Bioinformatics, vol. 9, no. 1, p. 40, Jan. 2008, doi: 10.1186/1471-2105-9-40.

[194] P. Kumar, S. Henikoff, and P. C. Ng, "Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm," Nat. Protoc., vol. 4, no. 7, Art. no. 7, Jul. 2009, doi: 10.1038/nprot.2009.86.

[195] I. A. Adzhubei et al., "A method and server for predicting damaging missense mutations," Nat. Methods, vol. 7, no. 4, Art. no. 4, Apr. 2010, doi: 10.1038/nmeth0410-248.

[196] P. Bragoszewski et al., "Retro-translocation of mitochondrial intermembrane space proteins," Proc. Natl. Acad. Sci., vol. 112, no. 25, pp. 7713–7718, Jun. 2015, doi: 10.1073/pnas.1504615112.

[197] A. B. Harbauer, R. P. Zahedi, A. Sickmann, N. Pfanner, and C. Meisinger, "The Protein Import Machinery of Mitochondria—A Regulatory Hub in Metabolism, Stress, and Disease," Cell Metab., vol. 19, no. 3, pp. 357–372, Mar. 2014, doi: 10.1016/j.cmet.2014.01.010.

[198] M. J. Baker, C. S. Palmer, and D. Stojanovski, “Mitochondrial protein quality control in health and disease,” Br. J. Pharmacol., vol. 171, no. 8, pp. 1870–1889, 2014, doi: 10.1111/bph.12430.

[199] D. Poveda-Huertes et al., “An Early mtUPR: Redistribution of the Nuclear Transcription Factor Rox1 to Mitochondria Protects against Intramitochondrial Proteotoxic Aggregates,” Mol. Cell, vol. 77, no. 1, pp. 180-188.e9, Jan. 2020, doi: 10.1016/j.molcel.2019.09.026.